Some new tests for normality based on U–processes *

Miguel A. Arcones †

Department of Mathematical Sciences, Binghamton University, Binghamton, NY 13902

Yishi Wang ‡

Department of Mathematical Sciences, Binghamton University, Binghamton, NY 13902

June 14, 2005

Abstract

We present two new tests for normality based on U–processes. These tests improve on the Lilliefors tests. We obtain the consistency and asymptotic null distribution of these tests. We present simulations of the power of these tests and of several classical tests for some fixed alternatives. These simulations show that the presented tests are competitive.

1 Introduction

Many statistical procedures assume that the observations are normally distributed. Hence, testing for normality should be done before using many statistical analysis. A very complete review of normality tests is in Thode (2002). The classical normality tests are the ones by Lilliefors (1967) and Shapiro and Wilk (1965, 1968). Shapiro, Wilk and Chen (1968) show that the Lilliefors test is not as powerful as the Shapiro–Wilk test. Several authors have provided different approaches to test normality. Mardia (1980) reviews tests for normality based on skewness and kurtosis. We present two modifications of Lilliefors test. The proposed tests are based on the Lévy characterization of the normal distribution. Csörgő, Seshadri, and Yalovsky (1973) and Nguyen and Dinh (2003) discuss several tests of normality based on other characterizations of the normal distribution. We will estimate the Lévy characterization of the normal distribution using the distribution function. Del Barrio, Cuesta–Albertos, and

* *Key words and phrases:* test of normality, U–processes.

† E-mail: arcones@math.binghamton.edu

‡ E-mail: wang@math.binghamton.edu

Matrán (2000) review goodness of fit tests using empirical processes. Our test statistics are U–processes. Several authors have considered normality tests based on U–statistics (see Csörgő, 1986; Epps and Pulley, 1983; and Henze and Zirkler, 1990).

We study the testing problem

$$H_0 : F \text{ has a normal distribution, versus } H_1 : F \text{ does not,} \tag{1.1}$$

based on a random sample $X_1, \dots, X_n$ of size n from F. Lilliefors (1968) proposed the test which rejects the null hypothesis if $L_n \geq a_{n,\alpha}$, where

$$L_n := \sup_{t \in \mathbb{R}} |F_n(t) - \Phi(s_n^{-1}(t - \bar{X}_n))|, \tag{1.2}$$

where F_n is the empirical cdf, Φ is the cdf of the standard normal distribution, $\bar{X}_n := \frac{1}{n}\sum_{=1}^{n} X_j$, $s_n^2 := \frac{1}{n-1}\sum_{j=1}^{n}(X_j - \bar{X}_n)^2$,

$$a_{n,\alpha} = \inf\{\lambda \geq 0 : \mathbb{P}_\Phi\{L_n < \lambda\} \geq 1 - \alpha\}, \tag{1.3}$$

and $\mathbb{P}_\Phi$ is the probability distribution for which the data follows a standard normal distribution. Since the distribution of L_n is a location and scale invariant, we have that for each μ and each $\sigma > 0$, $\mathbb{P}_{\Phi_{\mu,\sigma}}\{L_n \geq a_{n,\alpha}\} \leq \alpha$, where $\Phi_{\mu,\sigma}$ is the cdf of a normal r.v. with mean μ and variance σ^2. Hence, the type I error of the test is less or equal than α.

We present two normality tests which improve on the Lilliefors test. By the Lévy characterization of the normal distribution (see e.g. Theorem 20.2.A in Loève, 1977), given $m \geq 1$ and cdf F, $m^{-1/2}\sigma_F^{-1}\sum_{j=1}^{m}(X_j - \mu_F)$ has a standard normal distribution, where $X_1, \dots, X_m$ are i.i.d.r.v.'s with cdf F, mean μ_F and variance σ_F^2, if and only if F has a normal distribution. Hence, F has a normal distribution if and only if for some $m \geq 1$,

$$D_m(F) := \sup_{t \in \mathbb{R}} |\mathbb{P}_F\{\sigma_F^{-1} m^{-1/2} \sum_{j=1}^{m}(X_j - \mu_F) \leq t\} - \Phi(t)| = 0, \tag{1.4}$$

where $\mathbb{P}_F$ is the probability for which the i.i.d.r.v.'s $X_1, \dots, X_m$ have distribution F.

Let

$$D_{n,m} := \sup_{t \in \mathbb{R}} \left| \frac{(n-m)!}{n!} \sum_{(i_1,\dots,i_m) \in I_m^n} I(\hat{\sigma}_{n,m}^{-1} m^{-1/2} \sum_{j=1}^{m}(X_{i_j} - \bar{X}_n) \leq t) - \Phi(t) \right|, \tag{1.5}$$

where $I_m^n = \{(i_1, \dots, i_m) \in \mathbb{N}^m : 1 \leq i_j \leq n, i_j \neq i_k \text{ if } j \neq k\}$ and

$$\hat{\sigma}_{n,m}^2 := \frac{(n-m)!}{n!} \sum_{(i_1,\dots,i_m) \in I_m^n} m^{-1}(\sum_{j=1}^{m}(X_{i_j} - \bar{X}_n))^2.$$

Note that

$$
\begin{aligned}
\hat\sigma^2_{n,m} &:= \tfrac{(n-m)!}{n!}\textstyle\sum_{(i_1,\dots,i_m)\in I^n_m} m^{-1}(\sum_{j=1}^m (X_{i_j}-\bar X_n))^2 \qquad (1.6)\\
&= \tfrac{(n-m)!}{mn!}\textstyle\sum_{(i_1,\dots,i_m)\in I^n_m}\left(\sum_{j=1}^m (X_{i_j}-\bar X_n)^2+\sum_{1\le j\ne k\le m}(X_{i_j}-\bar X_n)(X_{i_k}-\bar X_n)\right)\\
&= \tfrac{1}{n}\textstyle\sum_{j=1}^n (X_j-\bar X_n)^2+\tfrac{m-1}{n(n-1)}\sum_{1\le j\ne k\le n}(X_j-\bar X_n)(X_k-\bar X_n)\\
&= \tfrac{1}{n}\textstyle\sum_{j=1}^n (X_j-\bar X_n)^2+\tfrac{m-1}{n(n-1)}\sum_{j,k=1}^n (X_j-\bar X_n)(X_k-\bar X_n)-\tfrac{m-1}{n(n-1)}\sum_{j=1}^n (X_j-\bar X_n)^2\\
&= \tfrac{n-m}{n(n-1)}\textstyle\sum_{j=1}^n (X_j-\bar X_n)^2.
\end{aligned}
$$

When $m=1$, the statistics L_n and $D_{n,m}$ are very similar. The difference is that the standardization is different: $\hat\sigma^2_{n,1}=\frac{1}{n}\sum_{j=1}^n (X_j-\bar X_n)^2\ne s_n^2$. The empirical cdf used in the Lilliefors test has jumps of size n^{-1}. The empirical cdf in (1.5) has jumps of size $(n-m)!/n!$. The proposed test statistic uses an empirical cdf which is smoother than the usual empirical cdf.

Given $1>\alpha>0$, let

$$b_{n,m,\alpha}=\inf\{\lambda\ge 0:\mathbb{P}_\Phi\{D_{n,m}<\lambda\}\ge 1-\alpha\}.$$

Then, the proposed test rejects the null hypothesis if $D_n\ge b_{n,m,\alpha}$. It is easy to see that the distribution of $D_{n,m}$ is invariant by changes of location and scale. So, the distribution of $D_{n,m}$ is the same for all normal distributions. Hence, the probability of type error I of the test is less or equal than α. As noted by Henze and Zirkler (1990), the Kolmogorov–Smirnov test is only invariant in the one dimensional situation. The same happens for the proposed tests. Using the proposed method, it is not possible to use the multivariate cdf and get invariant test statistics for multivariate data.

An alternative expression for the statistic in (1.5) is

$$\max_{1\le j\le k_n}\left(\max(|k_n^{-1}j-\Phi(U_{(j)})|,|k_n^{-1}(j-1)-\Phi(U_{(j)})|)\right),$$

where $k_n=\binom{n}{m}$ and $U_{(1)}\le\cdots\le U_{(k_n)}$ are the order values $\hat\sigma^{-1}_{n,m}m^{-1/2}\sum_{j=1}^m (X_{i_j}-\bar X_n)$.

A variation in the previous characterization of the normal distribution is as follows. Given $m\ge 2$ and a nondegenerate cdf F, $m^{-1/2}\sum_{j=1}^m (X_j-E[X_j])$ and $X_1-\mu$ have the same distribution, where $X_1,\dots,X_m$ are i.i.d.r.v.'s with cdf F, if and only if F has a normal distribution. Hence, F has a normal distribution if and only if for some $m\ge 2$,

$$\tilde D_m(F):=\sup_{t\in\mathbb{R}}|\mathbb{P}_F\{m^{-1/2}\sum_{j=1}^m (X_j-\mu_F)\le t\}-\mathbb{P}_F\{X_1-\mu_F\le t\}|=0. \tag{1.7}$$

An estimator of the previous quantity is

$$\tilde D_{n,m}:=\sup_{t\in\mathbb{R}}\left|\frac{(n-m)!}{n!}\sum_{(i_1,\dots,i_m)\in I^n_m} I(m^{-1/2}\sum_{j=1}^m (X_{i_j}-\bar X_n)\le t)-n^{-1}\sum_{j=1}^n I(X_j-\bar X_n\le t)\right| \tag{1.8}$$

It is easy to see that the distribution of $\tilde{D}_{n,m}$ is invariant by changes of location and scale. Given $1 > \alpha > 0$, let

$$c_{n,\alpha} = \inf\{\lambda \geq 0 : \mathbb{P}_\Phi\{\tilde{D}_{n,m} < \lambda\} \geq 1-\alpha\}.$$

Then, the test rejects the null hypothesis if $\tilde{D}_{n,m} \geq c_{n,m,\alpha}$.

Section 2 contains the main results. In Section 3, we present some simulations. Section 4 contains the proofs of the theorems in Section 2.

2 Main results

The statistics $D_{n,m}$ and $\tilde{D}_{n,m}$ are based on a collection of U–statistics. Given r.v.'s $X_1, \ldots, X_n$ with values in a measurable space $(S, \mathcal{S})$ and a measurable function $h : (S^m, \mathcal{S}^m) \to \mathbb{R}$, the U–statistic with kernel h is defined by

$$U_{n,m}(h) := \frac{(n-m)!}{n!} \sum_{(i_1,\ldots,i_m)\in I_m^n} h(X_{i_1}, \ldots, X_{i_m}).$$

General references on U–statistics are Lee (1990) and de la Peña and Giné (1999). The main property of U–statistics which we will use is the Hoeffding decomposition, which we describe next. Suppose that $X_1, \ldots, X_n$ are i.i.d.r.v.'s. We define

$$\pi_{k,m}h(x_1, \ldots, x_k) = (\delta_{x_1} - P)\cdots(\delta_{x_m} - P)P^{m-k}h,$$

where $Q_1 \cdots Q_m h = \int \cdots \int h(x_1, \ldots, x_m)\, dQ_1(x_1)\cdots dQ_m(x_m)$ and δ_x denotes the Dirac measure at x. Then, it is known that

$$U_{n,m}(h) = \sum_{k=0}^{m} \binom{m}{k} U_{n,k}(\pi_{k,m}h).$$

Observe that $\pi_{k,m}f$ is a function of k variables. If $E[(h(X_1, \ldots, X_m))^2] < \infty$, then

$$n^{1/2}|U_{n,m}(h) - E[h(X_1, \ldots, X_m)] - mU_{n,1}(\pi_{1,m}h)| \xrightarrow{\Pr} 0.$$

This allows to reduce the asymptotics of a U–statistic to the asymptotics of a sum of i.i.d.r.v.'s. We will need to work with a class of U–statistics. Given a class of functions $\mathcal{F}$ from S^m into $\mathbb{R}$, the U–process indexed by $\mathcal{F}$ is

$$U_{n,m}(h) := \frac{(n-m)!}{n!} \sum_{(i_1,\ldots,i_m)\in I_m^n} h(X_{i_1}, \ldots, X_{i_m}),\ h \in \mathcal{F}.$$

Limit theorems for U-processes were developed by Arcones and Giné (1993) and de la Peña and Giné (1999). The statistics $D_{n,m}$ and $\tilde{D}_{n,m}$ are the supremum of a U-process.

Sometimes, we will take the expectation over some of the variables present. If $X_1, X_2, \dots, X_m$ are independent r.v.'s, then $E_{X_{i_1}, X_{i_2}, \dots, X_{i_k}}[h(X_1, X_2, \dots, X_m)]$ denotes the expectation with respect to the variables $X_{i_1}, X_{i_2}, \dots, X_{i_k}$.

The first proposed test is consistent. Precisely, we have the following theorem:

Theorem 2.1. *Suppose that $\{X_j\}_{j=1}^{\infty}$ is a sequence of i.i.d.r.v.'s from a continuous cdf F with finite second moment. Then, $D_{n,m} \overset{a.s.}{\to} D_m(F)$, as $n \to \infty$.*

This implies that $b_{n,m,\alpha} \to 0$, as $n \to \infty$, and that if F does not have a normal distribution, then, for each $1 > \alpha > 0$, $\mathbb{P}_F\{D_{n,m} \geq b_{n,m,\alpha}\} \to 1$, as $n \to \infty$.

Next theorem gives the asymptotic null distribution of the first test.

Theorem 2.2. *Suppose that $\{X_j\}_{j=1}^{\infty}$ is a sequence of i.i.d.r.v.'s from a normal cdf with mean μ and variance $\sigma^2 > 0$. Then,*

$$n^{1/2} D_{n,m} - \sup_{t \in \mathbb{R}} |n^{-1/2} \sum_{j=1}^{n} (g(\sigma^{-1}(X_j - \mu), t) - E[g(\sigma^{-1}(X_j - \mu), t)])| \overset{\Pr}{\to} 0.$$

where

$$g(x,t) = m\Phi((m-1)^{-1/2}(m^{1/2}t - x)) + (m^{1/2}x + 2^{-1}(x^2 - 1)t)\phi(t),$$

and ϕ is the pdf of a standard normal distribution.

Consequently,

$$n^{1/2} D_{n,m} \overset{\mathrm{d}}{\to} \sup_{t \in \mathbb{R}} |U(t)|,$$

where $\{U(t) : t \in \mathbb{R}\}$ is a Gaussian process with zero means and covariance given by

$$E[U(s)U(t)] = \mathrm{Cov}(g(Z_1, s), g(Z_1, t)), s, t \in \mathbb{R},$$

and Z_1 is a standard normal r.v.

For the second proposed test, we have asymptotics similar to those of the first test:

Theorem 2.3. *Suppose that $\{X_j\}_{j=1}^{\infty}$ is a sequence of i.i.d.r.v.'s from a continuous cdf F with finite first moment. Then, $\tilde{D}_{n,m} \overset{a.s.}{\to} \tilde{D}_m(F)$, as $n \to \infty$.*

Theorem 2.4. *Suppose that $\{X_j\}_{j=1}^{\infty}$ is a sequence of i.i.d.r.v.'s from a normal cdf with mean μ and variance $\sigma^2 > 0$. Then,*

$$n^{1/2} \tilde{D}_{n,m} - \sup_{t \in \mathbb{R}} |n^{-1/2} \sum_{j=1}^{n} (h(\sigma^{-1}(X_j - \mu), t) - E[h(\sigma^{-1}(X_j - \mu), t)])| \overset{\Pr}{\to} 0.$$

where

$$h(x,t) = m\Phi((m-1)^{-1/2}(m^{1/2}t - x)) - I(x \leq t) + (m^{1/2} - 1)x\phi(t).$$

Consequently, $n^{1/2} \tilde{D}_{n,m} \overset{\mathrm{w}}{\to} \sup_{s \in \mathbb{R}} |V(s)|$, where $\{V(s) : s \in \mathbb{R}\}$ is a mean zero Gaussian process with covariance given by

$$E[V(s)V(t)] = \mathrm{Cov}(h(Z_1, s), h(Z_1, t)), s, t \in \mathbb{R}.$$

3 Simulations.

Table 1 shows the values of $b_{n,m,\alpha}$ and $c_{n,m,\alpha}$, for some values of n and m. Table 1 was obtained by doing 10000 simulations from a standard normal distribution.

TABLE 1

$b_{n,2,\alpha}$	$\alpha = 0.10$	$\alpha = 0.05$	$\alpha = 0.01$
$n = 10$	0.127767	0.146294	0.1881883
$n = 15$	0.09380524	0.10899795	0.14135184
$n = 20$	0.07639177	0.08738504	0.11396659
$n = 30$	0.05665779	0.06585072	0.08662073
$n = 50$	0.04040402	0.04640940	0.05884147

$b_{n,3,\alpha}$	$\alpha = 0.10$	$\alpha = 0.05$	$\alpha = 0.01$
$n = 10$	0.08116030	0.09628613	0.12998585
$n = 15$	0.05554883	0.06750267	0.09593262
$n = 20$	0.04399126	0.05336427	0.07580943
$n = 30$	0.03307612	0.04010797	0.05588256
$n = 50$	0.02454210	0.02908864	0.04048928

$c_{n,2,\alpha}$	$\alpha = 0.10$	$\alpha = 0.05$	$\alpha = 0.01$
$n = 10$	0.2333333	0.2444444	0.2777778
$n = 15$	0.1809524	0.2000000	0.2285714
$n = 20$	0.1578947	0.1710526	0.1973684
$n = 30$	0.1287356	0.1390805	0.1655172
$n = 50$	0.0983673	0.1065306	0.1240816

$c_{n,3,\alpha}$	$\alpha = 0.10$	$\alpha = 0.05$	$\alpha = 0.01$
$n = 10$	0.2500000	0.2750000	0.3250000
$n = 15$	0.1934066	0.2124542	0.2439560
$n = 20$	0.1649123	0.1789474	0.2070175
$n = 30$	0.1354680	0.1460591	0.1690476
$n = 50$	0.1034184	0.1116837	0.1287245

We will compare our tests with several other normality tests. Table 2 shows the power, when $\alpha = 0.05$, of the tests in Lilliefors (1967), Shapiro and Wilk (1965), Csörgő (1986), Epps-Pulley (1983) (BHEP test), and the tests in (1.5) and in (1.8) for several alternatives. The columns $D_{n,2}$, $D_{n,3}$, $\tilde{D}_{n,2}$ and $\tilde{D}_{n,3}$ in Table 2 were obtained by doing 10000 simulations from the test statistics using the mentioned alternatives. These columns represent the proportions of simulations which are bigger than the corresponding cutpoints from Table 1. To find the column L, first we found $a_{n,0.05}$ in (1.3) using 10000 simulations from a standard normal distribution. Then, the proportion of times such that the statistic L_n is bigger than the cutpoint was obtained. Let $X_{(1)},\dots,X_{n)}$ be the order statistics from a random sample with a standard normal distribution. Let $m=(m_1,\dots,m_n)'$ and let $V=(v_{i,j})_{1\le i,j\le n}$, where $m_i=E[X_{(j)}]$ and $v_{i,j}=\mathrm{Cov}(X_{(i)},X_{(j)})$. The Shapiro-Wilks test is significative for small values of

$$W:=(\sum_{j=1}^n(X_j-\bar{X}_n)^2)^{-1}(\sum_{j=1}^n a_jX_j)^2, \tag{3.1}$$

where

$$a'=(a_1,\dots,a_n)=(m'V^{-2}M)^{-1/2}m'V^{-1}. \tag{3.2}$$

To find the column SW, we found W using the vector a in Table 5 in Shapiro and Wilk (1965). The column represents the proportions of simulations of W which are smaller than the cutpoint from the Table 6 (level 0.05) in Shapiro and Wilk (1965). The test in Csörgő (1986) is significative for nonnormality for large values of the statistic

$$\begin{aligned}&\sup_{|t|\le T}\left||\varphi_n(t)|^2-|\varphi(t)|^2\right| \qquad (3.3)\\ =&\sup_{|t|\le T}\left|\left(n^{-1}\sum_{j=1}^n\cos(t\hat{\sigma}_n^{-1}X_j)\right)^2+\left(n^{-1}\sum_{j=1}^n\sin(t\hat{\sigma}_n^{-1}X_j)\right)^2-\exp(-t^2)\right|\\ =&\sup_{|t|\le T}\left|n^{-2}\sum_{j,k=1}^n\left(\cos(t\hat{\sigma}_n^{-1}X_j)\sin(t\hat{\sigma}_n^{-1}X_k)+\sin(t\hat{\sigma}_n^{-1}X_j)\sin(t\hat{\sigma}_n^{-1}X_k)\right)-\exp(-t^2)\right|,\end{aligned}$$

where $\varphi_n(t):=n^{-1}\sum_{j=1}^n\exp(it\hat{\sigma}_n^{-1}(X_j-\bar{X}_n))$ is an empirical ch.f., $\varphi(t)=\exp(-2^{-1}t^2)$ is the ch.f. of a standard normal r.v., $\bar{X}_n=n^{-1}\sum_{j=1}^n X_n$ and $\hat{\sigma}_n^2=n^{-1}\sum_{j=1}^n(X_j-\bar{X}_n)^2$. Since the supremum over a continuous is incalculable, as recommended in Csörgő (1986), we use in the simulations

$$\sup_{-10^2\le j\le 10^2}|\varphi_n((1.47)10^{-2}j)|^2-|\varphi((1.47)10^{-2}j)|^2|. \tag{3.4}$$

The $BHEP$ (Epps and Pulley, 1983; and Baringhaus and Henze, 1988) normality test is significative for nonnormality for large values of the statistic

$$\begin{aligned}&\int_{\mathbb{R}}|n^{-1}\textstyle\sum_{j=1}^n\exp(it\hat{\sigma}_n^{-1}(X_j-\bar{X}_n))-\exp(-2^{-1}t^2)|^2\phi(t)\,dt, \qquad (3.5)\\ =&\; n^{-2}\textstyle\sum_{j,k=1}^n\exp(-2^{-1}\delta^2\hat{\sigma}_n^{-2}(X_j-X_k)^2)\\ &-2^{1/2}n^{-1}\textstyle\sum_{j=1}^n\exp(-2^{-2}\hat{\sigma}_n^{-1}(X_j-\bar{X}_n))^2)+3^{-1/2}.\end{aligned}$$

The column $BHEP$ was found first estimating the cutpoint of the test by doing simulations.

TABLE 2

	L	SW	CS	$BHEP$	$D_{n,2}$ test	$D_{n,3}$ test	$\tilde{D}_{n,2}$ test	$\tilde{D}_{n,3}$ test
$n = 10$	0.2990	0.4150	0.3183	0.4150	0.3688	0.3709	0.2166	0.1949
$n = 15$	0.4630	0.6834	0.4923	0.6289	0.5650	0.5477	0.2688	0.2973
$n = 20$	0.5818	0.8468	0.5778	0.7961	0.7088	0.6975	0.3430	0.3730
$n = 30$	0.7867	0.9646	0.7274	0.9536	0.8898	0.8753	0.4783	0.5452
$n = 50$	0.9644	0.9994	0.8920	0.9960	0.9897	0.9891	0.7580	0.7485

Alternative: exponential distribution

	L	SW	CS	$BHEP$	$D_{n,2}$ test	$D_{n,3}$ test	$\tilde{D}_{n,2}$ test	$\tilde{D}_{n,3}$ test
$n = 10$	0.1416	0.1478	0.1860	0.1627	0.1781	0.1766	0.0505	0.0579
$n = 15$	0.1884	0.2113	0.2849	0.2058	0.2284	0.2371	0.0367	0.0665
$n = 20$	0.2200	0.2665	0.3592	0.2855	0.3037	0.3030	0.0332	0.0808
$n = 30$	0.2861	0.3239	0.4729	0.3557	0.4027	0.3860	0.0401	0.1233
$n = 50$	0.4464	0.3925	0.6658	0.5152	0.5643	0.5313	0.0813	0.0823

Alternative: double exponential distribution

	L	SW	CS	$BHEP$	$D_{n,2}$ test	$D_{n,3}$ test	$\tilde{D}_{n,2}$ test	$\tilde{D}_{n,3}$ test
$n = 10$	0.5823	0.5870	0.6085	0.5956	0.6044	0.6100	0.3525	0.3617
$n = 15$	0.7408	0.7594	0.8032	0.7627	0.7846	0.7826	0.4330	0.5299
$n = 20$	0.8437	0.8703	0.9016	0.8894	0.8826	0.8787	0.4896	0.6326
$n = 30$	0.8437	0.9515	0.9724	0.9603	0.9666	0.9615	0.6138	0.7927
$n = 50$	0.9933	0.9925	0.9983	0.9964	0.9969	0.9970	0.8020	0.8060

Alternative: Cauchy distribution

	L	SW	CS	$BHEP$	$D_{n,2}$ test	$D_{n,3}$ test	$\tilde{D}_{n,2}$ test	$\tilde{D}_{n,3}$ test
$n = 10$	0.0935	0.1230	0.0589	0.1075	0.0906	0.0931	0.1375	0.1158
$n = 15$	0.1384	0.2038	0.0682	0.1785	0.1242	0.0982	0.1591	0.1560
$n = 20$	0.1748	0.3080	0.0665	0.2795	0.1756	0.1250	0.1912	0.1995
$n = 30$	0.2720	0.5218	0.0526	0.8514	0.2556	0.1746	0.2334	0.2465
$n = 50$	0.4526	0.8820	0.0647	0.7101	0.4707	0.3238	0.3521	0.3920

Alternative: Beta$(2, 1)$ distribution

	L	SW	CS	$BHEP$	$D_{n,2}$ test	$D_{n,3}$ test	$\tilde{D}_{n,2}$ test	$\tilde{D}_{n,3}$ test
$n = 10$	0.0709	0.0759	0.0937	0.0825	0.0872	0.0935	0.0058	0.0446
$n = 15$	0.0726	0.1006	0.1328	0.0962	0.1016	0.1156	0.0453	0.0340
$n = 20$	0.0818	0.1170	0.1586	0.1383	0.1288	0.1396	0.0400	0.0475
$n = 30$	0.0956	0.1234	0.1953	0.1334	0.1548	0.1776	0.0355	0.0450
$n = 50$	0.1115	0.1305	0.2826	0.1589	0.2037	0.2243	0.0425	0.0404

Alternative: Logistic(1) distribution

	L	SW	CS	$BHEP$	$D_{n,2}$ test	$D_{n,3}$ test	$\tilde{D}_{n,2}$ test	$\tilde{D}_{n,3}$ test
$n = 10$	0.0642	0.0809	0.0142	0.0480	0.0418	0.0413	0.1601	0.1131
$n = 15$	0.0813	0.1208	0.0073	0.0783	0.0395	0.0204	0.1801	0.1585
$n = 20$	0.0989	0.2052	0.0028	0.1546	0.0437	0.0160	0.2269	0.1841
$n = 30$	0.1444	0.4100	0.0018	0.2503	0.0419	0.0120	0.2650	0.2488
$n = 50$	0.2618	0.8580	0.3142	0.5410	0.0631	0.0078	0.4173	0.4050
	Alternative: uniform distribution							

We get from Table 2 that there is not too much difference between the tests $D_{n,2}$ and $D_{n,3}$. This seems to indicate that there is not much gain by choosing a value of m, different from $m = 2$. The Lilliefors test is the less powerful of all the considered tests. The Csörgő test is very uneven. For some distributions, it is the most powerful test, but for other distributions does not have too much power. The tests $\tilde{D}_{n,2}$ and $\tilde{D}_{n,3}$ often do not rank between the most powerful tests. The three best tests are the Shapiro–Wilks test, the BHEP test and the $D_{n,2}$ tests. These tests are comparable. In the case of a uniform alternative, we get that the power of $D_{n,2}$ and $D_{n,3}$ tests is less than 0.05. This means that these tests are not unbiased. The test in Csörgő (1986) is not unbiased either.

4 Proofs

Proof of Theorem 2.1. Given $t \in \mathbb{R}$, define $h_t(x_1, \dots, x_m) = t - m^{-1/2}\sigma_F^{-1}\sum_{j=1}^m (X_j - \mu_F)$. Since the class of functions $\{h_t : t \in \mathbb{R}\}$ is a one dimensional affine space of functions, the class of sets

$$\{\{(x_1, \dots, x_m) \in \mathbb{R}^m : m^{-1/2}\sigma_F^{-1}\sum_{j=1}^m (x_j - \mu_F) \le t\} : t \in \mathbb{R}\}$$

is a VC class (see Theorem 4.2.1 in Dudley, 1999). By the LLN for U-processes indexed by a VC class (see e.g. Corollary 3.3 in Arcones and Giné, 1993),

$$\sup_{s\in\mathbb{R}} \left| \tfrac{(n-m)!}{n!} \textstyle\sum_{(i_1,\dots,i_m)\in I_m^n} I(m^{-1/2}\sigma_F^{-1}\sum_{j=1}^m (X_{i_j} - \mu_F) \le s) - H_{F,m}(s)\right| \stackrel{a.s.}{\to} 0.$$

where

$$H_{F,m}(s) := \mathbb{P}_F\{m^{-1/2}\sigma_F^{-1}\sum_{j=1}^m (X_j - \mu_F) \le s\}.$$

Taking $s = \sigma_F^{-1}m^{1/2}(\bar{X}_n - \mu_F) + \sigma_F^{-1}\hat{\sigma}_{n,m}t$, we get that

$$\begin{aligned}\sup_{t\in\mathbb{R}} \Big| &\tfrac{(n-m)!}{n!} \textstyle\sum_{(i_1,\dots,i_m)\in I_m^n} I(m^{-1/2}\hat{\sigma}_{n,m}^{-1}\sum_{j=1}^m (X_{i_j} - \bar{X}_n) \le t) \\ &-H_{F,m}(\sigma_F^{-1}m^{1/2}(\bar{X}_n - \mu_F) + \sigma_F^{-1}\hat{\sigma}_{n,m}t)\Big| \stackrel{a.s.}{\to} 0.\end{aligned} \tag{4.1}$$

Since $H_{F,m}(\cdot)$, is a uniformly continuous function, we have that

$$\sup_{t\in\mathbb{R}} \left|H_{F,m}(\sigma_F^{-1}m^{1/2}(\bar{X}_n-\mu_F)+\sigma_F^{-1}\hat{\sigma}_{n,m}t)-H_{F,m}(t)\right| \overset{a.s.}{\to} 0. \tag{4.2}$$

Note that by the strong law of the large numbers, $\bar{X}_n \overset{a.s.}{\to} \mu$ and $\hat{\sigma}^2_{n,m} \overset{a.s.}{\to} \sigma^2$. The claim follows from (4.1) and (4.2). $\square$

We will need the following lemma:

Lemma 4.1. *Let $\{X_j\}$ be a sequence of i.i.d.r.v.'s from a standard normal distribution. Then, for each $m \geq 2$,*

$$\begin{aligned}&\sup\nolimits_{t\in\mathbb{R}} |n^{-1/2}\textstyle\sum_{j=1}^n(\Phi((m-1)^{-1/2}(m^{1/2}(m^{1/2}\bar{X}_n+\hat{\sigma}_{n,m}t)-X_j))\\ &-\Phi((m-1)^{-1/2}(m^{1/2}t-X_j))-\Phi(m^{1/2}\bar{X}_n+\hat{\sigma}_{n,m}t)+\Phi(t))| \overset{\Pr}{\to} 0.\end{aligned}$$

Proof. Since $\Phi((m-1)^{-1/2}(m^{1/2}t-x))$ is increasing in t, the class of functions, $\{\Phi((m-1)^{-1/2}(m^{1/2}t-x)) : t\in\mathbb{R}\}$ is a VC subgraph class of functions. Hence, by the central limit theorem for VC subgraph classes (see e.g. Corollary 6.3.16 in Dudley, 1999), $\{U_n(t) : t\in\mathbb{R}\}$ converges weakly, where

$$U_n(t) := n^{-1/2}\sum_{j=1}^n \left(\Phi((m-1)^{-1/2}(m^{1/2}t-X_j))-\Phi(t)\right).$$

Observe that

$$\begin{aligned}&E[\Phi((m-1)^{-1/2}(m^{1/2}t-X_1))] = E[I(X_2+\cdots+X_m \leq m^{1/2}t-X_1)]\\ =\ &E[I(m^{-1/2}(X_1+\cdots+X_m)\leq t)] = \Phi(t).\end{aligned}$$

The weak convergence of the process $\{U_n(t) : t\in\mathbb{R}\}$ implies that for each $\tau>0$,

$$\lim_{\delta\to 0}\limsup_{n\to\infty}\mathbb{P}\{\sup_{d(t,s)\leq\delta}|U_n(s)-U_n(t)|\geq\tau\}=0,$$

where

$$d^2(s,t)=\mathrm{Var}(\Phi((m-1)^{-1/2}(m^{1/2}s-X_1))-\Phi((m-1)^{-1/2}(m^{1/2}t-X_1)))$$

(see Theorem 3.7.2 in Dudley, 1999). By the Cauchy-Schwartz inequality, for each $s,t\in R$,

$$\begin{aligned}&d^2(s,t)=\mathrm{Var}(\Phi((m-1)^{-1/2}(m^{1/2}s-X_1))-\Phi((m-1)^{-1/2}(m^{1/2}t-X_1)))\\ \leq\ &E[(\Phi((m-1)^{-1/2}(m^{1/2}s-X_1))-\Phi((m-1)^{-1/2}(m^{1/2}t-X_1)))^2]\\ =\ &E_{X_1}[(E_{X_2,\ldots,X_m}[I(m^{-1/2}\textstyle\sum_{j=1}^m X_j\leq s)-I(m^{-1/2}\sum_{j=1}^m X_j\leq t)])^2]\\ \leq\ &E_{X_1}[E_{X_2,\ldots,X_m}[(I(m^{-1/2}\textstyle\sum_{j=1}^m X_j\leq s)-I(m^{-1/2}\sum_{j=1}^m X_j\leq t))^2]]=|\Phi(t)-\Phi(s)|.\end{aligned}$$

Using the previous estimation, we get that for each $0<M<\infty$,

$$\sup_{t\in\mathbb{R}}\sup_{|a|,|b|\leq M} d^2(an^{-1/2}+(1+bn^{-1/2})t,t)\leq\sup_{t\in\mathbb{R}}\sup_{|a|,|b|\leq M}|\Phi(an^{-1/2}+(1+bn^{-1/2})t)-\Phi(t)|\to 0,$$

as $n \to \infty$. Hence, for each $0 < M < \infty$,

$$\sup_{t\in\mathbb{R}} \sup_{|a|,|b|\le M} |U_n(an^{-1/2} + (1+bn^{-1/2})t) - U_n(t)| \xrightarrow{\Pr} 0. \tag{4.3}$$

Plugging $(m^{1/2}n^{1/2}\bar{X}_n, n^{1/2}(\hat{\sigma}_{n,m}-1))$ as (a,b) in (4.3), the result follows. $\square$

Proof of Theorem 2.2. Since the distribution of the test is invariant by changes of location and scale, we can assume that the r.v.'s are standard normal. By the central limit theorem for U-processes indexed by a VC class (see Theorem 4.9 in Arcones and Giné, 1993),

$$\{n^{1/2}\frac{(n-m)!}{n!} \sum_{(i_1,\ldots,i_m)\in I_m^n} (I(m^{-1/2}\sum_{j=1}^m X_{i_j} \le t) - \Phi(t)) : t\in\mathbb{R}\}$$

converges weakly. By Corollary 4.2 in Arcones and Giné (1993), the U-process is asymptotically equivalent to the first term in its Hoeffding decomposition:

$$\begin{aligned} n^{1/2}\sup_{t\in\mathbb{R}}|\frac{(n-m)!}{n!} \sum_{(i_1,\ldots,i_m)\in I_m^n} (I(m^{-1/2}\sum_{j=1}^m X_{i_j} \le t) - \Phi(t)) \\ -mn^{-1}\sum_{j=1}^n(\Phi((m-1)^{-1/2}(m^{1/2}t - X_j)) - \Phi(t))| \xrightarrow{\Pr} 0. \end{aligned} \tag{4.4}$$

Observe that

$$\begin{aligned} & E_{X_2,\ldots,X_m}[I(m^{-1/2}\textstyle\sum_{j=1}^m X_j \le t)] = E_{X_2,\ldots,X_m}[I(\sum_{j=2}^m X_j \le m^{1/2}t - X_1)] \\ = & \Phi((m-1)^{-1/2}(m^{1/2}t - X_1)). \end{aligned}$$

Plugging $m^{1/2}\bar{X}_n + \hat{\sigma}_{n,m}t$ into t in (4.4), we get that

$$\begin{aligned} & n^{1/2}\sup_{t\in\mathbb{R}}|\tfrac{(n-m)!}{n!}\textstyle\sum_{(i_1,\ldots,i_m)\in I_m^n}(I(m^{-1/2}\sum_{j=1}^m(X_{i_j}-\bar{X}_n) \le \hat{\sigma}_{n,m}t) - \Phi(m^{1/2}\bar{X}_n + \hat{\sigma}_{n,m}t)) \\ & -mn^{-1}\textstyle\sum_{j=1}^n(\Phi((m-1)^{-1/2}(m^{1/2}(m^{1/2}\bar{X}_n + \hat{\sigma}_{n,m}t) - X_j) - \Phi(m^{1/2}\bar{X}_n + \hat{\sigma}_{n,m}t))| \xrightarrow{\Pr} 0. \end{aligned}$$

From the previous limit and Lemma 4.1, we get that

$$\begin{aligned} n^{1/2}\sup_{t\in\mathbb{R}}|\tfrac{(n-m)!}{n!}\textstyle\sum_{(i_1,\ldots,i_m)\in I_m^n}(I(m^{-1/2}\sum_{j=1}^m(X_{i_j}-\bar{X}_n) \le \hat{\sigma}_{n,m}t) - \Phi(m^{1/2}\bar{X}_n + \hat{\sigma}_{n,m}t)) \\ -mn^{-1}\textstyle\sum_{j=1}^n(\Phi((m-1)^{-1/2}(m^{1/2}t - X_j)) - \Phi(t))| \xrightarrow{\Pr} 0. \end{aligned} \tag{4.5}$$

Hence, we have that

$$\begin{aligned} 0 \xleftarrow{\Pr} \; & n^{1/2}D_{n,m} - \sup_{t\in\mathbb{R}}|mn^{-1/2}\textstyle\sum_{j=1}^n \Phi((m-1)^{-1/2}(m^{1/2}t - X_j)) - \Phi(t)) \\ & \qquad + n^{1/2}(\Phi(m^{1/2}\bar{X}_n + \hat{\sigma}_{n,m}t) - \Phi(t))| \\ = \; & n^{1/2}D_{n,m} - \sup_{t\in\mathbb{R}}|mn^{-1/2}\textstyle\sum_{j=1}^n \Phi((m-1)^{-1/2}(m^{1/2}t - X_j)) - \Phi(t)) \\ & \qquad + n^{1/2}(m^{1/2}\bar{X}_n + (\hat{\sigma}_{n,m}-1)t)\phi(t)| + o_P(1). \end{aligned}$$

Finally, the claim follows noticing that

$$
\begin{aligned}
& n^{1/2}(\hat{\sigma}_{n,m}-1) = (\hat{\sigma}_{n,m}+1)^{-1}n^{1/2}(\hat{\sigma}^2_{n,m}-1) = (\hat{\sigma}_{n,m}+1)^{-1}n^{1/2}(\tfrac{n-m}{n(n-1)}\textstyle\sum_{j=1}^n (X_j-\bar{X})^2-1) \\
= \ & (\hat{\sigma}_{n,m}+1)^{-1}n^{1/2}\left((n-1)^{-1}(n-m)\left(n^{-1}\textstyle\sum_{j=1}^n X_j^2-(\bar{X}_n)^2\right)-1\right) \\
= \ & 2^{-1}n^{-1/2}\textstyle\sum_{j=1}^n (X_j^2-1)+o_P(1).
\end{aligned}
$$

□

The proof of Theorem 2.3 is similar to that of Theorem 2.1 and it is omitted.

Lemma 4.2. *Let $\{X_j\}$ be a sequence of i.id.r.v.'s from a standard normal distribution. Then,*

$$n^{1/2}\sup_{t\in\mathbb{R}}|n^{-1}\textstyle\sum_{j=1}^n \left(I(X_j-\bar{X}_n\le t)-I(X_j\le t)-\Phi(\bar{X}_n+t)+\Phi(t)\right)| \xrightarrow{\Pr} 0. \tag{4.6}$$

Proof. We proceed as in Lemma 4.1. By the Donsker theorem, $\{Y_n(t):t\in\mathbb{R}\}$ converges weakly, where

$$Y_n(t) := n^{-1/2}\sum_{j=1}^n \left(I(X_j\le t)-\Phi(t)\right).$$

This implies that for each $\tau>0$,

$$\lim_{\delta\to 0}\limsup_{n\to\infty}\mathbb{P}\{\sup_{d(t,s)\le\delta}|Y_n(s)-Y_n(t)|\ge\tau\}=0,$$

where

$$d^2(s,t)=\mathrm{Var}(I(X_1\le s),I(X_1\le t))\le|\Phi(s)-\Phi(t)|.$$

It is easy to see that for each $0<M<\infty$,

$$\sup_{t\in\mathbb{R}}\sup_{|s|\le M} d^2(t+n^{-1/2}s,t)\to 0,$$

as $n\to\infty$. Hence, for each $0<M<\infty$,

$$\sup_{t\in\mathbb{R}}\sup_{|s|\le M}|Y_n(t+n^{-1/2}s)-Y_n(t)| \xrightarrow{\Pr} 0. \tag{4.7}$$

The claim follows plugging $n^{1/2}\bar{X}_n$ as s in (4.7). □

Proof of Theorem 2.4. As before, we may assume that the r.v.'s are standard normal. The arguments leading to (4.5) but using $\hat{\sigma}_{m,n}=1$ give that

$$
\begin{aligned}
n^{1/2}\sup_{t\in\mathbb{R}}|\tfrac{(n-m)!}{n!}\textstyle\sum_{(i_1,\ldots,i_m)\in I_m^n}(I(m^{-1/2}\sum_{j=1}^m(X_{i_j}-\bar{X}_n)\le t)-\Phi(m^{1/2}\bar{X}_n+t)) \\
-mn^{-1}\textstyle\sum_{j=1}^n(\Phi((m-1)^{-1/2}(m^{1/2}t-X_j))-\Phi(t))| \xrightarrow{\Pr} 0.
\end{aligned}
\tag{4.8}
$$

By the previous limit and Lemma 4.2

$$\begin{array}{rl} 0 \stackrel{\Pr}{\leftarrow} & n^{1/2}\tilde{D}_{n,m} - n^{1/2}\sup_{t\in\mathbb{R}}\Big|mn^{-1}\sum_{j=1}^{n}(\Phi((m-1)^{-1/2}(m^{1/2}t - X_j)) - \Phi(t)) \\ & -n^{-1}\sum_{j=1}^{n}(I(X_j \le t) - \Phi(t)) + \Phi(m^{1/2}\bar{X}_n + t) - \Phi(\bar{X}_n + t)\Big| \\ = & n^{1/2}\tilde{D}_{n,m} - n^{1/2}\sup_{t\in\mathbb{R}}\Big|mn^{-1}\sum_{j=1}^{n}(\Phi((m-1)^{-1/2}(m^{1/2}t - X_j)) - \Phi(t)) \\ & -n^{-1}\sum_{j=1}^{n}(I(X_j \le t) - \Phi(t)) + (m^{1/2}-1)\bar{X}_n\phi(t)\Big| + o_P(1). \end{array}$$

□

References

[1] Arcones, M.A. and Giné, E. (1993). Limit theorems for U–processes. *Ann. Probab.* **21** 1494–1542.

[2] Baringhaus, L. and Henze, N. (1988). A consistent test for multivariate normality based on the empirical characteristic function. *Metrika* **35** 339–348.

[3] Csörgő, S. (1986). Testing for normality in arbitrary dimension. *Ann. Statist.* **14** 708–723.

[4] Csörgő, M.; Seshadri, V. and Yalovsky, M. (1973). Some exact tests for normality in the presence of unknown parameters. *J. Roy. Statist. Soc. B* **35** 507-522.

[5] de la Peña, V. H. and Giné, E. (1999). *Decoupling. From Dependence to Independence.* Springer–Verlag, New York.

[6] del Barrio, E.; Cuesta–Albertos, J. A. and Matrán, C. (2000). Contributions of empirical and quantile processes to the asymptotic theory of goodness–of–fit tests. With discussion. *Test* **9** 1–96.

[7] Dudley, R. M. (1999). *Uniform Central Limit Theorems.* Cambridge University Press, Cambridge.

[8] Epps, T. W. and Pulley, L. B. (1983). A test for normality based on the empirical characteristic function. *Biometrika* **70** 723–726.

[9] Henze, N. and Zirkler, B. (1990). A class of invariant consistent tests for multivariate normality. *Comm. Statist. Theory Methods* **19** 3595–3617.

[10] Lee, A. J. (1990). *U–statistics, Theory and Practice.* Marcel Dekker, New York.

[11] Lilliefors, H. (1967). On the Kolmogorov–Smirnov test for normality with mean and variance unknown. *J. Amer. Statist. Assoc.* **62** 399-402.

[12] Loève, M. (1977). *Probability Theory I*, 4–th Edition. Springer, New York.

[13] Mardia, K. (1980). Tests of univariate and multivariate normality. Volume 1 of Handbook of Statistics (ed. P.R. Krishnaiah), pp. 297-320. Amsterdam: North Holland.

[14] Nguyen, T. T. and Dinh, K. T. (2003). Characterizations of normal distributions and EDF goodness-of-fit tests. *Metrika* **58** 149–157.

[15] Shapiro, S. S. and Wilk, M. B. (1965). An analysis of variance test for normality: complete samples. *Biomet.* **52** 591–611.

[16] Shapiro, S. S. and Wilk, M. B. (1968). Approximations for the null distribution of the W statistic. *Technom.* **10** 861–866.

[17] Shapiro, S. S.; Wilk, M. B. and Chen, H. J. (1968). A comparative study of various tests for normality. *J. Amer. Statist. Assoc.* **63** 1343–1372.

[18] Thode, H. C., Jr. (2002). *Testing for Normality.* Marcel Dekker, New York.