

UNIVERSITAT DE BARCELONA

TRANSFORMED EMPIRICAL PROCESSES
AND MODIFIED KOLMOGOROV-SMIRNOV TESTS
FOR MULTIVARIATE DISTRIBUTIONS

by

A. Cabaña and E. M. Cabaña

AMS Subject Classification: Primary: 62G30, 62G20, 62G10.
Secondary: 60G15



Mathematics Preprint Series No. 208

June 1996

Transformed Empirical Processes and Modified Kolmogorov-Smirnov Tests for multivariate distributions.

A. Cabaña *and E. M. Cabaña ††

Departament d'Estadística
Facultat de Matemàtiques
Universitat de Barcelona
Gran Vía 585
08007 Barcelona, Spain.

Centro de Matemática
Facultad de Ciencias
Universidad de la República
Eduardo Acevedo 1139
11200 Montevideo, Uruguay.

Abstract

A general way of constructing classes of goodness-of-fit tests for multivariate samples is presented. These tests are based on a random signed measure that plays the same role as the empirical process in the construction of the classical Kolmogorov-Smirnov tests. The resulting tests are consistent against any fixed alternative, and, for each sequence of contiguous alternatives, a test in each class can be chosen so as to optimize the discrimination of those alternatives.

Résumé

Processus empiriques transformés et tests de Kolmogorov-Smirnov modifiés, pour des distributions multivariantes.

On considère ici une méthode générale pour construire certaines classes de tests d'ajustement pour des échantillons multivariés. Ces tests sont basés sur des mesures aléatoires signées qui jouent le même rôle que le processus empirique pour la construction du test de Kolmogorov-Smirnov classique. Les tests résultants sont efficaces et consistents sous toute alternative, et, pour chaque suite d'alternatives contiguës, on peut choisir dans chaque classe un test que maximise la discrimination de ces alternatives.

AMS subject classification. Primary: 62G30, 62G20, 62G10 ; secondary: 60G15

*Permanent address: Departamento de Matemáticas, Instituto Venezolano de Investigaciones Científicas, Apartado 21827, Caracas 1020-A, Venezuela.

†Present address: Departamento de Estadística e Investigación Operativa, Facultad de Ciencias, Universidad de Valladolid, Prado de la Magdalena s/n, 47005 Valladolid, Spain.

††Joint work supported by Project 94/007, CONICYT (National Council for Scientific and Technological Research), Montevideo, Uruguay.



1 Introduction.

The empirical process associated to the sample $\{X\}$ of size 1 with respect to the probability distribution F_0 on $\mathcal{E} = \mathbf{R}^d$, $d \geq 1$, is the signed measure $b_x : A \mapsto b_x(A) = 1_{\{x \in A\}} - F_0(A)$, evaluated on the random point $x = X$.

It is well known that when X has distribution F_0 , b_X satisfies the properties

$\mathbf{E}b_X(A) = 0$ for every measurable set A , and

$\mathbf{E}b_X(A)b_X(B) = F_0(A \cap B) - F_0(A)F_0(B)$ for every measurable A, B , which mean that b_X has the same first- and second-order moments as an F_0 -Brownian bridge.

As n goes to infinity, the empirical process

$$b_n = \frac{1}{\sqrt{n}} \sum_{i=1}^n b_{X_i} \quad (1)$$

of the sample $\{X_1, \dots, X_n\}$ of i.i.d. variables with distribution F converges in distribution to an F_0 -Brownian bridge b^{F_0} when $F = F_0$. When $F \neq F_0$ instead, b_n behaves asymptotically as the sum of an F -bridge b^F and the deterministic term $\sqrt{n}(F - F_0)$ that tends to infinity in the supremum norm as $n \rightarrow \infty$. This gives a well known justification for the classical Kolmogorov-Smirnov test of $\mathcal{H}_0 : F = F_0$ (see [5]).

Let us introduce a family of alternatives $F^{(\tau)}$ (τ in a neighbourhood of 0) contiguous to $F^{(0)} = F_0$, with density $f^{(\tau)}$ with respect to F_0 , and such that there exists an $L^2(\mathcal{E}, dF_0)$ function k satisfying

$$\left\| \frac{1}{\tau} \left(\sqrt{f^{(\tau)}} - 1 \right) - \frac{k}{2} \right\|_{L^2} \rightarrow 0 \text{ as } \tau \rightarrow 0^+. \quad (2)$$

When (2) holds, the function k necessarily satisfies

$$\int k(x) dF_0(x) = 0, \quad (3)$$

and can be obtained as the $L^1(\mathcal{E}, dF_0)$ -limit of $(f^{(\tau)} - 1)/\tau$ as $\tau \rightarrow 0$.

We shall assume that we are especially interested in detecting the sequence of alternatives $\mathcal{H}_n : F = F_{(n)} = F^{(\delta/\sqrt{n})}$, where n is the sample size, and δ is a fixed parameter introduced for further convenience of notation.

Under \mathcal{H}_n , $b_n(A)$ converges in distribution to the sum of b^{F_0} plus the deterministic drift

$$\lim_{n \rightarrow \infty} \sqrt{n} \mathbf{E}b_{X_{\delta/\sqrt{n}}}(A) = \delta \frac{\partial \mathbf{E}b_{X_\tau}(A)}{\partial \tau} \Big|_{\tau=0} = \delta \int_A k(x) dF_0(x)$$

(c.f. [13]). That drift measures the sensitivity of the test based on b_n with respect to \mathcal{H}_n .

The aim of this paper is the construction of goodness-of-fit tests based on signed measures \hat{w}_x ($x \in \mathcal{E}$) such that

- when x is replaced by a random variable X , the resulting measure $\hat{w}_X(A)$ evaluated on any measurable set A is a random variable,
- the measure $\hat{w}_n = n^{-1/2} \sum_{i=1}^n \hat{w}_{X_i}$ associated to the sample has some normalized limit distribution under \mathcal{H}_0 , and also
- the asymptotic distributions of \hat{w}_n under \mathcal{H}_0 and under \mathcal{H}_n differ as much as possible.

These random measures will play, in the construction of our tests, the same role as the empirical processes, in classical Kolmogorov-Smirnov tests. Each of them will be called a *Transformed Empirical Proces* (in short: *TEP*).

For technical convenience, we shall impose that the second-order moments of \hat{w}_X coincide with those of a Wiener process, normalized by the requirement that its total variance be one. Thus, under a Central Limit property to be established, the limit distribution of $\hat{w}_n = n^{-1/2} \sum_{i=1}^n w_{X_i}$ will be a normalized Gaussian process with independent increments.

As a consequence, we pose ourselves the problem of finding signed measures \hat{w}_x verifying

- (i) $\mathbf{E}\hat{w}_X(A) = 0$, for all measurable sets A and the random variable X distributed as F_0 ,
- (ii) $\mathbf{E}\hat{w}_X(A)\hat{w}_X(B) = V(A \cap B)$, for all measurable sets A, B , the random variable X distributed as F_0 , and some probability V on \mathcal{E} , and
- (iii) $\left. \frac{\partial \mathbf{E}\hat{w}_{X_\tau}(\mathcal{E})}{\partial \tau} \right|_{\tau=0} \geq \left. \frac{\partial \mathbf{E}\hat{w}'_{X_\tau}(\mathcal{E})}{\partial \tau} \right|_{\tau=0}$ for any \hat{w}'_x satisfying the analogue of (i) and (ii), and the random variable X_τ distributed as $F^{(\tau)}$.

Conditions (i) and (ii) do not require further explanation. As for Condition (iii), notice that it imposes the drift of our TEP to be as large as possible on the set \mathcal{E} of maximum variance. The adequacy of such heuristic requirement will be verified a posteriori, from the properties of the resulting tests.

Let us point out that the problem of finding processes \hat{w}_X satisfying (i), (ii), (iii), is implicitly solved in [1], [2], where goodness-of-fit tests asymptotically efficient are obtained for univariate samples. In this article, we solve it explicitly, extend the solution to a multivariate context, and, for each family of alternatives as in (2), provide multivariate consistent tests, efficient for that family.

In the sequel, \mathcal{E} is always set equal to \mathbf{R}^d , and the examples in §8 are developed for $d = 1$ and $d = 2$, but most results apply to more general spaces, particularly the extension of isometries described in §7.

2 A formal L^2 construction of the TEP.

2.1 Isometries and associated TEPs.

Let us assume for simplicity that the probability V appearing in (ii) is absolutely continuous with respect to F_0 , and call a^2 its density, that is:

$$V(A) = \int_A a^2(x) dF_0(x), \quad \int a^2(x) dF_0(x) = 1.$$

We shall assume further that a is F_0 -a.e. different from zero, that is, F_0 and V are absolutely continuous with respect to each other.

We shall denote $\langle \cdot, \cdot \rangle$ and $\langle \cdot, \cdot \rangle_V$ the inner products in $L^2(\mathcal{E}, dF_0)$ and $L^2(\mathcal{E}, dV)$ respectively. From an orthonormal basis $\Psi = (\psi_0 = 1, \psi_1, \psi_2, \dots)$ of $L^2(\mathcal{E}, dF_0)$, we may construct the new sequence of functions

$$\Psi^V = (\psi_i^V)_{i=0,1,\dots}, \quad \psi_i^V(x) = \psi_i(x)/a(x),$$

which is an orthonormal basis for $L^2(\mathcal{E}, dV)$.

Conditions (i) and (ii) can be replaced by the requirement that the Fourier coefficients

$$c_i(x) = \int \psi_i^V(y) d\hat{w}_x(y)$$

of $d\hat{w}_x/dV$ with respect to Ψ^V satisfy

$$\mathbf{E}c_i(X) = \langle c_i, 1 \rangle = \mathbf{E} \int \psi_i^V(x) d\hat{w}_X(x) = 0$$

and

$$\mathbf{E}c_i(X)c_j(X) = \langle c_i, c_j \rangle = \mathbf{E} \int \psi_i^V(y) d\hat{w}_X(y) \int \psi_j^V(z) d\hat{w}_X(z)$$

$$= \int \psi_i^V(y) \psi_j^V(y) dV(y) = \langle \psi_i^V, \psi_j^V \rangle_V = \delta_{i,j}.$$

Consequently, $C = (c_i)_{i=0,1,\dots}$ is required to be an orthonormal system in $L^2(\mathcal{E}, dF_0)$ with all its elements orthogonal to the constant 1, in order that (i) and (ii) hold.

Expand now $\hat{w}_{X_\tau}(\mathcal{E}) = \sum_i c_i(X_\tau) \langle \psi_i^V, 1 \rangle$, and compute

$$\begin{aligned} \left. \frac{\partial \mathbf{E} \hat{w}_{X_\tau}(\mathcal{E})}{\partial \tau} \right|_{\tau=0} &= \lim_{\tau \rightarrow 0} \sum_i \frac{\mathbf{E}(c_i(X_\tau) - c_i(X))}{\tau} \langle \psi_i^V, 1 \rangle_V = \\ &= \sum_i \int c_i(x) k(x) dF_0(x) \langle \psi_i, a \rangle = \sum_i \langle c_i, k \rangle \langle \psi_i, a \rangle = \sum_i k_i^C a_i, \end{aligned} \quad (4)$$

where k_i^C , a_i are the i -th Fourier coefficients of k and a with respect to the orthonormal systems C and Ψ , respectively.

Let us introduce now the isometry \mathcal{T} that maps Ψ onto C , that is,

$$\mathcal{T} : L^2(\mathcal{E}, dF_0) \rightarrow L^2(\mathcal{E}, dF_0), \quad \psi_i \mapsto \mathcal{T} \psi_i = c_i, \quad i = 1, 2, \dots \quad (5)$$

The last term in (4) equals $\langle k, \mathcal{T} a \rangle$, and therefore Condition (iii) holds when C and a are chosen to minimize the angle between $\mathcal{T} a$ and k . This is accomplished by selecting $\mathcal{T} a = k / \|k\|$, provided the span of C contains k . This will be ensured by imposing that the span of C , that is, the range of \mathcal{T} , be the orthogonal complement 1^\perp of 1 in $L^2(\mathcal{E}, dF_0)$.

The conclusions obtained so far can be summarized in the following statement:

Proposition 1 *The measure \hat{w}_x with formal Fourier expansion*

$$\hat{w}_x(A) = \sum_{i=0}^{\infty} c_i(x) \langle \hat{a} 1_A, \psi_i \rangle \quad (6)$$

satisfies (i), (ii), (iii), when $\Psi = (\psi_i)_{i=1,2,\dots}$ is an orthonormal basis of $L^2(\mathcal{E}, dF_0)$, \mathcal{T} is an isometry on $L^2(\mathcal{E}, dF_0)$ with range 1^\perp , $C = (c_i)_{i=1,2,\dots}$ is the image of Ψ by \mathcal{T} and the function \hat{a} has the property $\mathcal{T} \hat{a} = k / \|k\|$.

In that case, the objective of our optimization has the value

$$\left. \frac{\partial \mathbf{E} \hat{w}_{X_\tau}(\mathcal{E})}{\partial \tau} \right|_{\tau=0} = \langle k, \mathcal{T} \hat{a} \rangle = \|k\|. \quad (7)$$

Remark 1.1 *The preceding proposition gives a formal solution to our problem for each orthonormal basis of $L^2(\mathcal{E}, dF_0)$ and each isometry \mathcal{T} with range 1^\perp .*

Remark 1.2 By replacing $c_i(x)$ by $\mathcal{T}\psi_i(x)$ in (6), and using the linearity of \mathcal{T} we obtain

$$\hat{w}_\bullet(A) = \mathcal{T} \sum_{i=0}^{\infty} \langle \hat{a}1_A, \psi_i \rangle \psi_i = \mathcal{T}(\hat{a}1_A), \quad (8)$$

and this implies in particular that \hat{w}_x depends only on the isometry \mathcal{T} , but not on the orthonormal basis Ψ .

Let us finally introduce the notation

$$w_\bullet^{(a,\mathcal{T})}(A) = \mathcal{T}(a1_A) \quad (9)$$

for the measure associated to the isometry \mathcal{T} and the score function $a \in L^2(\mathcal{E}, dF_0)$.

After Remark 1.2, we may reformulate our previous statement:

Proposition 2 *The measure (9) satisfies (i), (ii), (iii) and (7), when \mathcal{T} is an isometry on $L^2(\mathcal{E}, dF_0)$ with range 1^\perp , and the score function is chosen as $\hat{a} = \mathcal{T}^{-1}k/\|k\|$.*

2.2 Constructing the TEP as a stochastic integral with respect to the empirical process.

From equation (9) and the orthogonality of the range of \mathcal{T} with respect to 1, the expression

$$\begin{aligned} w_X^{(a,\mathcal{T})}(A) &= \mathcal{T}(a1_A)(X) = \int \mathcal{T}(a1_A) d1_{\{\bullet \leq X\}} \\ &= \int \mathcal{T}(a1_A) d(1_{\bullet \leq X} - F_0) = \int \mathcal{T}(a1_A) db_X \end{aligned} \quad (10)$$

follows.

From (10) we derive an expression for

$$w_n^{(a,\mathcal{T})} = \frac{1}{\sqrt{n}} \sum_{i=1}^n w_{X_i}^{(a,\mathcal{T})} \quad (11)$$

in terms of b_n , namely:

$$w_n^{(a,\mathcal{T})}(x) = \int \mathcal{T}(a1_x) db_n. \quad (12)$$

2.3 Transformed empirical process in $\mathcal{E} = \mathbf{R}^d$.

We adopt the expressions (10) and (12) as the definition of the transformed empirical process:

Definition 1 *The transformed empirical process of the sample $\{X_1, X_2, \dots, X_n\}$ of \mathcal{E} -valued random variables, associated to the probability distribution F_0 on \mathcal{E} , the isometry \mathcal{T} on $L_2 = L_2(\mathcal{E}, dF_0)$ with range equal to the orthogonal complement 1^\perp of the constant function 1, and the L_2 -function a with $\|a\| = 1$, is*

$$w_n^{(a, \mathcal{T})}(A) = \int \mathcal{T}(a1_A) db_n. \quad (13)$$

No attention has been paid to the convergence of the Fourier expansions involved in the arguments that motivated Definition 1, but straightforward computation of moments shows that properties §1 (i), (ii) hold when \hat{w} is replaced by $w_X^{(a, \mathcal{T})}$ or $w_n^{(a, \mathcal{T})}$. Moreover

$$\left. \frac{\partial \mathbf{E} w_{X_\tau}^{(a, \mathcal{T})}(\mathcal{E})}{\partial \tau} \right|_{\tau=0} = \int k(x) (\mathcal{T}a)(x) dF_0(x),$$

so that (iii) holds with $w_X^{(\hat{a}, \mathcal{T})}$ substituted for \hat{w}_X , where $\hat{a} = \mathcal{T}^{-1}k/\|k\|$.

We show in section §3 that, under suitable assumptions on \mathcal{T} , a , a Central Limit Theorem holds for the TEP so defined.

3 Asymptotic properties of the TEP.

Let us consider the TEPs $(w_n^{(a, \mathcal{T})})_{n=1,2,\dots}$ with respect to F_0 constructed over triangular arrays of i.i.d. variables $X_{n,1}, X_{n,2}, \dots, X_{n,n}$ with distribution function $F^{(n)}$. We describe separately the limit properties of the TEPs for $n \rightarrow \infty$ in two cases, when $F^{(n)} = F \neq F_0$ is the same fixed distribution for all samples, and when $F^{(n)}$ is a sequence contiguous to F_0 .

3.1 Unboundedness under fixed alternatives.

The expectation of $w_X^{(a, \mathcal{T})}(A)$ when X has distribution F is

$$\int \mathcal{T}(a1_A)(x) d(F(x) - F_0(x)) = \int \mathcal{T}(a1_A)(x) dF(x) = \langle \mathcal{T}(a1_A), dF/dF_0 \rangle.$$

Let us call g the projection of dF/dF_0 on the range of \mathcal{T} . Then $\mathbf{E} w_X^{(a, \mathcal{T})}(A) = \langle a1_A, g \rangle = \int_A a g dF_0$. From this expression we are lead to the following.

Theorem 1 *If $\int_A a g dF_0 \neq 0$ for some A in a family of sets \mathcal{J} , then*

$$\lim_{n \rightarrow \infty} \sup_{A \in \mathcal{J}} |w_n(A)| = +\infty \text{ a.s.} \quad (14)$$

Corollary 1.1 *When $\mathcal{J} = \{(a, b] : a, b \in \mathbf{R}^d\}$ or $\mathcal{J} = \{(-\infty, b] : b \in \mathbf{R}^d\}$, then (14) holds.*

The conclusion follows from the assumptions that \mathcal{T} has range 1^\perp and a is a.e. nonvanishing.

3.2 Gaussian limit under the null hypothesis and contiguous alternatives.

Replacing the empirical process b_n in (13) by the F_0 -Brownian bridge b , one obtains a Wiener process $w^{(a, \mathcal{T})}(A) = \int \mathcal{T}(a 1_A) db$, indexed on A (A in a given family of sets \mathcal{J}). Following Ossiander ([14]), the convergence in distribution of the transformed empirical process $w_n^{(a, \mathcal{T})}$ to $w^{(a, \mathcal{T})}$ is guaranteed, under the null hypothesis \mathcal{H}_0 , by the assumptions

$$\mathcal{T}(a 1_A) \leq G \text{ for some } G \in L_2(\mathcal{E}, dF_0) \text{ and all } A \in \mathcal{J}, \quad (15)$$

of uniform boundedness of the family

$$\mathcal{G} = \{\mathcal{T}(a 1_A) : A \in \mathcal{J}\}, \quad (16)$$

and

$$\int_0^1 \sqrt{\log N_{[]}^{(2)}(\varepsilon, \mathcal{G}, F_0)} d\varepsilon < \infty \quad (17)$$

about the boundedness of the $L_2(F_0)$ -metric entropy with bracketing $N_{[]}^{(2)}$, which is defined by:

$$N_{[]}^{(2)}(\varepsilon, \mathcal{G}, F_0) = \min\{\kappa : \text{there exist sets } \mathcal{U} \text{ and } \mathcal{L} \text{ with cardinal } \kappa \text{ of } L_2(F_0)\text{-functions such that for each } f \in \mathcal{G} \text{ there exist } u \in \mathcal{U} \text{ and } l \in \mathcal{L} \text{ such that } l < f < u \text{ and } \|u - l\|^2 \leq \varepsilon^2\}. \quad (18)$$

As for the asymptotic distribution of $w_n^{(a, \mathcal{T})}(A)$ under \mathcal{H}_n , it follows from Le Cam Third Lemma and the assumption (2) (see [11],[13]) that it is the same as under \mathcal{H}_0 plus the bias $\delta \int_A k(\mathcal{T}a) dF_0$.

The following statement summarizes the asymptotic behaviour indicated above. The assumptions in (i) must be verified for each particular isometry, as will be done below for the examples in §8.

Theorem 2 Let $\{X_1, \dots, X_n\}$ be a sample of \mathcal{E} -valued i.i.d. random variables with distribution F and $w_n^{(a, \mathcal{T})}(A)$ the transformed empirical process of that sample associated to the probability distribution F_0 on \mathcal{E} , the isometry \mathcal{T} on $L_2 = L_2(\mathcal{E}, dF_0)$ with range orthogonal to the constant function 1, and the L_2 -function a ($\|a\| = 1$), as introduced in Definition 1.

- (i) When the assumptions (15) and (17) hold, $\{w_n^{(a, \mathcal{T})}(A) : A \in \mathcal{J}\}$ converges in distribution to the Wiener process $\{w^{(a, \mathcal{T})}(A) : A \in \mathcal{J}\}$ with covariance function

$$\mathbf{E}w^{(a, \mathcal{T})}(A)w^{(a, \mathcal{T})}(B) = \int_{A \cap B} a^2 dF_0,$$

under the null hypothesis $\mathcal{H}_0 : "F = F_0"$.

- (ii) When, in addition, the family of probabilities $F^{(\tau)}$ on \mathcal{E} with density $f^{(\tau)}$ with respect to F_0 satisfies (2), then $\{w_n^{(a, \mathcal{T})}(A) : A \in \mathcal{J}\}$ converges in distribution to $\{w^{(a, \mathcal{T})}(A) + \delta \int_A k(\mathcal{T}a) dF_0 : A \in \mathcal{J}\}$, under the sequence of alternatives $\mathcal{H}_n : "F = F^{(\delta/\sqrt{n})}"$.

These results justify the test procedure described in next section.

4 The goodness-of-fit tests.

We propose to test $\mathcal{H}_0 : "F = F_0"$ by means of the critical region

$$\sup_{A \in \mathcal{J}} |w_n^{(\hat{a}, \mathcal{T})}(A)| > c(\alpha)$$

with $\mathcal{T}\hat{a} = k/\|k\|$ and $c(\alpha)$ such that

$$\mathbf{P}\{\sup_{A \in \mathcal{J}} |w^{(\hat{a}, \mathcal{T})}(A)| > c(\alpha)\} \leq \alpha,$$

where α is an upper bound of the asymptotic level desired for the test.

The family \mathcal{J} of measurable subsets of \mathcal{E} is chosen rich enough to generate the σ -field of measurable sets in \mathcal{E} , but not so large that $\sup_{A \in \mathcal{J}} |w^{(\hat{a}, \mathcal{T})}(A)|$ be unbounded. It is assumed that \mathcal{J} contains \mathcal{E} . Moreover, the compromise in choosing \mathcal{J} is that a CLT

$$\lim_{n \rightarrow \infty} \sup_{A \in \mathcal{J}} |\hat{w}_n(A) - w^V(A)| = 0 \quad (19)$$

holds for copies \hat{w}_n of $w_n^{(\hat{a}, \mathcal{T})}$ and a V -Wiener process w^V , and also

$$\text{for } F \neq F_0, \text{ there is } A \in \mathcal{J} \text{ such that } F(A) \neq F_0(A). \quad (20)$$

We shall choose for \mathcal{J} either the family of all generalized intervals

$$\mathcal{J}_1 = \{\times_{i=1}^d (p_i, q_i] : -\infty \leq p_i < q_i \leq \infty, i = 1, 2, \dots, d\}$$

or the family of translations of the negative orthant

$$\mathcal{J}_0 = \{\times_{i=1}^d (-\infty, q_i] : q_i \leq \infty, i = 1, 2, \dots, d\}.$$

Since $\mathcal{E} \in \mathcal{J}$, the asymptotic power of the test for \mathcal{H}_n is greater than

$$\mathbf{P}\{w^{(\hat{a}, \mathcal{T})}(\mathcal{E}) + \delta \|k\| > c(\alpha)\} = \Phi(-c(\alpha) + \delta \|k\|).$$

An argument as the one used in [2] or [4] shows that when the level α and the probability β of type II error are sufficiently small, the asymptotic relative efficiency of our test with respect to the Maximum Likelihood Ratio Test can be chosen as close to one as desired. In this sense, the test has optimum efficiency.

For other values of α and β there is a reduction in the efficiency, due to the use as the test variable of the supremum of $w_n^{(\hat{a}, \mathcal{T})}$ over a family of sets instead of the value of the same process on \mathcal{E} . This is the price to be paid for the consistency against all alternatives.

5 On the construction of isometries and their associated transformed empirical processes.

Let us notice that any isometry \mathcal{T} on $L^2(\mathcal{E}, dF_0)$ with range contained in $1^\perp = \{f \in L^2(\mathcal{E}, dF_0) : \langle f, 1 \rangle = 0\}$ induces an orthonormal system $\Psi(\mathcal{T}) = (1, \mathcal{T}1, \mathcal{T}^2 1, \dots, \mathcal{T}^i 1, \dots)$.

Conversely, given any orthonormal system $\Psi = (1, \psi_1, \psi_2, \dots)$, the linear transformation $\mathcal{T}(\Psi)$ on $L^2(\mathcal{E}, dF_0)$ that maps each ψ_i onto ψ_{i+1} , $i = 0, 1, \dots$, is an isometry with range in 1^\perp . In addition, when Ψ is a basis, the range is 1^\perp and therefore it contains any function k with the property (3).

These observations imply that the isometries needed for the construction of TEPs may possibly be obtained from known orthonormal systems in $L^2(\mathcal{E}, dF_0)$.

In subsection §5.1 we indicate the analytical form of the shift \mathcal{T}_L induced by the Laguerre Polynomials on $L^2(\mathbf{R}^+, e^{-x} dx)$. A simple analytical

property of \mathcal{T}_L reflected in equation (23) gives an alternative way to obtain this isometry.

For a given orthonormal system Ψ , it is not easy in general to find the explicit analytical expression of the shift operator that maps each ψ_n onto ψ_{n+1} . We describe it for a particular example in §5.2: the orthonormal system of Tchebyshev Polynomials on $L^2_T = L^2([-1, 1], dx/(\pi\sqrt{1-x^2}))$.

The normalized Hermite polynomials $h_n(x) = H_n(x)/\sqrt{n!}$, $n = 0, 1, \dots$, where $e^{tx-t^2/2} = \sum_{n=0}^{\infty} H_n(x) \frac{t^n}{n!}$ ([17]), constitute an orthonormal basis of $L^2_H = L^2(\mathbf{R}, e^{-x^2/2} dx/\sqrt{2\pi})$. An explicit writing of the shift that maps h_n onto h_{n+1} in terms of integrals of the Poisson kernel (see [12] for the form of the inverse shift), leads to cumbersome computations so that the alternative approach, namely, the generalization of equation (23), is adopted in §5.3 to construct an isometry \mathcal{T}_H on that space.

By means of changes of variables, the isometries in §5.1 §5.2 and §5.3 induce others on different L^2 spaces on \mathbf{R} , as described in §6, and the latter ones can be used in the construction of new isometries on L^2 spaces on \mathbf{R}^d as shown in §7.

The *Laguerre shift* (29) has been used many times in probability theory and statistical inference: for instance, let us mention that Brownian bridge is constructed from Wiener process by means of its inverse in [9], that Efron and Johnstone [6] and Ritov and Wellner [16] use (29) and its inverse in connection with hazard rates (Some properties of both isometries are described in [7]), and that E.V.Khmaladze ([10]) introduced (29) in statistical inference, specially emphasizing some associated martingale properties. We use it as our main example because of its very simple analytical expression, but do not apply any martingale approach.

5.1 The Laguerre shift.

The well known Laguerre polynomials $\{L_n : n = 0, 1, \dots\}$ are an orthonormal basis of $L^2(\mathbf{R}, dF_0)$, with $F_0(x) = F_0([0, x]) = 1 - e^{-x}$, and they are obtained by means of the iterated application of the mapping

$$h(x) = \mathcal{T}_L g(x) = g(x) - \int_0^x g(t) dt \quad (21)$$

to the first element of the basis, the polynomial of degree zero $L_0 = 1$:

$$L_n(x) = (-1)^n \mathcal{T}_L^n L_0(x) \quad (22)$$

(see, for instance, [17], as a general reference).

This implies that \mathcal{T}_L is an isometry and, for each $g \in L^2(\mathbf{R}, dF_0)$, $\mathcal{T}_L g$ is orthogonal to the constant 1.

Both properties of \mathcal{T}_L follow immediately after plain calculations, and, conversely, imply (22). The clue to show that \mathcal{T}_L is an isometry is the equation

$$\mathcal{T}_L g(x) \mathcal{T}_L h(x) f_0(x) = g(x) h(x) f_0(x) - \frac{d}{dx} \left(f_0(x) \int_0^x g(t) dt \int_0^x h(t) dt \right) \quad (23)$$

and the fact that $\lim_{x \rightarrow \infty} f_0(x) \int_0^x g(t) dt \int_0^x h(t) dt = 0$.

5.2 The Tchebychev shift.

The Tchebyshev Polynomials $T_0(x) = 1$, $T_n(x) = \sqrt{2} \cos(n \arccos x)$, $n = 1, 2, \dots$ are an orthonormal basis of $L^2_T = L^2([-1, 1], dx/(\pi\sqrt{1-x^2}))$. The isometry \mathcal{T}_T that maps each T_n onto T_{n+1} can be described as follows:

Given $u(x) \in L^2_T$, let us assume first that $\langle u, 1 \rangle = \int_{-1}^1 u(x) dx / (\pi\sqrt{1-x^2}) = 0$. Perform the change of variable $\tilde{u}(|t|) = u(\cos t)$, $-\pi < t \leq \pi$, obtain an analytic function $h^{(u)}$ on $D = \{\zeta = re^{it} : |\zeta| < 1\}$ such that $\tilde{u}_r(t) = \Re h^{(u)}(re^{it})$ converges in $L^2((-\pi, \pi], dt)$ -norm to $\tilde{u}(|t|)$ as $r \rightarrow 1$ and take the $L^2((-\pi, \pi], dt)$ -limit

$$v(\cos t) = \lim_{r=|\zeta| \rightarrow 1} \Re \zeta h^{(u)}(\zeta).$$

This limit is the image $v = \mathcal{T}_T u$ of u .

The function $h^{(u)}$ is determined up to an imaginary additive constant, and the one satisfying $\Im h^{(u)}(0) = 0$ is obtained by integrating the Poisson kernel:

$$\begin{aligned} h^{(u)}(\zeta) &= \frac{1}{2\pi i} \int_C u(\Re z) \frac{z + \zeta}{z - \zeta} \frac{dz}{z} \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} u(\cos s) \left(\frac{1 - r^2 + 2ir \sin(s-t)}{1 - 2r \cos(s-t) + r^2} \right) ds = u(\cos t) \\ &\quad + \frac{1}{2\pi} \int_{-\pi}^{\pi} (u(\cos s) - u(\cos t)) \left(\frac{1 - r^2 + 2ir \sin(s-t)}{1 - 2r \cos(s-t) + r^2} \right) ds. \end{aligned}$$

Taking the limit in

$$\begin{aligned} \Re \zeta h^{(u)}(\zeta) &= r \cos t u(\cos t) \\ &\quad + \frac{1}{2\pi} \int_{-\pi}^{\pi} (u(\cos s) - u(\cos t)) \left(\frac{r(1-r^2) \cos t - 2r^2 \sin(s-t) \sin t}{1 - 2r \cos(s-t) + r^2} \right) ds, \end{aligned}$$

one obtains

$$\begin{aligned}\mathcal{T}_T u(\cos t) &= \lim_{r \rightarrow 1} \Re \zeta h^{(u)}(\zeta) \\ &+ \frac{1}{2\pi} \int_{-\pi}^{\pi} (u(\cos s) - u(\cos t)) \left(\frac{-\sin(s-t) \sin t}{1 - \cos(s-t)} \right) ds \\ &= \cos t u(\cos t) - \frac{\sin(t)}{2\pi} \int_{-\pi}^{\pi} \frac{u(\cos s) - u(\cos t)}{\tan((s-t)/2)} ds.\end{aligned}$$

Any $u(x) \in L^2_T$ can be written as the sum of the constant $\langle u, 1 \rangle$ and the function $u(x) - \langle u, 1 \rangle$, orthogonal to 1. The image of $\langle u, 1 \rangle$ by \mathcal{T}_T is $\sqrt{2}\langle u, 1 \rangle x$, and the image of $u(x) - \langle u, 1 \rangle$ is obtained as indicated above.

It may be noticed, in particular, that the images by \mathcal{T}_T of functions such as the product $a(\cdot)1_x(\cdot)$ of a continuous nonvanishing function a times the indicator function of a half line are not bounded. This makes this isometry useless for our present purposes.

5.3 An isometry associated to normal distribution.

Let $\varphi(x) = e^{-x^2/2}/\sqrt{2\pi}$ denote the normal density. The analogue to (23)

$$\begin{aligned}& \mathcal{S}g(x)\mathcal{S}\langle(x)\varphi(x) \\ &= g(x)h(x)\varphi(x) - \frac{d}{dx} \left(\sqrt{|x|} \int_0^x \sqrt{|t|} g(t) dt \sqrt{|x|} \int_0^x \sqrt{|t|} h(t) dt \varphi(x) \right),\end{aligned}\quad (24)$$

where

$$\mathcal{S} : f(x) \mapsto f(x) - \operatorname{sgn}x \sqrt{|x|} \int_0^x \sqrt{|t|} f(t) dt \quad (25)$$

and

$$\lim_{x \rightarrow \infty} \sqrt{|x|} \int_0^x \sqrt{|t|} g(t) dt \sqrt{|x|} \int_0^x \sqrt{|t|} h(t) dt \varphi(x) = 0$$

imply that (25) defines an isometry \mathcal{S} on $L^2(\mathbf{R}, \varphi(x) dx)$.

\mathcal{S} maps even functions onto even functions orthogonal to $\sqrt{|x|}$, and odd functions onto odd functions orthogonal to $\operatorname{sgn}(x)\sqrt{|x|}$. As a consequence, the range of \mathcal{S} is orthogonal to $\operatorname{sgn}(x)\sqrt{|x|}$ and $\sqrt{|x|}$, but not to the constant 1 as required. The new isometry \mathcal{T}_H given by

$$\mathcal{T}_H f = \begin{cases} f, & f \text{ odd} \\ \mathcal{S}f + \frac{1}{1-\cos\gamma} \langle \mathcal{S}f, 1 \rangle (u-1), & f \text{ even,} \end{cases} \quad (26)$$

where $u(x) = \sqrt{|x|}/\|\sqrt{|\cdot|}\| = \sqrt[4]{\pi/2}\sqrt{|x|}$, $\cos\gamma = \langle 1, u \rangle = \sqrt[4]{1/\pi} \Gamma(3/4)$, has range orthogonal to 1.

Since any function f is the sum of an *odd part* $f^{\circ}(x) = (f(x) - f(-x))/2$ and an *even part* $f^e(x) = (f(x) + f(-x))/2$, (26) completely defines \mathcal{T}_H , which can also be written as

$$\mathcal{T}_H f(x) = f(x) - \sqrt{|x|} \int_0^{|x|} \sqrt{s} f^e(s) ds + \frac{1}{1 - \cos \gamma} \langle \mathcal{S} f^e, 1 \rangle (u(x) - 1). \quad (27)$$

The inverse of \mathcal{S} is given by

$$(\mathcal{S}^{-1}h)(x) = h(x) - \frac{\sqrt{|x|}}{\varphi(x)} \int_x^\infty \operatorname{sgnt} \sqrt{|t|} h(t) \varphi(t) dt$$

and hence

$$\begin{aligned} (\mathcal{T}_H^{-1}h)(x) &= h^{\circ}(x) + \mathcal{S}^{-1} \left(h^e - \frac{1}{1 - \cos \gamma} \langle h^e, u \rangle (u - 1) \right) \\ &= h(x) - \frac{1}{1 - \cos \gamma} \langle h^e, u \rangle (u - 1) \\ &\quad - \frac{\sqrt{|x|}}{\varphi(x)} \int_{|x|}^\infty \sqrt{t} \left[h^e(t) - \frac{1}{1 - \cos \gamma} \langle h, u \rangle (u(t) - 1) \right] \varphi(t) dt. \end{aligned} \quad (28)$$

In particular, (28) reduces to h when h is odd.

6 Construction of isometries on $L^2(\mathbb{R}, dF_0)$ for an arbitrary probability F_0 , by means of a change of variables.

Let us assume that we are given an isometry \mathcal{T}_F on $L_2(\mathcal{E}, dF)$ and wish to construct a new isometry \mathcal{T}_{F_0} on $L^2(\mathcal{E}, dF_0)$. The next lemma gives a simple and general procedure for such construction:

Lemma 1 *If \mathcal{T}_F is an isometry on $L_2(\mathcal{E}, dF)$, then the mapping $g_0 \mapsto \mathcal{T}_{F_0} g_0$ defined by*

$$(\mathcal{T}_{F_0} g_0)(x) = (\mathcal{T}_F g_0 \circ F_0^{-1} \circ F)(F^{-1}(F_0(x)))$$

is an isometry on $L^2(\mathcal{E}, dF_0)$.

Proof. Let us assume $X_0 \sim F_0$, and $\mathbf{E}g_0^2(X_0), \mathbf{E}h_0^2(X_0) < \infty$.

The change of variables $X = F^{-1}(F_0(X_0))$ maps X_0 onto $X \sim F$, and hence $\mathbf{E}g_0^2(X_0) < \infty, \mathbf{E}h_0^2(X_0) < \infty$ are equivalent to $\mathbf{E}g^2(X) < \infty, \mathbf{E}h^2(X) < \infty$, respectively, for $g = g_0 \circ F_0^{-1} \circ F$, and $h = h_0 \circ F_0^{-1} \circ F$.

We may then compute

$$\begin{aligned} \int \mathcal{T}_{F_0} g_0 \mathcal{T}_{F_0} h_0 dF_0 &= \mathbf{E} \mathcal{T}_{F_0} g_0(X_0) \mathcal{T}_{F_0} h_0(X_0) \\ &= \mathbf{E}(\mathcal{T}_F g_0 \circ F_0^{-1} \circ F)(F^{-1}(F_0(X_0)))(\mathcal{T}_F h_0 \circ F_0^{-1} \circ F)(F^{-1}(F_0(X_0))) \\ &= \mathbf{E}(\mathcal{T}_F g)(X)(\mathcal{T}_F h)(X). \end{aligned}$$

Since \mathcal{T}_F is an isometry on $L_2(\mathcal{E}, dF)$, the right-hand term equals

$$\mathbf{E}g(X)h(X) = \mathbf{E}g_0(X_0)h_0(X_0) = \int g_0 h_0 dF_0,$$

and this proves the Lemma.

6.1 Example 1. L-isometries.

From Lemma 1 applied to the isometry in §5.1, we get, for each F_0 , a new isometry

$$(\mathcal{T}_{L, F_0} g)(x) = g(x) - \int_{-\infty}^x \frac{g(t)}{1 - F_0(t)} dF_0(t), \quad (29)$$

on $L_2(\mathbf{R}, dF_0)$. Its inverse

$$(\mathcal{T}_{L, F_0}^{-1} h)(x) = h(x) + \frac{1}{1 - F_0(x)} \int_{-\infty}^x h(t) dF_0(t) \quad (30)$$

is obtained by solving (29) for g . In what follows, any isometry in the class defined by (29) will be called an *L-isometry*. Under suitable assumptions, \mathcal{T}_{L, F_0} satisfies (15) and (17), and hence Theorem 2 applies:

Lemma 2 *When $|a|/(1 - F_0)^\alpha$ belongs to $L^2(\mathbf{R}, dF_0)$ for some positive α , then \mathcal{T}_{L, F_0} satisfies (15) and (17) with $\mathcal{J} = \{(-\infty, x] : x \in \mathbf{R}\}$.*

Proof. The function $|\mathcal{T}_{L, F_0}(a1_{(-\infty, y]})|$ is bounded by

$$G = |a| + \int_{-\infty}^y \frac{|a(s)|}{1 - F_0(s)} dF_0(s)$$

uniformly in y . Let us assume with no loss of generality that $\alpha < 1/2$. The inequalities

$$\int_{-\infty}^{\infty} \left(\int_{-\infty}^x \frac{|a(s)|}{1 - F_0(s)} dF_0(s) \right)^2 dF_0(x)$$

$$\begin{aligned}
&\leq \int_{-\infty}^{\infty} \left(\int_{-\infty}^x \frac{a(s)^2}{(1-F_0(s))^{2\alpha}} dF_0(s) \int_{-\infty}^x \frac{dF_0(s)}{(1-F_0(s))^{2-2\alpha}} \right) dF_0(x) \\
&\leq \left\| \frac{|a|}{(1-F_0)^\alpha} \right\|^2 \frac{1}{1-2\alpha} \int_{-\infty}^{\infty} \left(\frac{1}{(1-F_0(x))^{1-2\alpha}} - 1 \right) dF_0(x) \\
&\leq \left\| \frac{|a|}{(1-F_0)^\alpha} \right\|^2 \frac{1}{2\alpha(1-2\alpha)}
\end{aligned}$$

and the assumptions on a imply that G is in $L^2(\mathbf{R}, dF_0)$, so that (15) holds.

Given a positive ε , let us construct a (finite) partition $(x_0 = -\infty, x_1, x_2, \dots, x_\nu = \infty)$ of \mathbf{R} such that for each $i = 1, 2, \dots, \nu$,

$$\int_{x_{i-1}}^{x_i} a^2(s) dF_0(s) \leq \varepsilon^2/8 \quad (31)$$

and also

$$\int_{x_{i-1}}^{x_i} \frac{a^2(s)}{(1-F_0(s))^{2\alpha}} dF_0(s) \leq \frac{\alpha(1-2\alpha)\varepsilon^2}{4},$$

so that

$$\int_{x_{i-1}}^{x_i} \left(\int_{x_{i-1}}^x \frac{|a|}{1-F_0} dF_0 \right)^2 dF_0(x) \leq \frac{\alpha(1-2\alpha)\varepsilon^2}{4} \frac{1}{2\alpha(1-2\alpha)} = \varepsilon^2/8. \quad (32)$$

Notice that a partition satisfying (31) can be constructed with $\nu \leq 1 + 8/\varepsilon^2$ intervals, and another one satisfying (32) requires at most $1 + \frac{4\|a/(1-F_0)^\alpha\|^2}{\alpha(1-2\alpha)\varepsilon^2}$. The partition obtained by joining the points in both ones satisfies our requirements, and can therefore be achieved with $\nu \leq C/\varepsilon^2$, where C may depend on a (and α) but not on ε .

For all $y \in [x_{i-1}, x_i]$, $\mathcal{T}(a1_{(-\infty, y]})$ is bounded from above by

$$u_i = a1_{(-\infty, x_{i-1})} + |a|1_{[x_{i-1}, x_i]} - \int_{-\infty}^{x_{i-1} \wedge \cdot} \frac{a}{1-F_0} dF_0 + \int_{x_{i-1} \wedge \cdot}^{x_i \wedge \cdot} \frac{|a|}{1-F_0} dF_0$$

and from below by

$$l_i = a1_{(-\infty, x_{i-1})} - |a|1_{[x_{i-1}, x_i]} - \int_{-\infty}^{x_{i-1} \wedge \cdot} \frac{a}{1-F_0} dF_0 - \int_{x_{i-1} \wedge \cdot}^{x_i \wedge \cdot} \frac{|a|}{1-F_0} dF_0.$$

These bounds satisfy

$$\|u_i - l_i\|^2 \leq \left\| 2 \left(|a|1_{[x_{i-1}, x_i]} + \int_{x_{i-1} \wedge \cdot}^{x_i \wedge \cdot} \frac{|a|}{1-F_0} dF_0 \right) \right\|^2 \leq \varepsilon^2$$

as a consequence of (31) and (32), hence $N_{[\cdot]}^{(2)} \leq C/\varepsilon^2$, and the condition (17) holds. This ends the proof of Lemma 2.



6.2 Example 2. N-isometries.

When F_0 is a symmetric probability distribution function ($F_0(x) + F_0(-x) = 1$), then $(g \circ F_0^{-1} \circ \Phi)^e = g^e \circ F_0^{-1} \circ \Phi$ and $(g \circ F_0^{-1} \circ \Phi)^o = g^o \circ F_0^{-1} \circ \Phi$, and hence, from the isometry in §5.3, we get the new isometry

$$\begin{aligned} (T_{H,F_0}g)(x) &= g(x) - \sqrt{|\Phi^{-1}(F_0(x))|} \int_0^{|x|} \frac{\sqrt{|\Phi^{-1}(F_0(s))|} g^e(s)}{\varphi(\Phi^{-1}(F_0(s)))} dF_0(s) \\ &+ \frac{1}{1 - \cos \gamma} \int \mathcal{S}(g^e \circ F_0^{-1} \circ \Phi)(z) \varphi(z) dz \left(u(\Phi^{-1}(F_0(x))) - 1 \right) \end{aligned} \quad (33)$$

on $L_2(\mathbf{R}, dF_0)$. A general version of (33) for nonsymmetric F_0 is equally easy to obtain, but its expression is even more complicated. The mappings given by (33) are called *N-isometries* in the following.

Lemma 3 *When $|a|^e / (\varphi(\Phi^{-1}(F_0(\cdot))))^\alpha$ belongs to $L^2(\mathbf{R}, dF_0)$ for some positive α , then T_{H,F_0} satisfies (15) and (17) with $\mathcal{J} = \{(-\infty, x] : x \in \mathbf{R}\}$.*

Proof. Proceed as in the proof of Lemma 2: Now $|T_{H,F_0}(a1_{(-\infty,y]})|$ is uniformly bounded by

$$\begin{aligned} G &= |a| + \sqrt{|\Phi^{-1}(F_0(x))|} \int_0^{|x|} \frac{\sqrt{|\Phi^{-1}(F_0(s))|} |a|^e(s)}{\varphi(\Phi^{-1}(F_0(s)))} dF_0(s) \\ &+ \frac{1}{1 - \cos \gamma} \|a\| \left(1 \vee u(\Phi^{-1}(F_0(s))) \right). \end{aligned}$$

The first and last term in G are square integrable, because $|a(\cdot)|$ and $u(\Phi^{-1}(F_0(\cdot)))$ are in $L^2(\mathbf{R}, dF_0)$. In order to verify that the middle term has the same property, we assume again $\alpha < 1/2$ and derive the inequalities

$$\begin{aligned} &\int_{-\infty}^{\infty} |\Phi^{-1}(F_0(x))| \left(\int_0^x \frac{\sqrt{|\Phi^{-1}(F_0(s))|} |a|^e(s)}{\varphi(\Phi^{-1}(F_0(s)))} dF_0(s) \right)^2 dF_0(x) \leq \\ &\int_{-\infty}^{\infty} |\Phi^{-1}(F_0(x))| \left| \int_0^x \frac{|\Phi^{-1}(F_0(s))| dF_0(s)}{\varphi^{2-2\alpha}(\Phi^{-1}(F_0(s)))} \int_0^x \frac{(|a|^e(s))^2 dF_0(s)}{\varphi^{2\alpha}(\Phi^{-1}(F_0(s)))} \right| dF_0(x) \\ &\leq M_a \int_{-\infty}^{\infty} |z| \left| \int_0^z \frac{t\varphi(t) dt}{\varphi^{2-2\alpha}(t)} \right| \varphi(z) dz \\ &= \frac{M_a}{1 - 2\alpha} \int_{-\infty}^{\infty} |z| \varphi(z) \left(\frac{1}{\varphi^{1-2\alpha}(z)} - \frac{1}{\varphi^{1-2\alpha}(0)} \right) dz < \infty, \end{aligned}$$

thus providing (15).

The estimation of the $L_2(F_0)$ -metric entropy with bracketing is made as in the proof of the previous Lemma, and it leads to conclude that condition (17) also holds.

7 Constructing isometries on $L_2(\mathcal{E}, dF_0)$, \mathcal{E} space of dimension greater than one.

When \mathcal{E} is the cartesian product $\mathcal{E} = \mathbf{R} \times \mathcal{E}_1$ the measure F_0 is written in terms of the *marginal measure* $\tilde{F}_0(J) = F_0(J \times \mathcal{E}_1)$ and the *conditional measures* $dF_0^c(y, \cdot)$ defined on \mathcal{E}_1 by $F_0(J \times B) = \int_{y \in J} d\tilde{F}_0(y) \int_{z \in B} dF_0^c(y, z)$.

In other words, if $X = (Y, Z) \sim F_0$, then $Y \sim \tilde{F}_0$ and the conditional distribution of Z given $Y = y$ is $\mathbf{P}\{Z \in B \mid Y = y\} = \int_B dF_0^c(y, \cdot)$.

For each g in $L_2(\mathcal{E}, dF_0)$, we denote

$$\tilde{g}(y) = \mathbf{E}(g(Y, Z) \mid Y = y). \quad (34)$$

This new function is in $L_2(\mathbf{R}, d\tilde{F}_0)$ and the mapping $g \mapsto \tilde{g}$ preserves the norm.

Lemma 4 *If $g \in L_2(\mathcal{E}, dF_0)$ and \tilde{T} is an isometry on $(\mathbf{R}, d\tilde{F}_0)$, then*

$$g \mapsto \mathcal{T}g(y, z) = g(y, z) - \tilde{g}(y) + \tilde{T}\tilde{g}(y) \quad (35)$$

is an isometry on $L_2(\mathcal{E}, dF_0)$.

Proof. Given $g, h \in L_2(\mathcal{E}, dF_0)$, let us compute $\langle \mathcal{T}g, \mathcal{T}h \rangle = \mathbf{E}(g(Y, Z) - \tilde{g}(Y) + \tilde{T}\tilde{g}(Y))(h(Y, Z) - \tilde{h}(Y) + \tilde{T}\tilde{h}(Y))$. The equalities $\mathbf{E}g(Y, Z)\tilde{h}(Y) = \mathbf{E}\tilde{h}(Y)\mathbf{E}(g(Y, Z) \mid Y) = \mathbf{E}\tilde{g}(Y)\tilde{h}(Y)$, $\mathbf{E}(g(Y, Z) - \tilde{g}(Y))\tilde{T}\tilde{h}(Y) = 0$, and similar ones obtained by interchanging g and h lead us to write $\langle \mathcal{T}g, \mathcal{T}h \rangle = \mathbf{E}g(Y, Z)h(Y, Z) - \mathbf{E}\tilde{g}(Y)\tilde{h}(Y) + \mathbf{E}\tilde{T}\tilde{g}(Y)\tilde{T}\tilde{h}(Y) = \langle g, h \rangle - \langle \tilde{g}, \tilde{h} \rangle + \langle \tilde{T}\tilde{g}, \tilde{T}\tilde{h} \rangle = \langle g, h \rangle$, and this ends the proof of the Lemma.

8 One application: Consistent goodness-of-fit to the standard normal distribution.

As an illustration of the general procedure for the design of consistent and efficient tests contained in the preceding sections, we describe the tests associated to the isometries in §5.1 and §5.3 for goodness-of-fit to standard normal distribution in \mathbf{R} and \mathbf{R}^2 . Two cases are considered: In Case 1, the tests are designed to have optimum sensibility against shifts of the mean, while in Case 2, the alternative to be detected is a change of dispersion.

8.1 Case 1. Tests designed for detection of shifts in the mean.

8.1.1 The one-dimensional test

Let $\mathcal{E} = \mathbf{R}$, $F_0(x) = \Phi(x) = \int_{-\infty}^x \varphi(t)dt$, $\varphi(t) = \frac{1}{\sqrt{2\pi}}e^{-t^2/2}$, and $F^\tau(x) = \Phi(x - \tau)$. The ratio of the densities is $\frac{f^{(\tau)}(x)}{f_0(x)} = 1 + x\tau + o(\tau)$, so that $k(x) = x$, $\|k\|^2 = \int_{-\infty}^{\infty} x^2\varphi(x)dx = 1$.

Example 8.1.1.1 *TEP associated to the L-isometry.*

The score function is $\hat{a}(x) = \mathcal{T}_{L,\Phi}^{-1}(x) = x + \frac{1}{1-\Phi(x)} \int_{-\infty}^x t\varphi(t)dt = x - \frac{\varphi(x)}{1-\Phi(x)}$ (see (30)), and hence the TEP for the sample $\{X\}$ of size one is

$$\begin{aligned} w_X(x) &= w_X((-\infty, x]) = \hat{a}(X)1_{\{X \leq x\}} - \int_{-\infty}^X \frac{a(t)1_{\{t \leq x\}}\varphi(t)dt}{1-\Phi(t)} \\ &= \left(X - \frac{\varphi(X)}{1-\Phi(X)} \right) 1_{\{X \leq x\}} + \int_{-\infty}^{X \wedge x} d\left(\frac{\varphi(t)}{1-\Phi(t)} \right) \\ &= X1_{\{X \leq x\}} + \frac{\varphi(x)}{1-\Phi(x)} 1_{\{x < X\}}. \end{aligned}$$

In order to describe the general shape of the TEPs and their response to changes in the position or in the dispersion of the samples, we introduce the fictitious sample of size 9: $(\Phi^{-1}(i/10))_{i=1,2,\dots,9}$, that will be referred as *the special sample* in the following. Then we compute the TEPs associated to the optimum score function \hat{a} for the special sample, and for the *shifted* and *dispersed special samples* $(\Phi^{-1}(i/10) + 1)_{i=1,2,\dots,9}$ and $(2\Phi^{-1}(i/10))_{i=1,2,\dots,9}$. The corresponding three diagrams are presented in the left-hand side of Figure 1.

Example 8.1.1.2 *TEP associated to the N-isometry.*

Since k is odd, then $\hat{a}(x) = \mathcal{T}_{N,\Phi}^{-1}k(x) = k(x) = x$, and

$$\begin{aligned} w_X(x) &= w_X((-\infty, x]) = \mathcal{T}_{N,\Phi}(\cdot 1_{(-\infty, x]}(\cdot))(X) \\ &= X1_{\{X \leq x\}} - \sqrt{|X|} \int_0^{|X|} \sqrt{s} \left(\frac{-s}{2} \right) 1_{\{|x| \leq s\}} ds \\ &\quad + \frac{u(X) - 1}{1 - \cos \gamma} \int \left(\frac{-s}{2} 1_{\{|x| \leq |s|\}} - \sqrt{|s|} \int_0^{|s|} \sqrt{t} \left(\frac{-t}{2} \right) 1_{\{|x| < t\}} dt \right) \varphi(s) ds \end{aligned}$$

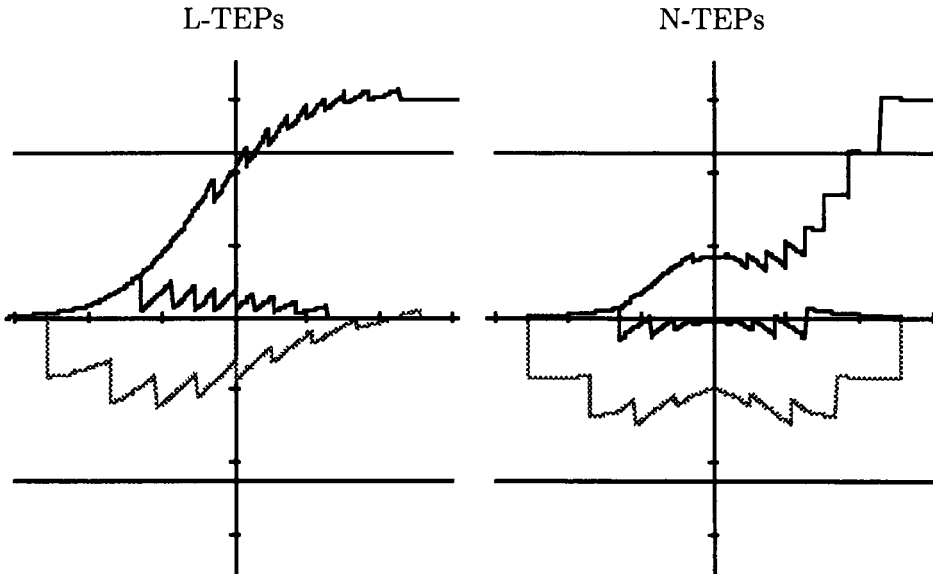


Figure 1: Responses to normal (—), shifted (- -) and dispersed (···) samples, of the TEPs associated to L-isometry (L-TEPs) and to N-isometry (N-TEPs), optimized to detect changes in position.

$$\begin{aligned}
 &= X1_{\{X \leq x\}} + \frac{\sqrt{|X|}}{5} \left(|X|^{5/2} - |x|^{5/2} \right)^+ \\
 &+ \frac{u(X) - 1}{1 - \cos \gamma} \left[-\varphi(|x|) + \frac{2}{5} \int_{|x|}^{\infty} \varphi(s) \sqrt{s} (s^{5/2} - |x|^{5/2}) ds \right].
 \end{aligned}$$

The right-hand side of Figure 1 shows the shapes of the new TEPs.

8.1.2 The test for $d = 2$.

Let $F^{(\tau)}$ be the normal distribution with mean $\begin{pmatrix} \tau \\ 0 \end{pmatrix}$ and variance matrix equal to the identity, and let $F_0 = F^{(0)}$. The density under \mathcal{H}_0 is $f_0(y, z) = \varphi(y)\varphi(z)$, and under the alternative, $f^{(\tau)}(y, z) = \varphi(y - \tau)\varphi(z)$. The densities ratio is $f^{(\tau)}(y, z)/f_0(y, z) = \varphi(y - \tau)/\varphi(y) = 1 + x\tau + o(\tau)$, and hence the drift is $k(y, z) = y$.

We choose now the direction of the shift, namely, the y axis, to project the plane measure on it, and obtain the corresponding marginal distribution

$\tilde{F}_0 = \Phi$. Lemma 4 is now applied to construct two isometries on $L^2(\mathbf{R}^2, dF_0)$, by substituting $\mathcal{T}_{L, \tilde{F}_0}$ and $\mathcal{T}_{H, \tilde{F}_0}$ for $\tilde{\mathcal{T}}$ in (35):

$$\mathcal{T}_{L, F_0} g(y, z) = g(y, z) - \mathbf{E}g(y, Z) + \mathcal{T}_{L, \Phi} \mathbf{E}g(\cdot, Z)(y),$$

$$\mathcal{T}_{H, F_0} g(y, z) = g(y, z) - \mathbf{E}g(y, Z) + \mathcal{T}_{H, \Phi} \mathbf{E}g(\cdot, Z)(y),$$

Z standard normal.

When $g(y, z) = \tilde{g}(y)$ does not depend on z , $\mathcal{T}g = \tilde{\mathcal{T}}\tilde{g}$. In particular, since the drift k depends only on y this implies that the score functions corresponding to \mathcal{T}_{L, F_0} and \mathcal{T}_{H, F_0} also depend only on y and each one is given by the formula obtained for $d = 1$, respectively

$$\hat{a}_{L, F_0}(y, z) = x - \frac{\varphi(y)}{1 - \Phi(y)} \quad \text{and} \quad \hat{a}_{H, F_0}(y, z) = y.$$

Consequently, the TEP for a single observation is, in the former case:

$$\begin{aligned} w_{Y, Z}^L(A) &= w_{Y, Z}^{(\mathcal{T}_{L, \hat{a}})}(A) = \hat{a}_{L, F_0}(Y, Z) 1_{\{(Y, Z) \in A\}} - \iint_A \hat{a}_{L, F_0}(y, z) \varphi(y) \varphi(z) dy dz \\ &\quad + \iint_A \frac{\hat{a}_{L, F_0}(y, z)}{1 - \Phi(y)} (1_{\{Y \leq y\}} - \Phi(y)) \varphi(y) \varphi(z) dy dz, \end{aligned}$$

and, for $A = (-\infty, y] \times (-\infty, z]$, it reduces to

$$w_{Y, Z}^L(y, z) = \left(Y - \frac{\varphi(Y)}{1 - \Phi(Y)} \right) 1_{\{Y \leq y, Z \leq z\}} + \Phi(z) \frac{\varphi(Y \wedge y)}{1 - \Phi(Y \wedge y)}.$$

As for the latter case:

$$\begin{aligned} w_{Y, Z}^H(y, z) &= w_{Y, Z}^{(\mathcal{T}_{H, \hat{a}})}((-\infty, y] \times (-\infty, z]) \\ &= Y 1_{\{Y \leq y, Z \leq z\}} + \frac{\sqrt{|Y|} \Phi(z)}{5} (|Y|^{5/2} - |y|^{5/2})^+ \\ &\quad - \frac{u(Y) - 1}{1 - \cos \gamma} \Phi(z) \left[\varphi(y) - 2 \int_{|y|}^{\infty} \varphi(s) \sqrt{s} \frac{s^{5/2} - |y|^{5/2}}{5} ds \right]. \end{aligned}$$

The goodness-of-fit is tested using the critical region

$$\max_{y, z \in \mathbf{R}} |w_n(y, z)| > c(\alpha)$$

with $w_n(y, z) = \frac{1}{\sqrt{n}} \sum_{i=1}^n w_{Y_i, Z_i}^L(y, z)$ or $w_n(y, z) = \frac{1}{\sqrt{n}} \sum_{i=1}^n w_{Y_i, Z_i}^H(y, z)$, respectively.

A conservative region of size α is obtained with $c(\alpha) = -\Phi^{-1}(\alpha/8)$, because of the well known estimate

$$\mathbf{P}\left\{\max_{y,z \in \mathbf{R}} w^V((-\infty, y], (-\infty, z]) \geq c\right\} \leq 4\Phi(-c)$$

that holds for a V -Wiener process w associated to any probability measure V .

The general shape of the TEPs, and how they are affected by changes in the samples, is sketched in Figure 2, for the TEPs associated to the L-isometry.

The upper part of Figure 2 shows the graph $(y, z, w(y, z))$, $-3 < y, z < 3$ of the TEP for the *special sample* $\{(\Phi^{-1}(i/6), \Phi^{-1}(j/6))\}_{i,j=1,\dots,5}$, corresponding to the L-isometry. As in the one-dimensional pictures, we have introduced an arbitrary *special (two-dimensional) sample* with an empirical distribution abnormally close to its theoretical distribution, assumed to be the standard Gaussian.

The other two graphs in the same figure show the TEPs for the *shifted sample* $\{(\Phi^{-1}(i/6) + 1, \Phi^{-1}(j/6) + 1)\}_{i,j=1,\dots,5}$ and the *dispersed sample* $\{(2\Phi^{-1}(i/6), 2\Phi^{-1}(j/6))\}_{i,j=1,\dots,5}$, in the same domain.

The small diagrams in the left-hand side, show the same graphs, with the direction of vision changed to horizontal, in order to show the position of maxima and minima. The critical planes at $c(5\%)$ and $-c(5\%)$ are also shown.

A graph with points over $c(5\%)$ or under $-c(5\%)$ leads to reject the null hypothesis of goodness of fit, at a level smaller than 5%.

8.2 Design of a test specially sensitive to changes in dispersion.

We choose now $F^\tau(x) = \Phi((1 - \tau)x)$, and so

$$\frac{f^\tau(x)}{f_0(x)} = (1 - \tau) \frac{\varphi(x - \tau x)}{\varphi(x)} = (1 - \tau)(1 + \tau x^2 + o(\tau)),$$

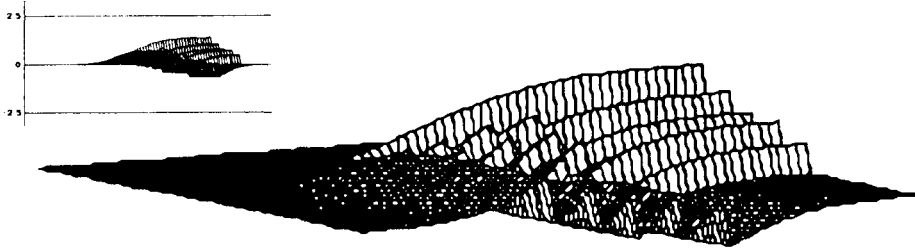
and $k(x) = x^2 - 1$, $\|k\|^2 = 2$.

Example 8.2.0.1 *TEP associated to the L-isometry.*

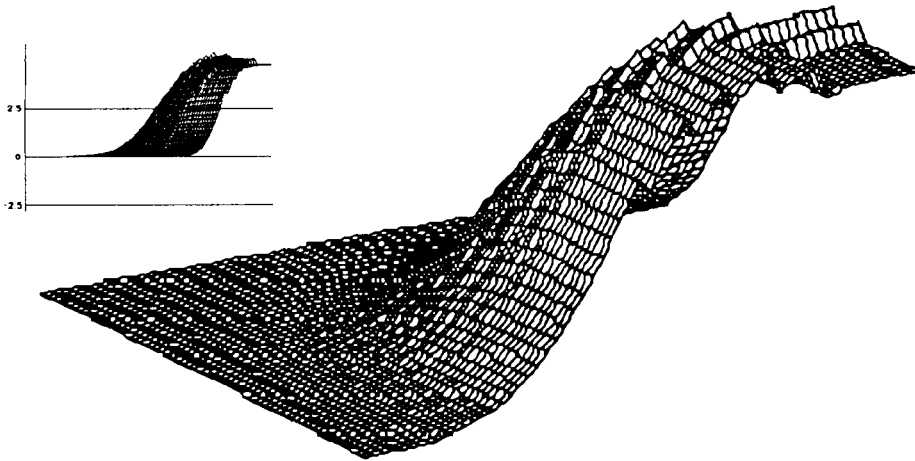
The score function is

$$\hat{a}(x) = \frac{1}{\sqrt{2}} \left[x^2 - 1 + \frac{1}{1 - \Phi(x)} \int_{-\infty}^x (t^2 - 1)\varphi(t)dt \right]$$

Original *special sample*:



Shifted sample:



Dispersed sample:

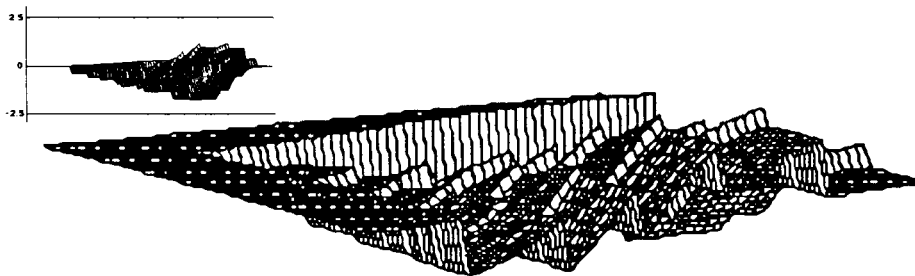


Figure 2: Graph of the TEPs for the two-dimensional *special sample*, corresponding to the L-Isometry, and for the same sample after changes in position or dispersion.

$$= \frac{1}{\sqrt{2}} \left[x^2 - 1 - \frac{x\varphi(x)}{1 - \Phi(x)} \right],$$

(see again (30)) and the TEP for the sample $\{X\}$ of size one is, consequently,

$$\begin{aligned} & \frac{1}{\sqrt{2}} \left[\left(X^2 - 1 - \frac{X\varphi(X)}{1 - \Phi(X)} \right) 1_{\{X \leq x\}} - \int_{-\infty}^{X \wedge x} \frac{t^2 - 1 - \frac{t\varphi(t)}{1 - \Phi(t)}}{1 - \Phi(t)} \varphi(t) dt \right] \\ &= \frac{1}{\sqrt{2}} \left[(X^2 - 1) 1_{\{X \leq x\}} + \frac{x\varphi(x)}{1 - \Phi(x)} 1_{\{x < X\}} \right]. \end{aligned}$$

The shape of the TEPs for the same special samples used in previous one-dimensional diagrams is shown in Figure 3.

Example 8.2.0.2 *TEP associated to the N -isometry.*

From (28) and $k(x) = (x^2 - 1)/\sqrt{2}$, we evaluate the score function \hat{a} numerically, and that evaluation is then used to compute the TEP $w_X(x) = \mathcal{T}_H(\hat{a}1_{\{\cdot \leq x\}})(X)$ by numerical integration in (27).

The shape of the TEPs for the special samples is shown in the right-hand part of Figure 3.

The extension to $d = 2$ is similar to the one in §8.1.2.

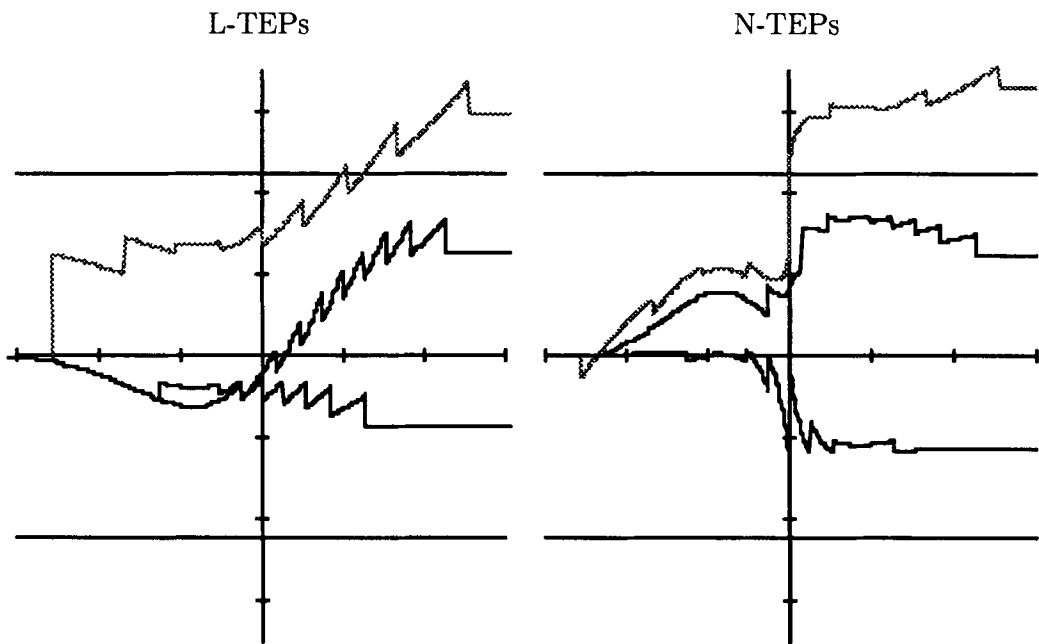


Figure 3: Responses to normal (—), shifted (—) and dispersed (.....) samples, of the TEPs associated to L-isometry (L-TEPs) and to N-isometry (N-TEPs), optimized to detect changes in dispersion.

References

- [1] Cabaña, A., *Transformations du pont brownien empirique et tests de type Kolmogorov-Smirnov*. C. R. Acad. Sci. Paris t. **317** Série I (1993),pp.315-318.
- [2] Cabaña, A., *Transformations of the empirical process and Kolmogorov-Smirnov tests*. Ann. Statist., to appear.
- [3] Cabaña, A. and Cabaña, E.M., *Goones-of-fit and comparison tests of the Kolmogorov-Smirnov type for bivariate populations*, Ann. Statist., Vol 22, No. 3, 1447-1459, 1994.
- [4] Cabaña, A. and Cabaña, E.M., *Bridge-to-bridge transformiosn and Kolmogorov-Smirnov tests*, Commun. Statist. - Theory Meth., **25** 227-234, 1996.
- [5] Donsker, M. D. *Justification and extension of Doob's heuristic approach to the Kolmogorov-Smirnov theorems*, Ann. Math. Statist. 23 277, 1952.
- [6] Efron, B. and Johnston, I. *Fisher's information in terms of hazard rate*. Ann. Statist. 18, 38-62, 1990.
- [7] Groeneboom, P. and Wellner, J.A., *Information Bounds and Non-parametric Maximum Likelihood Estimation.*, DMV Seminars, Band 19, Birkhauser, New York, 1992.
- [8] Hájek, J. and Sidák, Z., *Theory of Rank Tests*, Academic Press, New York, 1967.
- [9] Karatzas, I. and Shreve, S. *Brownian motion and stochastic calculus*. Springer Verlag, New York, 1988.
- [10] Khmaladze, E. V., *Martingale approach in the theory of Goodness-of-fit*. Theory of Probability and its applications, Vol XXVI, No. 2 (1981), 240-257.
- [11] Le Cam, L. and Yang, G.L. *Asymptotics in Statistics. Some basic concepts*. Springer-Verlag, New York, 1990.
- [12] Morán, M. and Urbina, W. *Some properties of the Gaussian Hilbert transform*. submitted.

- [13] Oosterhoff, J; van Zwet, W. R. *A note on contiguity and Hellinger distance*. Contributions to Statistics, 157-166, Reidel, Dordrecht, 1979.
- [14] Ossiander, M., *A central limit theorem under metric entropy with L_2 bracketing*. Ann. Probab. 15, 897-919, 1987.
- [15] Pfanzagl, J. *Contributions to a General Asymptotic Statistical Theory*. Lecture Notes in Statistics 13, Springer-Verlag, New York, 1982.
- [16] Ritov, J. and Wellner, J.A. *Censoring, martingales and the Cox model*. Contemp. Math. 80, 191-220, 1988.
- [17] Sansone, G. *Orthogonal functions*. Pure and Applied Mathematics Vol. IX, edited by R. Courant, L. Bers and J. J. Stoker, Interscience, N.Y., 1959.

Relació dels últims Preprints publicats:

- 191 *Tetravalent modal algebras and logics*. Josep Maria Font and Miquel Rius. AMS 1991 Subject Classification: 03G25, 03B50, 03B45, 06D30. December 1995.
- 192 *Stochastic differential equations with random coefficients*. Arturo Kohatsu-Higa, Jorge León and David Nualart. AMS Subject Classification: 60H10. December 1995.
- 193 *Variability reduction with multiweighing proceedings*. Miquel Salicrú, Carlos González and Juan J. Barreiro. AMS Subject Classification: 62N10. December 1995.
- 194 *Some compactness results in real interpolation for families of Banach spaces*. María J. Carro and Jaak Peetre. 1980 Mathematics Subject Classification (1985 Revision): 46M37, 47B07. January 1996.
- 195 *Separatrix splitting for a high-frequency perturbation of the pendulum*. V. G. Gelfreich. 1991 Mathematics Subject Classification: 34C37, 34C29, 58F30, 70H05. January 1996.
- 196 *A second order Stratonovich differential equation with boundary conditions*. Aureli Alabert and David Nualart. AMS Subject Classification: 60H10, 60H07. January 1996.
- 197 *A Gentzen system equivalent to the BCK-logic*. Romà J. Adillon and Ventura Verdú. Mathematics Subject Classification: 03B20. January 1996.
- 198 *On Gentzen systems associated with the finite linear MV-algebras*. Àngel J. Gil, Antoni Torrens Torrell and Ventura Verdú. Mathematics Subject Classification: 03B22, 03F03, 03G20, 03G99. February 1996.
- 199 *Exponentially small splitting of separatrices under fast quasiperiodic forcing*. Amadeu Delshams, Vassili Gelfreich, Àngel Jorba and Tere M. Seara. AMS classification scheme numbers: 34C37, 58F27, 58F36, 11J25. February 1996.
- 200 *Existence and regularity of the density for solutions to stochastic differential equations with boundary conditions*. Arturo Kohatsu-Higa and Marta Sanz-Solé. AMS 1990 SCI: 60H07, 60H10, 60H99. March 1996.
- 201 *A forcing construction of thin-tall Boolean algebras*. Juan Carlos Martínez. 1991 Mathematics Subject Classification: 03E35, 06E99, 54E12. March 1996.
- 202 *Weighted continuous metric scaling*. C. M. Cuadras and J. Fortiana. AMS Subject Classification: 62H25, 62G99. April 1996.
- 203 *Homoclinic orbits in the complex domain*. V.F. Lazutkin and C. Simó. AMS Subject Classification: 58F05. May 1996.
- 204 *Quasivarieties generated by simple MV-algebras*. Joan Gispert and Antoni Torrens. Mathematics Subject Classification: 03B50, 03G99, 06F99, 08C15. May 1996.
- 205 *Regularity of the law for a class of anticipating stochastic differential equations*. Carles Rovira and Marta Sanz-Solé. AMS Subject Classification: 60H07, 60H10. May 1996.
- 206 *Effective computations in Hamiltonian dynamics*. Carles Simó. AMS Subject Classification: 58F05, 70F07. May 1996.
- 207 *Small perturbations in a hyperbolic stochastic partial differential equation*. David Márquez-Carreras and Marta Sanz-Solé. AMS Subject Classification: 60H15, 60H07. May 1996.



