The Annals of Statistics
1999, Vol. 27, No. 4, 1230–1239

TESTS OF GOODNESS OF FIT BASED ON THE L_2-WASSERSTEIN DISTANCE[1]

By Eustasio del Barrio, Juan A. Cuesta-Albertos, Carlos Matrán and Jesús M. Rodríguez-Rodríguez

Universidad de Valladolid, Universidad de Cantabria, Universidad de Valladolid and Universidad de Valladolid

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

1. Introduction. Goodness-of-fit tests are often based on some distances between probability laws (p.l.'s). In this work we follow this approach by using the L_2-Wasserstein distance between a fixed distribution and a location scale family of probability distributions on $\mathbb{R}$. We focus on the (more interesting) normal case, but our approach can be used to cover other distributions (see [8] for details).

Let $\mathscr{P}_2(\mathbb{R})$ be the set of probabilities on $\mathbb{R}$ with finite second moment. For probabilities P_1 and P_2 in $\mathscr{P}_2(\mathbb{R})$, the L_2-Wasserstein distance between P_1 and P_2 is defined as

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

For distributions P_1 and P_2 on $\mathbb{R}$ the distance $\mathscr{W}$ can be explicitly calculated (see, e.g., [1]):

$$\mathscr{W}(P_1, P_2) = \left[\int_0^1 (F_1^{-1}(t) - F_2^{-1}(t))^2\, dt\right]^{1/2}, \tag{1.1}$$

where F_1^{-1} and F_2^{-1} are the quantile functions of P_1 and P_2, respectively [recall that for a distribution function F the quantile function F^{-1} is defined by $F^{-1}(t) = \inf\{s\colon F(s) \geq t\}$ and note that the right-hand side of (1.1) is equal to $[E(F_1^{-1}(U) - F_2^{-1}(U))^2]^{1/2}$, where U is a r.v. with uniform distribution on $(0, 1)$].

For simplification of notation we will often identify a probability law with its distribution function (d.f.). The d.f. of the standard normal law will be denoted by Φ and its density by ϕ. We write $\mathscr{H}_{\mathscr{N}} := \{H\colon H(x) = \Phi((x-\mu)/\sigma),\ \mu \in \mathbb{R},\ \sigma > 0\}$ for the class of normal laws on the line.

Received February 1998; revised June 1999.

[1]Supported in part by DGICYT Grants PB95-0715-c02 and PAPIJCL VA08/97.

AMS 1991 *subject classifications*. Primary 62F05, 62E20; secondary 60F25.

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

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

$$
\begin{aligned}
\mathscr{W}^2(P, \mathscr{H}_{\mathscr{N}}) &:= \inf\{\mathscr{W}^2(P, H),\ H \in \mathscr{H}_{\mathscr{N}}\} \\
&= \inf_{\sigma>0}\left\{\int_0^1 (F^{-1}(t) - \mu_0 - \sigma\Phi^{-1}(t))^2\, dt\right\} \\
&= \sigma_0^2 - \left(\int_0^1 (F^{-1}(t) - \mu_0)\Phi^{-1}(t)dt\right)^2 \\
&= \sigma_0^2 - \left(\int_0^1 F^{-1}(t)\Phi^{-1}(t)dt\right)^2.
\end{aligned}
\tag{1.2}
$$

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

Now let $X_1, X_2, \dots, X_n$ be a random sample with underlying d.f. F and let F_n denote its empirical d.f. and S_n^2 the sample variance. A sample version of $\mathscr{W}^2(P, \mathscr{H}_{\mathscr{N}})/\sigma_0^2$ is given by

$$
\mathscr{R}_n := \frac{\mathscr{W}^2(P_n, \mathscr{H}_{\mathscr{N}})}{S_n^2} = 1 - \frac{(\int_0^1 F_n^{-1}(t)\Phi^{-1}(t)\, dt)^2}{S_n^2}.
\tag{1.3}
$$

This quantity can be used as a test statistic for the hypothesis of normality. In fact, $\mathscr{R}_n$ is connected with correlation tests and in particular with the Shapiro–Wilk's test of normality; see [14]. This has been noted, in the context of normal probability plots, in [2]. Moreover, $\mathscr{R}_n$ is asymptotically equivalent to Shapiro–Wilk statistics, Shapiro–Francia statistics and De Wet–Venter statistics. This can be obtained from the results in [11] and [15].

For literature concerning statistics related to Shapiro–Wilk's W statistic, see [7], [9], [11], [15] or [12]. All the proofs of the asymptotic behavior of any of these statistics given in [11], [15] and [12] rely on the results in [9].

Some other statistical applications of L_2-Wasserstein statistics can be found in [13]. Here, a trimmed version of the L_2-Wasserstein distance is considered as a dissimilarity measure between distributions and the asymptotic normality of its empirical version is shown.

The purpose of this paper is to analyze the asymptotic behaviour of $\mathscr{R}_n$ through approximations of quantile processes by Brownian bridges, $B(t)$. This approach was also used in [3] to treat a simplified version of W, but the proof makes heavy use of the results in [9] to give a sense to the limit expression,

$$
Z := \int_0^1 \frac{B^2(t) - EB^2(t)}{(\phi(\Phi^{-1}(t)))^2}\, dt.
\tag{1.4}
$$

Note that Z is not defined because the set of trajectories of a Brownian bridge $B(t)$ for which the function $t \mapsto (B^2(t) - EB^2(t))/(\phi(\Phi^{-1}(t)))^2)$ is integrable has probability zero. This difficulty will be circumvented in Theorem 2, where we will show that the sequence

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

is an L_2-Cauchy sequence, so we can define Z as the L_2-limit of this sequence.

We should remark that the ambitious program on the convergence of integrals of empirical and quantile processes developed in [4], [5] and [6] does not cover our results.

We also remark that the present approach can be applied to other location scale families, including those with heavy tails (see [8] for details).

2. The results. The normal law closest to the empirical d.f. F_n has mean $\hat{\mu}_n = \mu(F_n) = \bar{X}_n$ and standard deviation

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

if we denote the order statistic by $X_{kn},\ k = 1, 2, \ldots, n$.

Our measure of nonnormality is then

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

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

The invariance of $\mathscr{R}_n$ 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 behavior of $\mathscr{R}_n$ through that of $S_n^2\mathscr{R}_n$. We use the following decomposition:

$$\begin{aligned} 0 \le \mathscr{R}_n^* := S_n^2\mathscr{R}_n &= \int_0^1 (F_n^{-1}(t) - \Phi^{-1}(t))^2\,dt \\ &\quad - \left(\int_0^1 (F_n^{-1}(t) - \Phi^{-1}(t))\,dt\right)^2 \\ &\quad - \left(\int_0^1 (F_n^{-1}(t) - \Phi^{-1}(t))\Phi^{-1}(t)\,dt\right)^2 \\ &:= \mathscr{R}_n^{(1)} - \mathscr{R}_n^{(2)} - \mathscr{R}_n^{(3)}. \end{aligned} \tag{2.1}$$

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

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

has a scaled χ_1^2 asymptotic law. Note that $n\mathscr{R}_n^{(1)}$ is the statistic L_n^0 of De Wet and Venter. We need a joint treatment of $(\mathscr{R}_n^{(1)}, \mathscr{R}_n^{(2)}, \mathscr{R}_n^{(3)})$. Note also that

$$\mathscr{R}_n^{(1)} = \int_0^1 \left(\frac{\rho_n(t)}{\phi(\Phi^{-1}(t))} \right)^2 dt, \tag{2.2}$$

$$\mathscr{R}_n^{(2)} = \left(\int_0^1 \frac{\rho_n(t)}{\phi(\Phi^{-1}(t))} \, dt \right)^2, \tag{2.3}$$

$$\mathscr{R}_n^{(3)} = \left(\int_0^1 \frac{\rho_n(t)\Phi^{-1}(t)}{\phi(\Phi^{-1}(t))} \, dt \right)^2, \tag{2.4}$$

where ρ_n is the quantile process defined by

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

We will use the following result on ρ_n.

THEOREM 1 (see Theorem 6.2.1 in [5]). *On a rich enough probability space we can define a sequence of Brownian bridges $\{B_n(t),\ 0 \le t \le 1\}_n$ such that*

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

Before we make use of the approximation of ρ_n, given in Theorem 1, we treat the behavior of the integrals in (2.2), (2.3) and (2.4) at the boundary.

PROPOSITION 1. *If $\{X_{in},\ i = 1, \dots, n\}$ is the ordered sample obtained from an i.i.d. sample with standard normal law, then*

$$n \int_0^{1/n} (X_{1n} - \Phi^{-1}(t))^2 \, dt \to_p 0 \quad \textit{and} \quad n \int_{1-1/n}^1 (X_{nn} - \Phi^{-1}(t))^2 \, dt \to_p 0.$$

PROOF. By symmetry it suffices to consider the behavior of $\{X_{1n}\}_n$. It is well known (see, e.g., [10]) that $a_n(X_{1n} - b_n)$ converges weakly to a nondegenerate limit law for some $a_n \to \infty$ and $b_n = \Phi^{-1}(1/n)$. Hence

$$n \int_0^{1/n} (X_{1n} - b_n)^2 \, dt = (X_{1n} - b_n)^2 \to_p 0.$$

The result follows from

$$n \int_0^{1/n} (b_n - \Phi^{-1}(t))^2 \, dt \to 0. \tag{2.5}$$

Claim (2.5) is an easy consequence of l'Hôpital's rule and the well-known equivalence $\phi(\Phi^{-1}(x)) \approx |\Phi^{-1}(x)|x$ as $x \to 0$, as follows:

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

□

PROPOSITION 2. *On an adequate probability space, there exists a sequence* $\{B_n(t)\}_n$ *of Brownian bridges such that the statistic* $n\mathscr{R}_n^* = n(S_n^2 - \hat{\sigma}_n^2)$ *fulfills*

$$n\mathscr{R}_n^* - \left(\int_{1/(n+1)}^{n/(n+1)} \left(\frac{B_n(t)}{\phi(\Phi^{-1}(t))}\right)^2 dt - \left(\int_{1/(n+1)}^{n/(n+1)} \frac{B_n(t)}{\phi(\Phi^{-1}(t))}\, dt\right)^2 - \left(\int_{1/(n+1)}^{n/(n+1)} \frac{B_n(t)\Phi^{-1}(t)}{\phi(\Phi^{-1}(t))}\, dt\right)^2\right) \to_p 0.$$

PROOF. From Proposition 1 and the obvious inequality (valid for every Borel set A),

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

it follows that

$$n\mathscr{R}_n^* - \left(\int_{1/(n+1)}^{n/(n+1)} \left(\frac{\rho_n(t)}{\phi(\Phi^{-1}(t))}\right)^2 dt - \left(\int_{1/(n+1)}^{n/(n+1)} \frac{\rho_n(t)}{\phi(\Phi^{-1}(t))}\, dt\right)^2 - \left(\int_{1/(n+1)}^{n/(n+1)} \frac{\rho_n(t)\Phi^{-1}(t)}{\phi(\Phi^{-1}(t))}\, dt\right)^2\right) \to_p 0.$$

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

$$\begin{aligned} L_n^{(1)} &:= \int_{1/(n+1)}^{n/(n+1)} \left(\frac{\rho_n(t)}{\phi(\Phi^{-1}(t))}\right)^2 dt \\ &\quad - \int_{1/(n+1)}^{n/(n+1)} \left(\frac{B_n(t)}{\phi(\Phi^{-1}(t))}\right)^2 dt \to_p 0, \\ L_n^{(2)} &:= \left(\int_{1/(n+1)}^{n/(n+1)} \frac{\rho_n(t)}{\phi(\Phi^{-1}(t))}\, dt\right)^2 \\ &\quad - \left(\int_{1/(n+1)}^{n/(n+1)} \frac{B_n(t)}{\phi(\Phi^{-1}(t))}\, dt\right)^2 \to_p 0 \quad \text{and} \\ L_n^{(3)} &:= \left(\int_{1/(n+1)}^{n/(n+1)} \frac{\rho_n(t)\Phi^{-1}(t)}{\phi(\Phi^{-1}(t))}\, dt\right)^2 \\ &\quad - \left(\int_{1/(n+1)}^{n/(n+1)} \frac{B_n(t)\Phi^{-1}(t)}{\phi(\Phi^{-1}(t))}\, dt\right)^2 \to_p 0. \end{aligned} \tag{2.6}$$

We will study first the asymptotic behavior of $L_n^{(1)}$. Theorem 1 guarantees the existence of a sequence of Brownian bridges such that, for every

$\nu \in (0, 1/2)$:

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

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

$$
\lim_{n \to \infty} n^{\alpha - 1} \int_{1/(n+1)}^{n/(n+1)} \frac{(t(1-t))^{\alpha}}{\phi(\Phi^{-1}(t))^2} dt = 0 \tag{2.7}
$$

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

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

Therefore $A_n^{(1)} \to_p 0$. On the other hand, also for $\nu \in (0, 1/2)$ [taking $\alpha = \nu + \frac{1}{2}$ in (2.7)],

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

thus $A_n^{(2)} \to_p 0$. This shows that $L_n^{(1)} \to_p 0$.

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

$$
L_n^{(2)} = \left(\int_{1/(n+1)}^{n/(n+1)} \frac{\rho_n(t) - B_n(t)}{\phi(\Phi^{-1}(t))} dt \right) \left(\int_{1/(n+1)}^{n/(n+1)} \frac{\rho_n(t) + B_n(t)}{\phi(\Phi^{-1}(t))} dt \right). \tag{2.8}
$$

The first factor on the right-hand side of (2.8) is bounded by

$$\left[\int_{1/(n+1)}^{n/(n+1)}\left(\frac{\rho_n(t)-B_n(t)}{\phi(\Phi^{-1}(t))}\right)^2 dt\right]^{1/2} \to_p 0, \tag{2.9}$$

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

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

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

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

It is easy to verify that $\sigma_1^2(x) \to 1$ as $x \to 0$. This shows

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

and, therefore, also

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

This and (2.9) show $L_n^{(2)} \to_p 0$. Similarly, we get

$$\begin{aligned}&\left|\int_{1/(n+1)}^{n/(n+1)} \frac{(\rho_n(t)-B_n(t))\Phi^{-1}(t)}{\phi(\Phi^{-1}(t))}\, dt\right|\\ &\qquad\le \left[\left(\int_0^1 (\Phi^{-1}(t))^2\, dt\right)\int_{1/(n+1)}^{n/(n+1)} \frac{(\rho_n(t)-B_n(t))^2}{(\phi(\Phi^{-1}(t)))^2}\, dt\right]^{1/2} \to_p 0.\end{aligned}$$

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

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

With similar arguments as for $L_n^{(2)}$ we get that $L_n^{(3)} \to_p 0$. This completes the proof of (2.6). □

In the next theorem we obtain the asymptotic law of $\mathscr{R}_n$ through its equivalent version based on the Brownian bridge. Note that the main dificulty is to give sense to expression Z, defined by (1.4), because, as stated in the introduction, the integrand in the definition of Z is a.s. not integrable (see Lemma 2.2 in [6]). Thus

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

does not exist.

However, it turns out that this limit does exist in the L_2-sense and we can define Z as this L_2-limit. This will be done in the proof of the next theorem.

THEOREM 2. *Let $\{X_n\}_n$ be a sequence of i.i.d. normal random variables. Then*

$$\begin{aligned} n(\mathscr{R}_n - a_n) \to_{\mathscr{L}} \int_0^1 \frac{B^2(t) - EB^2(t)}{(\phi(\Phi^{-1}(t)))^2}\, dt - \left(\int_0^1 \frac{B(t)}{\phi(\Phi^{-1}(t))}\, dt\right)^2 \\ - \left(\int_0^1 \frac{B(t)\Phi^{-1}(t)}{\phi(\Phi^{-1}(t))}\, dt\right)^2, \end{aligned} \tag{2.10}$$

where

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

PROOF. By the invariance of $\mathscr{R}_n$ 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

$$n(\mathscr{R}_n - a_n) - n(\mathscr{R}_n^* - a_n) = \frac{n}{S_n^2}\mathscr{R}_n^*(1 - S_n^2) = O_p(1)\sqrt{n}(\mathscr{R}_n^* - a_n + a_n) \to_p 0,$$

provided $n(\mathscr{R}_n^* - a_n) = O_p(1)$. It remains to show that $n(\mathscr{R}_n^* - a_n)$ converges to the right-hand side of (2.10). By Proposition 2, 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

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

then it can be shown that

$$\begin{aligned} EA_n^2 &= \int_{1/(n+1)}^{n/(n+1)} \int_{1/(n+1)}^{n/(n+1)} \frac{2(s \wedge t - st)^2}{(\phi(\Phi^{-1}(s))(\phi(\Phi^{-1}(t)))^2}\, ds\, dt \\ &\to \int_0^1 \int_0^1 \frac{2(s \wedge t - st)^2}{(\phi(\Phi^{-1}(s))\phi(\Phi^{-1}(t)))^2}\, ds\, dt < \infty. \end{aligned}$$

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 explicit expression for the limit law of $\mathscr{R}_n$; see [9]. A new proof of this result can be based on Theorem 2. The proof, which will not be given here, relies on a careful principal components expansion based on the eigenfunctions of the operator

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

see [8] for details.

THEOREM 3. *Let $\{X_n\}_n$ be a sequence of i.i.d. normal random variables. Then*

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

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

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

Acknowledgment. Some 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] BICKEL, P. and FREEDMAN, D. (1981).Some asymptotic theory for the bootstrap. *Ann. Statist.* **9** 1196–1217.

[2] BROWN, B. and HETTMANSPERGER, T. (1996).Normal scores, normal plots, and test for normality. *J. Amer. Statist. Assoc.* **91** 1668–1675.

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

[4] CSÖRGŐ, M. and HORVÁTH, L. (1988).On the distributions of L_p norms of weighted uniform empirical and quantile process. *Ann. Probab.* **16** 142–161.

[5] CSÖRGŐ, M. and HORVÁTH, L. (1993).*Weighted Approximations in Probability and Statistics*. Wiley, New York.

[6] CSÖRGŐ, M., HORVÁTH, L. and SHAO, Q.-M. (1993).Convergence of integrals of uniform empirical and quantile processes. *Stochastic Process. Appl.* **45** 283–294.

[7] CSÖRGŐ, M. and RÉVÉSZ, P. (1978).Strong approximations of the quantile process. *Ann. Statist.* **6** 882–894.

[8] DEL BARRIO, E., MATRÁN, C., RODRÍGUEZ-RODRÍGUEZ, J. and CUESTA-ALBERTOS, J. A. (1998).Descriptive statistics and tests of goodness of fit based on the L_2-Wasserstein distance. Technical report, Univ. Valladolid.

[9] DE WET, T. and VENTER, J. (1972).Asymptotic distributions of certain test criteria of normality. *South African Statist. J.* **6** 135–149.

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

[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. (1998).The probability plot: tests of fit based on the correlation coefficient. In *Handbook of Statistics 17: Order Statistics: Applications* (N. Balakrishnan and C.R. Rao, eds.) 453–473. North-Holland, Amsterdam.

[13] MUNK, A. and CZADO, C. (1998).Nonparametric validation of similar distributions and assesment of goodness of fit. *J. Roy. Statist. Soc. Ser. B* **60** 223–241.

[14] SHAPIRO, S. and WILK, M. (1965).An analysis of variance test for normality (complete samples). *Biometrika* **52** 591–611.

[15] 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.

E. DEL BARRIO
C. MATRÁN
J. M. RODRÍGUEZ-RODRÍGUEZ
DEPARTAMENTO DE ESTADÍSTICA E
INVESTIGACIÓN OPERATIVA
FACULTAD DE CIENCIAS
UNIVERSIDAD DE VALLADOLID
VALLADOLID, 47005
SPAIN
E-MAIL: tasio@westad.cis.uva.es

J. A. CUESTA-ALBERTOS
DEPARTAMENTO DE MATEMÁTICAS,
ESTADÍSTICA Y COMPUTACIÓN
FACULTAD DE CIENCIAS
UNIVERSIDAD DE CANTABRIA
SANTANDER, 39071
SPAIN
E-MAIL: cuesta@matesco.unican.es