The Annals of Statistics
1994, Vol. 22, No. 3, 1447–1459

GOODNESS-OF-FIT AND COMPARISON TESTS OF THE KOLMOGOROV–SMIRNOV TYPE FOR BIVARIATE POPULATIONS

BY ALEJANDRA CABAÑA AND ENRIQUE M. CABAÑA

Universidad Simón Bolívar and Universidad de la República

Statistical tests consistent under any alternative are provided for both goodness-of-fit and two-sample comparisons. They are based on (1) the derivation of estimates for the tails of the probability distribution of some particular functionals of a Wiener sheet and (2) the use of a mapping that carries normalized empirical processes (asymptotically distributed as Brownian bridge) onto new processes that converge in distribution to a Wiener sheet.

The proposed test statistics are easily obtained from the data. Hints for its computation and empirical estimations of the power of the tests are given.

1. Introduction. In this article, we consider the design of goodness-of-fit tests and two-sample comparison tests for bivariate distributions.

Let us suppose that the pair of real random variables X, Y has continuous joint distribution function F and consider an i.i.d. sample (X_1, Y_1), (X_2, Y_2), $\dots$, (X_n, Y_n). We let $F_n = (1/n)\sum_{i=1}^{n}\delta_{(X_i, Y_i)}$ denote the associated empirical measure and let $B_n^F = \sqrt{n}(F_n - F)$ denote the normalized empirical measure. Then B_n^F converges in distribution to the biparameter F-Brownian bridge B^F, a centred Gaussian measure characterized by $\mathbf{E}B^F(A_1)B^F(A_2) = F(A_1 \cap A_2) - F(A_1)F(A_2)$, for every measurable set A_1 and A_2 in $\mathbb{R}^2$. The F-Brownian bridge is also characterized as the conditional distribution of F-Wiener sheet W^F, given $W^F(\mathbb{R}^2) = 0$. Let us remember that an F-Wiener sheet is a centred Gaussian measure with covariances $\mathbf{E}W^F(A_1)W^F(A_2) = F(A_1 \cap A_2)$.

The probability distributions of functionals of a Brownian bridge, suitable for generalizing the univariate Kolmogorov–Smirnov statistics to bivariate samples, are unknown. Adler and Brown [1] have provided an estimate for the probability $\mathbf{P}\{\sup_{x,y} B^F((-\infty,x]\times(-\infty,y]) > u\}$ and claim that it can be used to construct a generalization of Kolmogorov–Smirnov tests. The present paper proposes a different approach: we are able to provide estimates for the tails of the probability distributions of functionals of W^F, which, in addition, do not depend on F. We then transform B_n^F into a new process that converges to the Wiener sheet and propose statistics based on this transformed process, suggested by analytical reasons mentioned in Section 3, to perform either goodness-of-fit tests or two-sample comparisons. These statistics are easily computed from the empirical data.

Received September 1991; revised June 1993.

AMS 1991 *subject classifications*. Primary 62G10; secondary 60G60, 60G15.

Key words and phrases. Wiener sheet, empirical measure transformations, bivariate Kolmogorov-Smirnov tests, tail estimates for functionals of Wiener sheet.

The transformation that maps a Brownian bridge onto a Wiener sheet and the estimation of the tails of functionals of a Wiener sheet are described in Sections 2 and 3, respectively. In Sections 4 and 5 we introduce the goodness-of-fit test and the two-sample comparison test. Some hints for the computations of the test statistics appear in Section 6, and Section 7 provides an empirical description of the performance of the tests.

2. A measure transformation from Brownian bridge to Wiener sheet. To each probability measure F on $\mathbb{R}^2$ we associate the *marginal measure* $\widetilde{F}$ on $\mathbb{R}$ defined by $\widetilde{F}((s,t]) = F((x,y)\colon s < x+y \le t)$, and, for each measurable $A \subset \mathbb{R}^2$, the restriction $F_A(B) = F(A \cap B)$. With this notation, the following statement holds.

THEOREM 1. *Let F be a probability measure on $\mathbb{R}^2$. Then the application*

$$\mathcal{W}^F B(A) = \sqrt{3}\left(\int \left(1 - \widetilde{F}(\tau)\right) d\widetilde{B}_A(\tau) + \int \widetilde{B}(\tau)\, d\widetilde{F}_A(\tau) \right) \tag{1}$$

transforms the F-Brownian bridge B^F into a Wiener sheet associated to the probability measure

$$\operatorname{Var} \mathcal{W}^F B^F(A) = 3 \int \left(1 - \widetilde{F}(\tau)\right)^2 d\widetilde{F}_A(\tau) = V(A). \tag{2}$$

A plain computation of covariances proves the assertion. One of the Editors has pointed out to the authors that the preceding statement is a special case of Example 4 of [6]. We omit the details for that reason.

REMARK 1. When $A = \mathbb{R}^2$, $\widetilde{F}_A(\tau)$ equals $\widetilde{F}(\tau)$, and the total variance reduces to 1.

REMARK 2. Statements (i) and (ii) follow from (2):

(i) The marginals $\widetilde{F}$ and $\widetilde{V}$ of F and V, respectively, are linked by the equation $1 - \widetilde{V}(t) = (1 - \widetilde{F}(t))^3$.

(ii) If (X_1, Y_1) has distribution F and (X_2, Y_2) has distribution V, then the conditional distributions of $X_1 - Y_1$ given $X_1 + Y_1$ and of $X_2 - Y_2$ given $X_2 + Y_2$ are the same.

REMARK 3. The test for goodness-of-fit to a distribution F described in Section 4 is simpler when the variance measure V associated to $\mathcal{W}^F B^F$ by equation (2) is the probability distribution of a pair of independent random variables. A simple example of such a pair (F_c, V_c) of probabilities on $\mathbb{R}^2 \times \mathbb{R}^2$ is obtained by arbitrarily choosing V_c as the measure with density $v_c(x,y) = \exp(-x - y)$. The marginal $\widetilde{V}(t) = 1 - (1+t)\exp(-t)$ is readily computed, and $\widetilde{F}(t) = 1 - ((1+t)\exp(-t))^{1/3}$ is obtained from the first part of Remark 2. Finally, the second part of the same remark leads us to conclude that the density of the distribution F_c is $f_c(x,y) = (1/3)/(x+y)\exp[-(x+y)/3](1+x+y)^{-2/3}$.

REMARK 4. The transformation $dw^F = db^F + [b^F/(1-F)]\,dF$ maps a one-dimensional Brownian bridge on a Wiener process, both associated to the same measure F, and has been used by Khmaladze [4] in the construction of statistical tests. We introduce the composition of that transformation with $dv = \sqrt{3}(1-F)\,dw^F$, in order to avoid the singularity at 1, and this leads to another Wiener process, associated to a different measure. The transformation in Theorem 1 generalizes this to the two-dimensional case, as in Khmaladze [5, 6].

3. Estimation of the probability tails for some functionals of the Wiener sheet. The wave components of a V-Wiener process W (a Wiener sheet associated to the probability measure of variance V) have been introduced in [2] and [3], motivated by interpreting the Wiener sheet as the solution of a wave equation with a random noise excitation. The second paper describes how these wave components allow us to state a strong Markov property for the sheet and provides estimates for the probability tails of functionals associated to the Wiener sheet.

We develop in the present paper similar estimates for other functionals, namely, differences of wave components, to be used in the tests here proposed. We do not need to be concerned here with the analytical motivation of the functionals we choose, and refer the interested reader to [2] and [3] for details.

DEFINITION 1. Let us consider the sets

$$\begin{aligned} C_1(x,y) &= \{(x',y'):\ x'+y' \le x+y,\ y' \le y\},\\ C_2(x,y) &= \{(x',y'):\ x'+y' \le x+y,\ x' \le x\},\\ S_t &= \{(x',y'):\ x'+y' \le t\}. \end{aligned}$$

The wave components of a V-Wiener process are the new processes (W_1, W_2) defined by the following:

$$\begin{aligned} W_1(x,y) &= W\big(C_1(x,y)\big) - \tfrac{1}{2}W(S_{x+y}) \text{ and}\\ W_2(x,y) &= W\big(C_2(x,y)\big) - \tfrac{1}{2}W(S_{x+y}); \end{aligned} \tag{3}$$

and

$$\begin{aligned} W_1(\infty,y) &= W\Big(\{(x,y'):\ y' \le y\}\Big) - \tfrac{1}{2}W(\mathbb{R}^2) \text{ and}\\ W_2(x,\infty) &= W\Big(\{(x',y):\ x' \le x\}\Big) - \tfrac{1}{2}W(\mathbb{R}^2). \end{aligned} \tag{3'}$$

REMARK 5. It may be noticed that the usual parametrization of a Wiener sheet by ordered pairs of real numbers is made by means of the sum of its wave components: $W_1(x,y) + W_2(x,y) = W(\{(x',y'):\ x' \le x,\ y' \le y\})$.

NOTATION 1. We call $\mathcal{A}_t$ the σ-field generated by

$$\{W_1(x,y) + W_2(x,y):\ x+y \le t\}.$$

The same field is generated either for $\{W_1(x,y): x+y \le t\}$ or $\{W_2(x,y): x+y \le t\}$.

THEOREM 2. *Let T be a stopping time with respect to the filtration $\{\mathcal{A}_t: t > 0\}$, and let $\mathcal{A}_{T+}$ be defined by $\mathcal{A}_{T+} = \{A: A \cap \{T < t\}$ is $\mathcal{A}_t$-measurable for all $t\}$.*
Then, conditional on $\{T < \infty\}$, the processes in $s \in \mathbb{R}^+$ $W_1(x+s, T-x) - W_1(x, T-x)$ and $W_2(T-y, y+s) - W_2(T-y, y)$ are centred and independent of $\mathcal{A}_{T+}$ for any x and any y, and the same happens with the random variables $W_1(\infty, T-x) - W_1(x, T-x)$ and $W_2(T-y, \infty) - W_2(T-y, y)$.

PROOF. This statement generalizes the Dynkin-Hunt theorem and can be verified by adapting a proof of this well-known result (see, e.g., [7]). □

NOTATION 2. Given the V-Wiener process W, we introduce the following associated processes and variables:

$$\begin{aligned}
m^+(t) &= \max\{W_1(x,y) - W_2(x',y'): x+y = x'+y' = t\};\\
m^+(\infty) &= \max\{W_1(\infty,y) - W_2(x,\infty): x,y \in \mathbb{R}\};\\
m^-(t) &= \max\{-W_1(x,y) + W_2(x',y'): x+y = x'+y' = t\};\\
m^-(\infty) &= \max\{-W_1(\infty,y) + W_2(x,\infty): x,y \in \mathbb{R}\};\\
M^+ &= \max_{t\in\mathbb{R}} m^+(t)\\
&= \max\{W_1(x,y) - W_2(x',y'): x,y,x',y' \in \mathbb{R}\};\\
M^- &= \max_{t\in\mathbb{R}} m^-(t)\\
&= \max\{-W_1(x,y) + W_2(x',y'): x,y,x',y' \in \mathbb{R}\};\\
M &= \max\{M^+, M^-\}.
\end{aligned}$$

THEOREM 3.

(i) *For any Wiener sheet with total variance equal to 1, and for each positive constant a, the estimate*

$$\mathbf{P}\{M^+ \ge a\} \le 8\Phi(-a) + 8a\varphi(a) \tag{4}$$

holds, with $\varphi(x) = (2\pi)^{-1/2}\exp(-\frac{1}{2}x^2)$ and $\Phi(x) = \int_{-\infty}^x \varphi(t)\,dt$.

(ii) *When the variance measure V is the joint distribution of two independent random variables, then*

$$\begin{aligned}
\mathbf{P}\{M^+ \ge a\} &\le 16\Phi(-a) - 8\Phi(-2a)\\
&\quad + 8\int_0^a (2\pi)^{-1/2}\exp\left(\frac{z^2}{2}\right)dz \int_z^{2a-z} v\exp\left(-\frac{v^2}{2}\right)\\
&\quad \times \exp\left(-\frac{1}{2}(4a^2 + v^2 - 4av)\right)dv.
\end{aligned} \tag{5}$$

(iii) *Both statements hold for M^- instead of M^+.*

PROOF. As a first step, we prove the inequality $\mathbf{P}\{M^+ \geq a\} \leq 2\mathbf{P}\{m^+(\infty) \geq a\}$. Define the stopping time $T = \sup\{t\colon m^+(s) < a \text{ for all } s \leq t\}$. When the event $\{M^+ \geq a\} = \{T < \infty\}$ occurs, then $m^+(T) = a$ and there exist two random variables X and Y such that $W_1(X, T-X) - W_2(T-Y, Y) = a$.

Since $D = (W_1(\infty, T-X) - W_1(X, T-X)) - (W_2(T-Y, \infty) - W_2(T-Y, Y))$ is centred and independent of $\mathcal{A}_{T^+}$, then

$$\begin{aligned}
&\mathbf{P}\{m^+(\infty) > a\} \\
&\quad = \mathbf{P}\{m^+(\infty) > a, W_1(X, T-X) - W_2(T-Y, Y) = a\} \\
&\quad \geq \mathbf{P}\Big\{\big(W_1(\infty, Y) - W_2(X, \infty)\big) > a, W_1(X, T-X) - W_2(T-Y, Y) = a\Big\} \\
&\quad = \mathbf{P}\{W_1(X, T-X) - W_2(T-Y, Y) = a, D > 0\} = \tfrac{1}{2}\mathbf{P}\{M^+ \geq a\}.
\end{aligned}$$

This ends the first step. The validity of (4) shall be proved by establishing the estimate

$$\mathbf{P}\{m^+(\infty) > a\} \leq 4\Phi(-a) + 4a\varphi(a).$$

Observe now that

$$\begin{aligned}
\mathbf{P}\{m^+(\infty) > a\} &= \mathbf{P}\Big\{\max_{x,y}\{W_1(\infty, y) - W_2(x, \infty)\} > a\Big\} \\
&\leq \mathbf{P}\Big\{\max_y W_1(\infty, y) > \tfrac{1}{2}a\Big\} + \mathbf{P}\Big\{\max_x\big(-W_2(x, \infty)\big) > \tfrac{1}{2}a\Big\} \\
&= 2\mathbf{P}\Big\{\max_t \xi(t) - \tfrac{1}{2}\xi(1) > \tfrac{1}{2}a\Big\},
\end{aligned}$$

because, if ξ is a standard Wiener process on $[0, 1]$, $W_1(\infty, y)$ and $-W_2(x, \infty)$ have both the distribution of $\xi(t) - \frac{1}{2}\xi(1)$, for $t = V(\mathbb{R} \times (-\infty, y])$ and $t = V((-\infty, x] \times \mathbb{R})$, respectively.

Now write

$$\begin{aligned}
&\mathbf{P}\Big\{\max_t \xi(t) - \tfrac{1}{2}\xi(1) > \tfrac{1}{2}a\Big\} \\
&\quad = \mathbf{E}\Big(\mathbf{E}\big(\mathbf{1}_{\{\max_t \xi(t) - (1/2)\xi(1) > (1/2)a\}} \mid \xi(1)\big)\Big) \\
&\quad = \int_{-\infty}^{\infty} \mathbf{P}\Big\{\max_t \xi(t) - \tfrac{1}{2}\xi(1) > \tfrac{1}{2}a \mid \xi(1) = z\Big\}\varphi(z)\,dz;
\end{aligned}$$

apply the reflection principle for a one-parameter Wiener process, to obtain

$$\mathbf{P}\Big\{\max_{0 \leq t \leq 1} \xi(t) > \tfrac{1}{2}(a+z) \mid \xi(1) = z\Big\} = \begin{cases} 1, & \text{if } |z| > a, \\ \varphi(a)/\varphi(z), & \text{if } |z| \leq a; \end{cases}$$

and replace in the integrand, to obtain

$$\mathbf{P}\Big\{\max_t \xi(t) - \tfrac{1}{2}\xi(1) > \tfrac{1}{2}a\Big\} = 2\Phi(-a) + 2a\varphi(a).$$

This ends the proof of (i).

As for (ii), we refer to the computations made in [2]. □

4. Conservative tests for goodness of fit. Given a random sample X_1, $X_2, \dots, X_n$ of the continuous distribution F, it is well known that the normalized empirical measure $B_n^F = \sqrt{n}(F_n - F)$ converges weakly to the F-Brownian bridge B^F.

Let us apply a representation theorem that states that there exist copies of B_n^F and B^F, for which we maintain the same notation, such that B_n^F converges to B^F, uniformly on $\mathbf{X} = \{C_1(x,y), C_2(x,y), S_t \colon x,y,t \in \mathbb{R}\}$. Such a theorem is an adaptation of 4.3.13 in [8] which we apply to the convergence of the wave components of normalized empirical processes to the wave components of the Brownian bridge, on the metric space of continuous pairs of functions $\mathbb{R}^2 \to \mathbb{R}$ with norm

$$\big\|\big(g_1(x,y), g_2(x,y)\big)\big\| = \max\left\{\sup_{x,y}|g_1(x,y)|, \sup_{x,y}|g_2(x,y)|\right\}.$$

After integration by parts, (1) can be written as

$$\mathcal{W}^F B(A) = \sqrt{3}\left(\int \widetilde{B}_A(\tau)\,d\widetilde{F}(\tau) + \int \widetilde{B}(\tau)\,d\widetilde{F}_A(\tau)\right), \tag{6}$$

and this exhibits the continuity of the mapping $\mathcal{W}^F$ that maps the processes with parameter set $\mathbf{X}$ onto processes with the same parameter set. This implies in particular that $W_n^F = \mathcal{W}^F B_n^F$ has the asymptotic distribution of the V-Wiener process $W = \mathcal{W}^F B^F$.

Equations (3) and (3′) can be used to define the wave components of any signed measure. In particular, when W, W_1 and W_2 are replaced by $W_n^{F_0}, W_{n,1}^{F_0}$ and $W_{n,2}^{F_0}$, respectively, the resulting equations define the wave components $W_{n,1}^{F_0}$ and $W_{n,2}^{F_0}$ of the transformed empirical process $W_n^{F_0} = \mathcal{W}^F B_n^{F_0}$.

When the null hypothesis H_0: $F = F_0$ is true, the continuous functionals of the wave components of the process $W_n^{F_0} = \mathcal{W}^{F_0} B_n^{F_0}$ also converge in distribution to the corresponding functionals of W. This is the case of the maxima:

$$\begin{aligned}
M_n^+ &= \max_{t\in\mathbb{R}} \max\left\{W_{n,1}^{F_0}(x,y) - W_{n,2}^{F_0}(x',y') \colon x+y = x'+y' = t\right\},\\
M_n^- &= \max_{t\in\mathbb{R}} \max\left\{-W_{n,1}^{F_0}(x,y) + W_{n,2}^{F_0}(x',y') \colon x+y = x'+y' = t\right\},\\
M_n &= \max\{M_n^+, M_n^-\}.
\end{aligned}$$

We propose the use of the statistics M_n^+, M_n^- or M_n for testing $F = F_0$, and verify below that a test with critical region M_n greater than some constant is consistent against any alternative $F \neq F_0$.

In order to test H_0 with level not greater than α, two applicable one-sided approximate critical regions are $\{M_n^+ > c(\alpha)\}$ and $\{M_n^- > c(\alpha)\}$, where $\mathbf{P}\{M^+ > c(\alpha)\} \le \alpha$, and $c(\alpha)$ is obtained by solving the equation

$$8\Phi\big(-c(\alpha)\big) + 8c(\alpha)\varphi\big(c(\alpha)\big) = \alpha, \tag{7}$$

after Theorem 3(i).

When F_0 is the particular probability distribution F_c introduced in Remark 3 after Theorem 1 (or any other sharing the property discussed in that remark), $c(\alpha)$ can be derived from the improved bound (5).

The two-sided critical region $\{M_n > c(\alpha)\}$ has level not greater than 2α.

THEOREM 4. *For any constant c, the rejection regions* $\{M_n > c\}$ *provide consistent tests for* H_0: $F = F_0$ *against any alternative* $F \neq F_0$, F_0 *continuous.*

PROOF. When F is different from F_0, since both are continuous, the signed measure $D = F - F_0$ is strictly positive on some region $R\{(x,y)\colon s < x+y < t, a < y < b\}$.

The normalized empirical process $B_n^{F_0} = \sqrt{n}(F_n - F_0)$ is the sum of the stochastically bounded term $B_n^F = \sqrt{n}(F_n - F)$ plus another one, namely, $\sqrt{n}D$, that tends to infinity on R under the alternative hypothesis. Now $W_n^{F_0}(R) = \mathcal{W}^{F_0}B_n^F(R) + \sqrt{n}\mathcal{W}^{F_0}D(R) \to \infty (n \to \infty)$, because the first term is stochastically bounded due to the continuity of $\mathcal{W}^{F_0}$, and

$$\mathcal{W}^{F_0}D(R) = \sqrt{3}\left(\int \widetilde{D}_R(\tau)\, d\widetilde{F}_0(\tau) + \int \widetilde{D}(\tau)\, d\widetilde{F}_{0R}(\tau) \right)$$

is strictly positive.

Let W_1 and W_2 denote the wave components of $W_n^{F_0}$. Notice that $W_n^{F_0}(R)$ is the double increment over R of W_1, namely, $W_n^{F_0}(R) = W_1(t-b,b) - W_1(t-a,a) - W_1(s-b,b) + W_1(s-a,a)$. This can be written as

$$\begin{aligned} W_n^{F_0}(R) = {}& \big(W_1(t-b,b) - W_1(t-a,a)\big) - \big(W_1(t-a,a) - W_1(t-a,a)\big) \\ & - \big(W_1(s-b,b) - W_2(s-b,b)\big) + \big(W_1(s-a,a) - W_2(s-b,b)\big). \end{aligned}$$

Since $W_n^{F_0}(R)$ tends to infinity, the same happens at least with one of the four terms in the previous sum, and this implies that M_n also tends to infinity. □

In designing an algorithm to carry out the computations required for performing the tests, one should include a transformation of the data to fit the canonical distribution F_c, in order to use the sharper bound (see Section 6). Empirical estimates of the power of the test in some particular cases are given in Section 7.

Table 1 provides some values of $c(\alpha)$ for the construction of the conservative critical regions described above.

5. Two-sample comparison tests. Let $(X_i, Y_i)_{i=1,2,\dots,m}$ be a sample of F, and let $(X_i, Y_i)_{i=m+1,\dots,m+n}$ be an independent sample of G, both unknown continuous bivariate distributions; let F_m and G_n denote the respective empirical measures. The conclusions of the next theorem can be applied for testing the null hypothesis H_0: $F = G$, by means of a functional of the difference of measures $F_m - G_n$.

TABLE 1

Upper bound for the level	Critical points for general distributions	Sharper critical points for product measures
20%	2.795	2.514
10%	3.057	2.807
5%	3.296	3.014
2.5%	3.515	3.160
1 %	3.721	3.344
1.25%	3.784	3.435
0.5%	3.974	3.568

In the following we shall replace the continuous marginal distribution $\widetilde{F}$ by a pure jump function, either $\widetilde{F}_m$ or the marginal of $H_{m,n}$ defined below. In these cases the integrand may have jumps at the same points as the integrator; hence, the integrals in (1) are not well defined. This ambiguity is avoided by choosing left-continuous versions for the integrands, and right-continuous ones for the integrators.

THEOREM 5. *Consider a family of samples* $(X_i, Y_i)_{i=1,2,\dots,m}$ *and* $(X_i, Y_i)_{i=m+1,\dots,m+n}$ *with sizes* m *and* n *tending to infinity in such a way that* $m/(m+n)$ *and* $n/(m+n)$ *have finite limits. Let* F_m *and* G_n *be the empirical measures, and let* $H_{m,n} = (mF_m + nG_n)/(m+n)$.

Then, under the null hypothesis $F = G$,

$$W_{m,n} = \mathcal{W}^{H_{m,n}}\left(\sqrt{\frac{mn}{m+n}}(F_m - G_n)\right)$$

has the asymptotic distribution of a V*-Wiener sheet (see Section* 3).

PROOF. Let $B_m^F = \sqrt{m}(F_m - F)$ and $B_n^G = \sqrt{n}(G_n - G)$ be the normalized empirical processes. Since B_m^F converges in distribution to the F-Brownian bridge, $\mathcal{W}^F B_m^F$ converges in distribution to a V-Wiener sheet. Let us verify first that the distributions of $\mathcal{W}^{F_m} B_m^F$ and $\mathcal{W}^{H_{m,n}} B_m^F$ also converge to the same limit, under the null hypothesis $F = G$. For that purpose, we apply again, as in the previous section, the representation theorem that states that there exist copies of F_m, G_n and the two F-Brownian bridges B^F and B^G, for which we maintain the same notation, such that F_m converges to F, G_n converges to G and B_m^F and B_n^G converge to B^F and B^G, respectively. All convergences are uniform on the family of sets $\mathbf{X}$.

Now, with the strongly convergent copies instead of the original processes, the difference

$$\begin{aligned}\mathcal{W}^{F_m} B_m^F(A) &- \mathcal{W}^F B_m^F(A)\\ &= \sqrt{3}\left(\int \widetilde{B}_{mA}^F(\tau)\, d\big(\widetilde{F}_m(\tau) - \widetilde{F}(\tau)\big) + \int \widetilde{B}_m^F(\tau)\, d\big(\widetilde{F}_{mA}(\tau) - \widetilde{F}_A(\tau)\big)\right)\end{aligned}$$

differs from

$$\sqrt{3}\left(\int \widetilde{B}_A^F(\tau)\,d\big(\widetilde{F}_m(\tau)-\widetilde{F}(\tau)\big)+\int \widetilde{B}^F(\tau)\,d\big(\widetilde{F}_{mA}(\tau)-\widetilde{F}_A(\tau)\big)\right)$$

by less than an arbitrary positive number ε, for m sufficiently large. These last integrals can be written as Wiener integrals with respect to the Brownian bridges by integrating by parts, namely,

$$-\sqrt{3}\left(\int \big(\widetilde{F}_m(\tau)-\widetilde{F}(\tau)\big)\,d\widetilde{B}_A^F(\tau)+\int \big(\widetilde{F}_{mA}(\tau)-\widetilde{F}_A(\tau)\big)\,d\widetilde{B}^F(\tau)\right),$$

and the uniform convergence of $\widetilde{F}_m$ and $\widetilde{F}_{mA}$ to $\widetilde{F}$ and $\widetilde{F}_A$ completes the argument. The previous considerations show that $\mathcal{W}^{F_m}B_m^F(A)$ and $\mathcal{W}^F B_m^F(A)$ have the same limit distribution. The verification that $\mathcal{W}^{H_{m,n}}B_m^F(A)$ and $\mathcal{W}^F B_m^F(A)$ have the same limit is made in the same way, provided the null hypothesis $F=G$ holds, which is required to ensure the convergence of $H_{m,n}$ to F.

The same procedure leads us to conclude that $\mathcal{W}^G B_n^G$, $\mathcal{W}^{G_n}B_n^G$ and $\mathcal{W}^{H_{m,n}}B_n^G$ also have the same limit distribution.

To end the proof, assume $F=G$ and write

$$\begin{aligned} W_{m,n} &= \mathcal{W}^{H_{m,n}}\left(\sqrt{\frac{mn}{m+n}}(F_m-G_n)\right)\\ &= \sqrt{\frac{n}{m+n}}\mathcal{W}^{H_{m,n}}B_m^F-\sqrt{\frac{m}{m+n}}\mathcal{W}^{H_{m,n}}B_n^G. \end{aligned}$$

Both terms are independent and each one tends to a multiple of a V-Wiener sheet, where V is the variance measure given by (2). Since the squares of the coefficients add to 1, the sum converges to a V-Wiener sheet.

Because the common value of F and G is assumed to be unknown, so is V. □

Extend the definition of wave components to the process $W_{m,n}$, as done in Section 4 with $W_n^{F_0}$, and define

$$\begin{aligned} M_{m,n}^+ &= \max_{t\in\mathbb{R}}\max\{W_{m,n,1}(x,y)-W_{m,n,2}(x',y'):\ x+y=x'+y'=t\},\\ M_{m,n}^- &= \max_{t\in\mathbb{R}}\max\{-W_{m,n,1}(x,y)+W_{m,n,2}(x',y'):\ x+y=x'+y'=t\},\\ M_{m,n} &= \max\{M_{m,n}^+,M_{m,n}^-\}. \end{aligned}$$

THEOREM 6. *The rejection regions* $\{|M_{m,n}|>c\}$ *based on the statistic* $M_{m,n}$ *provide consistent tests for* H_0: $F=G$ *against any alternative* $F\neq G$, F,G *continuous.*

The proof follows closely the one of Theorem 4, and will be omitted.

We propose the construction of approximate conservative tests, following a procedure similar to that in the case of goodness-of-fit tests, namely, to adopt the rejection regions $\{M_{m,n}^+>\text{constant}\}$, $\{M_{m,n}^->\text{constant}\}$ or $\{M_{m,n}>\text{constant}\}$, where the constant is obtained as in the above case.

6. Computing the test statistics.

6.1. *The comparison test.* Since $W_{m,n}$ is a discrete measure concentrated in the sample points (X_i, Y_i), its wave components $(W_{m,n,1}, W_{m,n,2})$ are sectionally constant. In fact, when the point (x,y) moves without crossing any of the lines $x+y = X_i + Y_i$ or $y = Y_j$, the first wave component $W_{m,n,1}(x,y)$ remains unchanged, and the same happens with $W_{m,n,2}(x,y)$ when (x, y) moves off the lines $x+y = X_i + Y_i$ and $x = X_j, i, j = 1, 2, \dots, m+n$.

Consequently,

$$\lim_{\delta \to 0^+} W_{m,n,1}(T_i - Y_j + \delta, Y_j) \quad \text{and} \quad \lim_{\delta \to 0^+} W_{m,n,2}(X_j, T_i - X_j + \delta),$$

$i, j = 1, 2, \dots, m+n$, describe completely the range of $W_{m,n,1}$ and $W_{m,n,2}$.

6.2. *The goodness-of-fit test.* The random measure $W_n^{F_0}$ can be written as

$$\begin{aligned} W_n^{F_0}(A) &= \sqrt{3n}\left(\int (\widetilde{F}_{nA} - \widetilde{F}_{0A})\, d\widetilde{F}_0 + \int (\widetilde{F}_n - \widetilde{F}_0)\, d\widetilde{F}_{0A} \right) \\ &= \sqrt{\frac{3}{n}}\Bigg[\sum_{(X_k, Y_k) \in A} (1 - \widetilde{F}_0(X_k)) \\ &\qquad - \sum_k F_0\Big(A \cap \{(x, y): x + y \le X_k + Y_k\}\Big)\Bigg]. \end{aligned} \tag{8}$$

For each t, $m_n^+(t)$ is the maximum of differences

$$W_{n,1}^{F_0}(x,y) - W_{n,2}^{F_0}(x',y') = W_n^{F_0}\big(C_1(x,y)\big) - W_n^{F_0}\big(C_2(x',y')\big)$$

on points (x,y) and (x',y') such that $x+y = x'+y' = t$, that is,

$$m_n^+(t) = \sup_y W_n^{F_0}\big(C_1(t-y,y)\big) - \inf_x W_n^{F_0}\big(C_2(x, t-x)\big).$$

A similar argument leads to $m_n^-(t) = \sup_x W_n^{F_0}\,(C_2(x,t-x)) - \inf_y W_n^{F_0}(C_1(t-y,y))$.

Since the first sum in (8) is a discrete measure concentrated on the sample points, while the second is absolutely continuous, $W_n^{F_0}(C_1(t-y, y))$ decreases as a function of y except possibly for positive jumps at $y = Y_k, k = 1, 2, \dots, n$. Therefore, the supremum of $W_n^{F_0}(C_1(t-y, y))$ is the right-hand limit of $W_n^{F_0}(C_1\ (t-y, y))$ at one of the jump points, and the infimum is one of the left-hand limits at the same points. For the same reason, the supremum and the infimum of $W_n^{F_0}(C_2(x, t-x))$ are, respectively, a right and a left limit of $W_n^{F_0}(C_2(x, t-x))$ at one of the points $x = X_k$. This can be applied to compute $m^+(t)$ and $m^-(t)$ for each fixed t.

It follows from (8) that the probability distribution F_0 prescribed in the null hypothesis H_0 has to be evaluated on a great number of sets. This is cumbersome, in general, and, for this reason, we propose to transform the sample in

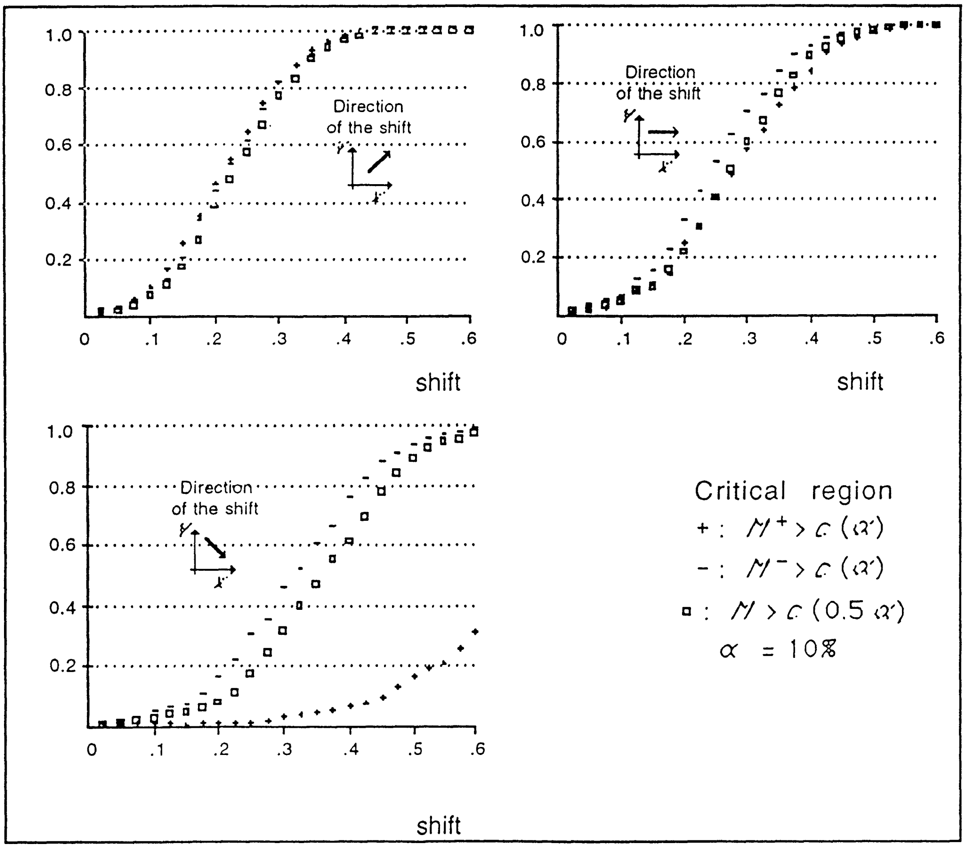

FIG. 1. *Estimated power of the comparison test for samples of Gaussian distribution with variance* $\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}$, *as a function of the relative shift; sizes* $m = n = 300$.

such a way that, under H_0, the new sample has the canonical distribution F_c (see Remark 3). After this transformation, no matter what was the original F_0, the computation is aided by an algorithm to compute F_c on the involved sets belonging to $\mathbf{X}$.

The choice of F_c allows, in addition, the use of the estimate in Theorem 3(ii).

For fixed (x,y) and t belonging to any of the intervals between points of $\{X_h + Y_h: h = 1,2,\ldots,n\}$, the partial derivative of $W_n^{F_0}(C_1(t-y,y)) - W_n^{F_0}(C_2(x,t-x))$ with respect to t reduces to $\sqrt{(3/n)}(x-y)f_c(t)$, which has constant sign. This means that the maximum of $W_n^{F_0}(C_1(t-y,y)) - W_n^{F_0}(C_2(x,t-x))$ occurs on $t = (X_h + Y_h)^+$ or $t = (X_h + Y_h)^-$, for some h in $\{1,\ldots,n\}$, and the same happens with the minimum.

Now, for those values of $t, W_n^{F_0}(C_1(t-y,y))$ is monotone as a function of y, except for possible jumps at $Y_1,\ldots,Y_n$. The behavior of $W_n^{F_0}(C_2(x,t-x))$ as a function of x is analogous. Hence, the evaluation of the test statistic is again reduced to a scanning of the sample points.

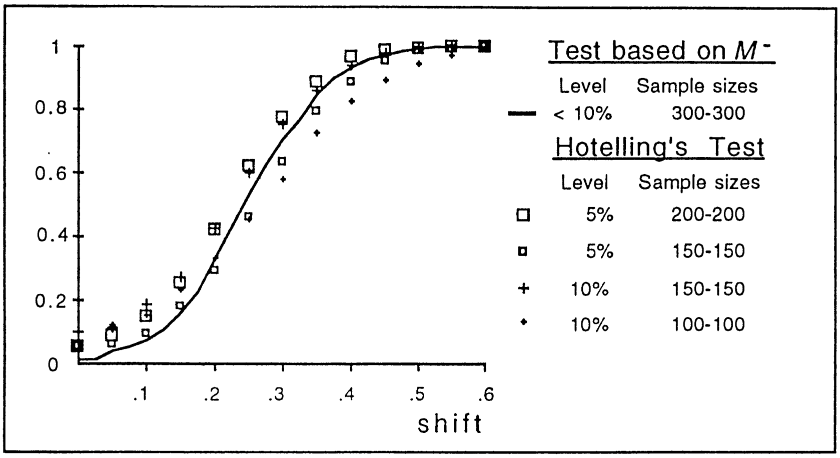

FIG. 2. *Comparison with Hotelling's test for Gaussian samples.*

7. Empirical estimation of the power. In the present section, we present a few figures intended to summarize the results of simulations of the application of the proposed tests. The number of replications has been set equal to 1000 when testing the null hypothesis, and 500 when the alternatives hold.

Figure 1 shows the performance of the comparison tests with critical regions $\{M > 3.296\}$, $\{M^+ > 3.057\}$ and $\{M^- > 3.057\}$ for Gaussian samples differing in location. The behaviour is seen to depend on the direction of the relative shift between samples.

The constants defining the critical regions have been chosen to ensure a level of 10% (see Table 1). The actual level is empirically estimated to be much smaller, but such an estimate depends on the shape of the parent distribution.

Figure 2 compares the performance of the test based on M^-, for shifts along the X axis, with Hotelling's T-test, both applied to isotropic Gaussian samples of equal sizes.

Figure 3 describes the ability of our test to detect differences between samples in two cases for which the performance of Hotelling's T is not expected to be good.

One case corresponds to the comparison between one sample of centred Gaussian distribution X, with variance $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, and other sample distributed as

$$\frac{1}{\sqrt{1-\alpha+\frac{1}{2}\alpha^2}}\left(\varepsilon\begin{pmatrix} 1 & 0 \\ 0 & 1-\alpha \end{pmatrix}X + (1-\varepsilon)\begin{pmatrix} 1-\alpha & 0 \\ 0 & 1 \end{pmatrix}X\right),$$

where ε is independent of X and assumes the values 0 or 1 with equal probabilities. A plain computation shows that the expectation and variance of Y are also $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. The comparison is made for different values of the anisotropy coefficient α.

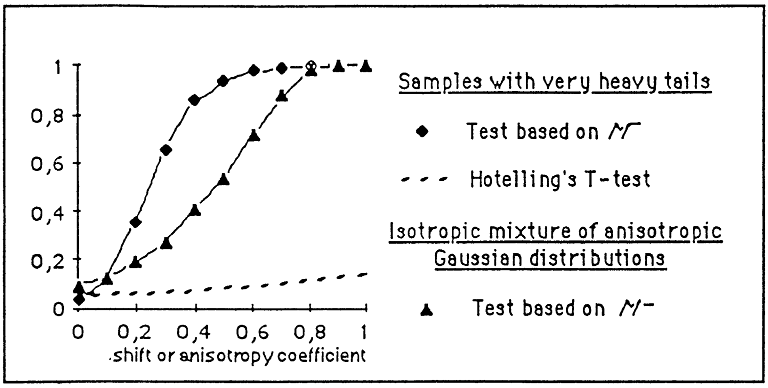

FIG. 3. *Comparison with Hotelling's test for non-Gaussian samples; upper bound of the level*; 10%.

The other case is the comparison between samples with different shifts, taken from an isotropic distribution with very heavy tails, namely,

$$\frac{U}{1-U}\begin{pmatrix}\cos(2\pi V)\\ \sin(2\pi V)\end{pmatrix},$$

with U and V independent on $[0,1]\times[0,1]$.

The goodness-of-fit tests behave very much like the comparison tests. Their actual level is closer to the upper bound used in the design, because the critical regions are determined by means of the sharper bound (5).

REFERENCES

[1] ADLER R. J. and BROWN, L. D. (1986). Tail behavior for suprema of empirical processes. *Ann. Probab.* **14** 1–30.
[2] CABAÑA, E. M. (1987). On two-parameter Wiener process and the wave equation. *Acta Cient. Venezolana* **38** 550–555.
[3] CABAÑA, E. M. (1990). A martingale characterization of two-parameter Wiener process. *Statist. Probab. Lett.* **10** 263–270.
[4] KHMALADZE, E. V. (1981). Martingale approach in the theory of goodness-of-fit tests. *Theory Probab. Appl.* **26** 240–257.
[5] KHAMALADZE, E. V. (1988). An innovation approach to goodness-of-fit tests in $\mathbf{R}_m$. *Ann. Statist.* **16** 1503–1516.
[6] KHAMALADZE, E. V. (1989). Theory of goodness of fit tests and scanning innovation martingales. Report BS-R8904, Center for Mathematics and Computer Science, Amsterdam.
[7] MCKEAN, H. P., JR.,(1968). *Stochastic Integrals*. Academic, New York.
[8] POLLARD, D. (1984). *Convergence of Stochastic Processes*. Springer, New York.

FACULTAD DE MATEMÁTICAS
DEPARTMENTO DE ESTADÍSTICA
UNIVERSIDAD DE BARCELONA
GRAN VIA 585
BARCELONA 08007
SPAIN

CENTRO DE MATEMÁTICA
FAC. DE CIENCIAS
AND FAC. DE CIENCIAS ECONÓMICAS Y ADM.
UNIVERSIDAD DE LA REPÚBLICA
EDUARDO A 'EVEOB 1139
MONTEVIDEO
URUGUAY