TESTS OF GOODNESS OF FIT BASED ON THE L_2-WASSERSTEIN DISTANCE *

E. del Barrio; C. Matrán; J. Rodríguez-Rodríguez
Departamento de Estadística e Investigación Operativa,
Universidad de Valladolid, Spain

J.A. Cuesta-Albertos
Departamento de Matemáticas, Estadística y Computación,
Universidad de Cantabria, Spain

Abstract

We consider the Wasserstein distance between a sample distribution and the set of normal distributions as a measure of non-normality. By considering the standardized version of this distance we obtain a version of Shapiro-Wilk's test of normality. The asymptotic behaviour of the associated statistic is studied through approximations to the quantile process by Brownian bridges. This method differs from the available "ad hoc" method by de Wet and Venter and permits a similar analysis for testing fit to location and scale families.

Key words and phrases: Wasserstein distance, correlation test, Shapiro-Wilk, goodness of fit, test of normality, quantile process, Brownian bridge, convergence of integrals.

A.M.S. 1991 subject classification: Primary: 62F05, 62E20. Secondary: 60F25.

1 Introduction

Goodness of fit tests are often based on some distance between distribution functions (d.f.'s) or between probability laws (p.l.'s). In this work we follow this methodology, through the L_2-Wasserstein distance, by analyzing the distance between a fixed distribution and a location and scale family of probability distributions in $\mathbb{R}$. We focus on the (more interesting) normal case, but our approach can be used to cover different distribution types (we refer to [1] for details).

Let $\mathcal{P}_2(\mathbb{R})$ be the set of probabilities on the line with finite second order moment. For probabilities P_1 and P_2 in $\mathcal{P}_2(\mathbb{R})$ the L_2-Wasserstein distance between P_1 and P_2 is defined as the lowest L_2-distance between random variables (r.v.'s), defined in any probability space, with these distribution laws:

$$\mathcal{W}(P_1, P_2) := \inf \left\{ \left[E \left(X_1 - X_2 \right)^2 \right]^{1/2}, \mathcal{L}(X_1) = P_1, \mathcal{L}(X_2) = P_2 \right\}.$$

*Research partially supported by DGICYT, grants PB95-0715-c02-00, 01, 02 and PAPIJCL VA08/97

A fact that makes $\mathcal{W}$ useful in statistics on the line is that it can be explicitely obtained in terms of quantile functions. If P_1 and P_2 have distribution functions F_1 and F_2 and quantile functions F_1^{-1} and F_2^{-1}, then (see e.g. [2])

$$\mathcal{W}(P_1, P_2) = \left[\int_0^1 \left(F_1^{-1}(t) - F_2^{-1}(t) \right)^2 dt \right]^{1/2} \tag{1}$$

(recall that F^{-1} is defined on $(0,1)$ by $F^{-1}(t) = \inf\{s : F(s) \geq t\}$ and verifies that its distribution function is F when considered as a r.v. defined on the unit interval).

To ease the notation we will often identify a probability law with its distribution function (d.f.). In particular Φ will denote the d.f. of the standard normal law and ϕ will denote its density function, while $\mathcal{H}_\mathcal{N} := \{H : H(x) = \Phi\left(\frac{x-\mu}{\sigma}\right), \mu \in \mathbb{R}, \sigma > 0\}$ will be the set of normal laws on the line.

Now, observe that if $P \in \mathcal{P}_2(\mathbb{R})$ has d.f. F, mean μ_0 and standard deviation σ_0, then

$$\begin{aligned} \mathcal{W}^2(P, \mathcal{H}_\mathcal{N}) &:= \inf\{\mathcal{W}^2(P, H),\ H \in \mathcal{H}_\mathcal{N}\} = \inf_{\sigma>0} \left\{ \int_0^1 \left(F^{-1}(t) - \mu_0 - \sigma\Phi^{-1}(t) \right)^2 dt \right\} \\ &= \sigma_0^2 - \left(\int_0^1 (F^{-1}(t) - \mu_0)\Phi^{-1}(t) dt \right)^2 = \sigma_0^2 - \left(\int_0^1 F^{-1}(t)\Phi^{-1}(t) dt \right)^2. \end{aligned} \tag{2}$$

Thus, the normal law closest to P is given by $\mu = \mu_0$, and $\sigma = \int_0^1 F^{-1}(t)\Phi^{-1}(t)dt$. Note also that the ratio $\mathcal{W}^2(P, \mathcal{H}_\mathcal{N})/\sigma_0^2$ is not affected by location or scale changes on P. Hence, it can be considered as a measure of non-normality.

Now let $X_1, X_2, ..., X_n$ be a simple random sample with underlying d.f. F and let F_n denote the associated sample d.f.. It is natural to try to employ the sample version based on F_n and the sample variance S_n^2:

$$\mathcal{R}_n := \frac{\mathcal{W}^2(P_n, \mathcal{H}_\mathcal{N})}{S_n^2} = 1 - \frac{\left(\int_0^1 F_n^{-1}(t)\Phi^{-1}(t)dt \right)^2}{S_n^2} \tag{3}$$

to test the hypothesis of normality. In fact $\mathcal{R}_n$ is connected with the so-called correlation tests, whose interest is largely motivated by Shapiro-Wilk's test of normality [13]. This has been noted, in the context of normal probability plots, looking for the best choice for the plotting positions, by Brown and Hettmansperger in [3].

Additional relevant literature concerning the family of statistics related to Shapiro-Wilk's W statistic includes (or can be obtained in) [9], [11], [14], [12], but the only direct proof of the asymptotic behaviour of any of these statistics is that one in [9].

The purpose of this paper is to analyze the asymptotic behaviour of $\mathcal{R}_n$ through approximations of quantile processes by Brownian bridges, $B(t)$. This approach was also used in [4] to obtain the law of a simplified version, but the proof depends heavily on the previous results in [9] because it requires to give a sense to the limit expression

$$Z := \int_0^1 \frac{B^2(t) - EB^2(t)}{(\phi(\Phi^{-1}(t)))^2} dt. \tag{4}$$

This difficulty will be circumvented in Theorem 4, where we will show that, although the set of trajectories of a Brownian bridge $B(t)$ for which the function $t \mapsto \frac{B^2(t)-EB^2(t)}{(\phi(\Phi^{-1}(t)))^2}$ is integrable has zero probability, the sequence

$$\left\{ \int_{\frac{1}{n}}^{1-\frac{1}{n}} \frac{B^2(t) - EB^2(t)}{(\phi(\Phi^{-1}(t)))^2} dt \right\}_n$$

is an L_2-Cauchy sequence, whence we can give an adequate sense to Z as an L_2-limit.

We should remark that the ambitious program on the convergence of integrals of empirical and quantile processes developed in [6], [7] and [8] does not cover our results.

We would also like to notice that, with the present approach, we have not only been able to obtain the asymptotic distributions of those statistics belonging to the Shapiro-Wilk family under normality, but we have also found (see [1] for details) limit laws in a more general setup, including heavier tailed distributions, which had not been previously reported in the literature concerning the correlation tests (see [12]).

2 The results

The normal law closest to the sample d.f. F_n, based on $X_1, X_2, ...X_n$, is given by the sample mean, $\hat{\mu}_n = \mu(F_n) = \bar{X}_n$, and (denoting the ordered statistic by $X_{kn}, k = 1, 2, ...n$)

$$\hat{\sigma}_n = \sum_{k=1}^{n} X_{kn} \int_{\frac{k-1}{n}}^{\frac{k}{n}} \Phi^{-1}(t)dt.$$

Our measure of non-normality is then

$$\mathcal{R}_n = \frac{\mathcal{W}^2(P_n, \mathcal{H}_\mathcal{N})}{S_n^2} = 1 - \frac{\hat{\sigma}_n^2}{S_n^2},$$

where $S_n^2 = \frac{1}{n}\sum_{i=1}^n (X_i - \bar{X}_n)^2$ is the variance of the sample distribution. We will study $\mathcal{R}_n$ under the hypothesis $F(x) = \Phi\left(\frac{x-\mu_0}{\sigma_0}\right)$.

The invariance of $\mathcal{R}_n$ with respect to location or scale changes, allows us to assume $F = \Phi$ and, by the convergence $S_n^2 \to \sigma^2(\Phi) = 1$ a.s., we can study the asymptotic behaviour of $\mathcal{R}_n$ through that of $S_n^2\mathcal{R}_n$ which, in turns, admits the following decomposition

$$\begin{aligned} 0 \le \mathcal{R}_n^* \quad := \quad & S_n^2\mathcal{R}_n = \int_0^1 (F_n^{-1}(t) - \Phi^{-1}(t))^2 dt - \left(\int_0^1 (F_n^{-1}(t) - \Phi^{-1}(t))dt\right)^2 \\ - \quad & \left(\int_0^1 (F_n^{-1}(t) - \Phi^{-1}(t))\Phi^{-1}(t)dt\right)^2 := \mathcal{R}_n^{(1)} - \mathcal{R}_n^{(2)} - \mathcal{R}_n^{(3)}. \qquad (5) \end{aligned}$$

Observe that $n\mathcal{R}_n^{(2)} = (n^{1/2}\bar{X}_n)^2$, which has a χ_1^2 asymptotic law. On the other hand

$$n\mathcal{R}_n^{(3)} = \left(n^{1/2}\left(\int_0^1 F_n^{-1}(t)\Phi^{-1}(t)dt - 1\right)\right)^2 = (n^{1/2}(\hat{\sigma}_n - 1))^2,$$

which has a scaled χ_1^2 asymptotic law. Finally note that $n\mathcal{R}_n^{(1)}$ is the statistic L_n^0 of De Wet and Venter. However, we need a joint treatment of $(\mathcal{R}_n^{(1)}, \mathcal{R}_n^{(2)}, \mathcal{R}_n^{(3)})$.

This joint treatment can be handled in terms of the quantile process ρ_n defined by

$$\rho_n(t) := n^{1/2}\phi(\Phi^{-1}(t))\left(\Phi^{-1}(t) - F_n^{-1}(t)\right),\ 0 \le t \le 1,$$

which, as can be easily showed, verifies the regularity conditions introduced in [5]. Thus,

Theorem 1 *(see Theorem 6.2.1 in [7]) We can define in a rich enough probability space a sequence of Brownian bridges $\{B_n(t), 0 \le t \le 1\}_n$ such that*

$$n^{(1/2)-\nu} \sup_{\frac{1}{n+1} \le t \le 1-\frac{1}{n+1}} \frac{|\rho_n(t) - B_n(t)|}{(t(1-t))^\nu} = \begin{cases} O_P(\log n), & \text{if } \nu = 0, \\ O_P(1), & \text{if } 0 < \nu \le \frac{1}{2}. \end{cases}$$

Theorem 1 allows to consider jointly the three integrals in (5) because, in terms of the general quantile process, decomposition (5) is equivalent to

$$n\left(S_n^2 - \hat{\sigma}_n^2\right) = \int_0^1 \left(\frac{\rho_n(t)}{\phi(\Phi^{-1}(t))}\right)^2 dt - \left(\int_0^1 \frac{\rho_n(t)}{\phi(\Phi^{-1}(t))} dt\right)^2 - \left(\int_0^1 \frac{\rho_n(t)\Phi^{-1}(t)}{\phi(\Phi^{-1}(t))} dt\right)^2.$$

To carry out our program we begin by considering the behaviour of the integrals at the extremes and then (Proposition 3) we show that the approximation given by Theorem 1 reduces the problem to the corresponding one in terms of a Brownian bridge.

Proposition 2 *If $\{X_{in}, i = 1, ..., n\}$ is the ordered sample obtained from an i.i.d. random sample of a standard nornal law, then:*

$$n\int_0^{\frac{1}{n}} \left(X_{1n} - \Phi^{-1}(t)\right)^2 dt \xrightarrow{p} 0 \text{ and } n\int_{1-\frac{1}{n}}^1 \left(X_{nn} - \Phi^{-1}(t)\right)^2 dt \xrightarrow{p} 0.$$

PROOF: By simmetry it suffices to consider the behaviour of $\{X_{1n}\}_n$. It is well known (see e.g. [10]) that $a_n\left(X_{1n} - b_n\right) \xrightarrow{\mathcal{L}}$ for some $a_n \to \infty$ and $b_n = \Phi^{-1}(1/n)$. Hence

$$n\int_0^{1/n} \left(X_{1n} - b_n\right)^2 dt = \left(X_{1n} - b_n\right)^2 \xrightarrow{p} 0.$$

Then, by Schwarz's inequality and the following decomposition:

$$\begin{aligned} n\int_0^{1/n} \left(X_{1n} - \Phi^{-1}(t)\right)^2 dt &= n\int_0^{1/n} \left(X_{1n} - b_n\right)^2 dt + n\int_0^{1/n} \left(b_n - \Phi^{-1}(t)\right)^2 dt \\ &\quad +2n\left(X_{1n} - b_n\right)\int_0^{1/n} \left(b_n - \Phi^{-1}(t)\right) dt, \end{aligned}$$

we only need to prove that the second summand on the right hand side converges to 0. But this is an easy consequence of L'Hôpital's rule and the well known equivalence $\phi(\Phi^{-1}(x)) \approx \left|\Phi^{-1}(x)\right| x$ as $x \to 0$:

$$\lim_{x\to 0}\frac{1}{x}\int_0^x (\Phi^{-1}(x)-\Phi^{-1}(t))^2 dt = \lim_{x\to 0}\frac{2\int_0^x \Phi^{-1}(x)-\Phi^{-1}(t)dt}{\phi(\Phi^{-1}(x))} = \lim_{x\to 0}\frac{-2x}{\Phi^{-1}(x)\phi(\Phi^{-1}(x))} = 0.$$

□

Proposition 3 *There exists, on an adequate probability space, a sequence* $\{B_n(t)\}_n$ *of Brownian bridges such that the statistic* $n\mathcal{R}_n^* = n\left(S_n^2 - \hat{\sigma}_n^2\right)$ *verifies*

$$n\mathcal{R}_n^* - \left(\int_{\frac{1}{(n+1)}}^{\frac{n}{n+1}} \left(\frac{B_n(t)}{\phi(\Phi^{-1}(t))}\right)^2 dt - \left(\int_{\frac{1}{(n+1)}}^{\frac{n}{n+1}} \frac{B_n(t)}{\phi(\Phi^{-1}(t))}dt\right)^2 - \left(\int_{\frac{1}{(n+1)}}^{\frac{n}{n+1}} \frac{B_n(t)\Phi^{-1}(t)}{\phi(\Phi^{-1}(t))}dt\right)^2\right) \xrightarrow{p} 0.$$

PROOF: From Proposition 2 and the obvious inequality (valid for every Borel set A)

$$\int_A (F_n^{-1}(t)-\Phi^{-1}(t))^2 dt \geq \left(\int_A (F_n^{-1}(t)-\Phi^{-1}(t))dt\right)^2 \vee \left(\int_A (F_n^{-1}(t)-\Phi^{-1}(t))\Phi^{-1}(t)dt\right)^2,$$

it follows that

$$n\mathcal{R}_n^* - \left(\int_{\frac{1}{(n+1)}}^{\frac{n}{n+1}} \left(\frac{\rho_n(t)}{\phi(\Phi^{-1}(t))}\right)^2 dt - \left(\int_{\frac{1}{(n+1)}}^{\frac{n}{n+1}} \frac{\rho_n(t)}{\phi(\Phi^{-1}(t))}dt\right)^2 - \left(\int_{\frac{1}{(n+1)}}^{\frac{n}{n+1}} \frac{\rho_n(t)\Phi^{-1}(t)}{\phi(\Phi^{-1}(t))}dt\right)^2\right) \xrightarrow{p} 0.$$

Therefore, our claim reduces to showing that (on an adequate space)

$$\begin{aligned} L_n^{(1)} &:= \int_{\frac{1}{(n+1)}}^{\frac{n}{n+1}} \left(\frac{\rho_n(t)}{\phi(\Phi^{-1}(t))}\right)^2 dt - \int_{\frac{1}{(n+1)}}^{\frac{n}{n+1}} \left(\frac{B_n(t)}{\phi(\Phi^{-1}(t))}\right)^2 dt \xrightarrow{p} 0, \\ L_n^{(2)} &:= \left(\int_{\frac{1}{n+1}}^{\frac{n}{n+1}} \frac{\rho_n(t)}{\phi(\Phi^{-1}(t))}dt\right)^2 - \left(\int_{\frac{1}{n+1}}^{\frac{n}{n+1}} \frac{B_n(t)}{\phi(\Phi^{-1}(t))}dt\right)^2 \xrightarrow{p} 0 \quad \text{and} \\ L_n^{(3)} &:= \left(\int_{\frac{1}{n+1}}^{\frac{n}{n+1}} \frac{\rho_n(t)\Phi^{-1}(t)}{\phi(\Phi^{-1}(t))}dt\right)^2 - \left(\int_{\frac{1}{n+1}}^{\frac{n}{n+1}} \frac{B_n(t)\Phi^{-1}(t)}{\phi(\Phi^{-1}(t))}dt\right)^2 \xrightarrow{p} 0. \end{aligned} \tag{6}$$

We will study first the asymptotic behaviour of $L_n^{(1)}$. Theorem 1 guarantees the existence of a sequence of Brownian bridges such that, for every $\nu \in (0, 1/2)$:

$$\begin{aligned} &\left|\int_{\frac{1}{n+1}}^{\frac{n}{n+1}} \left(\frac{\rho_n(t)}{\phi(\Phi^{-1}(t))}\right)^2 dt - \int_{\frac{1}{n+1}}^{\frac{n}{n+1}} \left(\frac{B_n(t)}{\phi(\Phi^{-1}(t))}\right)^2 dt\right| \\ &\leq \int_{\frac{1}{n+1}}^{\frac{n}{n+1}} \left(\frac{\rho_n(t)-B_n(t)}{\phi(\Phi^{-1}(t))}\right)^2 dt + 2\int_{\frac{1}{n+1}}^{\frac{n}{n+1}} \frac{|\rho_n(t)-B_n(t)||B_n(t)|}{\phi(\Phi^{-1}(t))^2}dt \\ &\leq O_p(1)n^{2\nu-1}\int_{\frac{1}{n+1}}^{\frac{n}{n+1}} \frac{(t(1-t))^{2\nu}}{\phi(\Phi^{-1}(t))^2}dt + O_p(1)n^{\nu-\frac{1}{2}}\int_{\frac{1}{n+1}}^{\frac{n}{n+1}} \frac{(t(1-t))^{\nu}|B_n(t))|}{\phi(\Phi^{-1}(t))^2}dt \\ &:= A_n^{(1)} + A_n^{(2)} \end{aligned}$$

But if $0 < \alpha < 1$, then

$$\lim_{n\to\infty} n^{\alpha-1} \int_{\frac{1}{n+1}}^{\frac{n}{n+1}} \frac{(t(1-t))^\alpha}{\phi(\Phi^{-1}(t))^2} dt = 0 \tag{7}$$

because the equivalence $|x|\Phi(x) \approx \phi(x)$, as $x \to -\infty$, easily shows that

$$\begin{aligned} & n^{\alpha-1} \int_{\frac{1}{n+1}}^{\frac{1}{2}} \frac{t^\alpha}{\phi(\Phi^{-1}(t))^2} dt \\ = \quad & \frac{-n^{\alpha-1}}{(n+1)^\alpha} \frac{\Phi^{-1}(\frac{1}{n+1})}{\phi(\Phi^{-1}(\frac{1}{n+1}))} - n^{\alpha-1} \int_{\frac{1}{n+1}}^{\frac{1}{2}} \frac{\alpha t^{\alpha-1}\phi(\Phi^{-1}(t)) + t^\alpha \Phi^{-1}(t)}{\phi(\Phi^{-1}(t))^2} \Phi^{-1}(t) dt \to 0. \end{aligned}$$

Therefore $A_n^{(1)} \xrightarrow{p} 0$. On the other hand, also for $\nu \in (0, 1/2)$ (taking $\alpha = \nu + \frac{1}{2}$ in (7))

$$E\left[n^{\nu-\frac{1}{2}} \int_{\frac{1}{n}}^{\frac{n}{n+1}} \frac{(t(1-t))^\nu |B_n(t))|}{\phi(\Phi^{-1}(t))^2} dt \right] = n^{\nu-\frac{1}{2}} \int_{\frac{1}{n}}^{\frac{n}{n+1}} \frac{(t(1-t))^{\nu+\frac{1}{2}}}{\phi(\Phi^{-1}(t))^2} dt \to 0,$$

thus $A_n^{(2)} \xrightarrow{p} 0$, which shows that $L_n^{(1)} \xrightarrow{p} 0$.

Let us consider now $L_n^{(2)}$. We can rewrite it as follows:

$$L_n^{(2)} = \left(\int_{\frac{1}{n+1}}^{\frac{n}{n+1}} \frac{\rho_n(t) - B_n(t)}{\phi(\Phi^{-1}(t))} dt \right) \left(\int_{\frac{1}{n+1}}^{\frac{n}{n+1}} \frac{\rho_n(t) + B_n(t)}{\phi(\Phi^{-1}(t))} dt \right). \tag{8}$$

The first factor on the right hand side of (8) is bounded by

$$\left[\int_{\frac{1}{n+1}}^{\frac{n}{n+1}} \left(\frac{\rho_n(t) - B_n(t)}{\phi(\Phi^{-1}(t))} \right)^2 dt \right]^{1/2} \xrightarrow{p} 0, \tag{9}$$

where the last convergence is a consequence of the convergence $A_n^{(1)} \xrightarrow{p} 0$ showed above. Moreover, it is well known that the law of

$$\int_{\frac{1}{n+1}}^{\frac{n}{n+1}} \frac{B_n(t)}{\phi(\Phi^{-1}(t))} dt$$

is $N(0, \sigma_1^2(1/(n+1)))$, with

$$\sigma_1^2(x) := \int_x^{1-x} \int_x^{1-x} \frac{u \wedge v - uv}{\phi(\Phi^{-1}(u))\phi(\Phi^{-1}(v))} dudv.$$

It is easy to verify that $\sigma_1^2(x) \to 1$ as $x \to 0$, from which

$$\int_{\frac{1}{n+1}}^{\frac{n}{n+1}} \frac{B_n(t)}{\phi(\Phi^{-1}(t))} dt = O_p(1),$$

and, therefore, also

$$\int_{\frac{1}{n+1}}^{\frac{n}{n+1}} \frac{\rho_n(t)}{\phi(\Phi^{-1}(t))} dt = O_p(1),$$

which, combined with (9), shows $L_n^{(2)} \xrightarrow{p} 0$. Similarly,

$$\left| \int_{\frac{1}{n+1}}^{\frac{n}{n+1}} \frac{(\rho_n(t) - B_n(t))\,\Phi^{-1}(t)}{\phi(\Phi^{-1}(t))} dt \right| \leq \left[\left(\int_0^1 (\Phi^{-1}(t))^2 dt \right) \int_{\frac{1}{n+1}}^{\frac{n}{n+1}} \frac{(\rho_n(t) - B_n(t))^2}{(\phi(\Phi^{-1}(t)))^2} dt \right]^{1/2} \xrightarrow{p} 0.$$

Since $\int_{\frac{1}{n+1}}^{\frac{n}{n+1}} \frac{B_n(t)\Phi^{-1}(t)}{\phi(\Phi^{-1}(t))} dt$ has a $N(0, \sigma_2^2(\frac{1}{n+1}))$ law, where

$$\sigma_2^2(x) := \int_x^{1-x} \int_x^{1-x} \frac{u \wedge v - uv}{\phi(\Phi^{-1}(u))\phi(\Phi^{-1}(v))} \Phi^{-1}(u)\Phi^{-1}(v) du dv \to 1/2 \text{ as } x \to 0,$$

we can use the reasoning employed for $L_n^{(2)}$ to show that also $L_n^{(3)} \xrightarrow{p} 0$, completing the proof of (6). □

Now, showing convergence and describing the limit law are easier tasks. In the next theorem we obtain the asymptotic law of $\mathcal{R}_n$ through its equivalent version based on the Brownian bridge. Note that the main dificulty is to give sense to expression Z, defined by (4), because, as stated in the Introduction, the function involved is a.s. not integrable (see Lemma 2.2 in [8]). Thus, we cannot assume the existence of

$$\lim_n \int_{\frac{1}{n+1}}^{\frac{n}{n+1}} \frac{B^2(t) - EB^2(t)}{(\phi(\Phi^{-1}(t)))^2} dt.$$

But it turns out that this limit does exist in L_2-sense and we can define Z as this L_2-limit. This process is carried out in the next theorem.

Theorem 4 *Let $\{X_n\}_n$ be a sequence of i.i.d. normal random variables. Then*

$$n(\mathcal{R}_n - a_n) \xrightarrow{\mathcal{L}} \int_0^1 \frac{B^2(t) - EB^2(t)}{(\phi(\Phi^{-1}(t)))^2} dt - \left(\int_0^1 \frac{B(t)}{\phi(\Phi^{-1}(t))} dt \right)^2 - \left(\int_0^1 \frac{B(t)\Phi^{-1}(t)}{\phi(\Phi^{-1}(t)} dt \right)^2,$$

where

$$a_n = \frac{1}{n} \int_{\frac{1}{n+1}}^{\frac{n}{n+1}} \frac{t(1-t)}{[\phi(\Phi^{-1}(t))]^2} dt.$$

PROOF: By the invariance of $\mathcal{R}_n$ we can assume, without loss of generality, that X_n has a standard normal law. Then, by the asymptotic normality of S_n^2, we have

$$n(\mathcal{R}_n - a_n) - n(\mathcal{R}_n^* - a_n) = \frac{n}{S_n^2} \mathcal{R}_n^* (1 - S_n^2) = O_p(1)\sqrt{n}(\mathcal{R}_n^* - a_n + a_n) \xrightarrow{p} 0 \qquad (10)$$

provided $n(\mathcal{R}_n^* - a_n) = O_p(1)$. In particular this shows that we can prove the theorem by showing that the asymptotic law of $n(\mathcal{R}_n^* - a_n)$ is that of the functional of the Brownian bridge involved in the statement of the theorem and, by Proposition 3, it even suffices to give a limit sense to

$$\int_0^1 \frac{B^2(t) - EB^2(t)}{(\phi(\Phi^{-1}(t)))^2} dt.$$

If

$$A_n := \int_{\frac{1}{n+1}}^{\frac{n}{n+1}} \frac{B^2(t) - EB^2(t)}{(\phi(\Phi^{-1}(t)))^2} dt,$$

then it can be shown that

$$EA_n^2 = \int_{\frac{1}{n+1}}^{\frac{n}{n+1}} \int_{\frac{1}{n+1}}^{\frac{n}{n+1}} \frac{2(s \wedge t - st)^2}{(\phi(\Phi^{-1}(s))(\phi(\Phi^{-1}(t)))^2} dsdt \to \int_0^1 \int_0^1 \frac{2(s \wedge t - st)^2}{(\phi(\Phi^{-1}(s))\phi(\Phi^{-1}(t)))^2} dsdt < \infty.$$

From this it is easy to see that $E(A_n - A_m)^2 \to 0$ as $n, m \to \infty$ and, hence, that A_n converges in L_2. □

The next theorem provides the known (see [9]) explicit expression for the limit law just obtained. Its proof, which will not be detailed here, relies on a careful principal components expansion (see [1] for details) based on the eigenfunctions of the operator

$$Lf(t) := \int_0^1 \frac{s \wedge t - st}{\phi(\Phi^{-1}(s))\phi(\Phi^{-1}(t))} f(s) ds.$$

Theorem 5 *Let $\{X_n\}_n$ be a sequence of i.i.d. normal random variables. Then*

$$n(\mathcal{R}_n - a_n) \xrightarrow{\mathcal{L}} -\frac{3}{2} + \sum_{j=3}^{\infty} \frac{Z_j^2 - 1}{j},$$

where $\{Z_n\}_n$ is a sequence of independent $N(0,1)$ random variables and

$$a_n = \frac{1}{n} \int_{\frac{1}{n+1}}^{\frac{n}{n+1}} \frac{t(1-t)}{[\phi(\Phi^{-1}(t))]^2} dt.$$

Remarks: Recall decomposition (5). It can be shown that the only distribution for which the asymptotic terms corresponding to $n\mathcal{R}_n^{(2)}$ and $n\mathcal{R}_n^{(3)}$ just cancel out some terms of the principal components expansion of the limit law of $n\mathcal{R}_n^{(1)}$ is the normal.

The asymptotic equivalence of $\mathcal{R}_n$ with the statistics of Shapiro-Wilk, Shapiro-Francia or De Wet-Venter can be obtained through the available results in [11] or in [14].

Acknowledgement. Part of the present ideas are the final result of helpful conversations with Prof. Evarist Giné, while carrying out other joint work with him.

References

[1] del Barrio, E.; Matrán, C.; Rodríguez-Rodríguez, J.; Cuesta-Albertos, J. A. (1998) Descriptive statistics and tests of goodness of fit based on the L_2-Wasserstein distance. Technical Report

[2] Bickel, P. and Freedman, D. (1981) Some asymptotic theory for the bootstrap. *Ann. Statist.*, 9, 1196-1217.

[3] Brown, B. and Hettmansperger, T. (1996) Normal scores, normal plots, and test for normality. *Jour. Amer. Stat. Assoc.*, 91, 1668-1675.

[4] Csörgő, M. (1983) *Quantile Processes with Statistical Applications.* SIAM, Philadelphia.

[5] Csörgő, M. and Révész, P. (1978) Strong approximations of the quantile process. *Ann. Statist.*, 6, 882-894.

[6] Csörgő, M. and Horváth, L. (1988) On the distributions of L_p norms of weighted uniform empirical and quantile process. *Ann. Prob.*, 16, 142-161.

[7] Csörgő, M. and Horváth, L. (1993) *Weighted Approximations in Probability and Statistics.* John Wiley and Sons.

[8] Csörgő, M., Horváth, L. and Shao, Q.-M. (1993). Convergence of integrals of uniform empirical and quantile processes *Stochastic. Process. Appl.*, 45, 283-294.

[9] De Wet, T. and Venter, J. (1972). Asymptotic distributions of certain test criteria of normality. *South Africa Statistical Journal*, 6, 135-149.

[10] Galambos, J. (1987). *The Asymptotic Theory of Extreme Order Statistics* 2nd ed.. Krieger, Melbourne, Florida.

[11] Leslie, J.R.; Stephens, M.A. and Fotopoulos, S. (1986). Asymptotic distribution of the Shapiro-Wilk W for testing for normality. *Ann. Statist.*, 14, 1497-1506.

[12] Lockhart, R.A. and Stephens, M.A. (1996). The Probability Plot: Test of Fit Based on the Correlation Coefficient. *in N. Balakrishnan and C.R. Rao eds., Handbook of Statistics.*

[13] Shapiro, S. and Wilk, M. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, 52, 591-611.

[14] Verrill, S. and Johnson, R. (1987). The asymptotic equivalence of some modified Shapiro-Wilk statistics- Complete and censored sample cases. *Ann. Statist.*, 15, 413-419.