Modified Anderson-Darling test with selective power improvement.

Alejandra Cabaña and Enrique M. Cabaña

ABSTRACT

Transformed empirical processes (TEPs) have been used by the authors in a previous paper to construct consistent and selectively efficient goodness-of-fit tests of the Kolmogorov - Smirnov type.
A straightforward application of the same ideas to the construction of tests of the Cramér - von Mises type with the same properties leads to cumbersome computations.
This short note exhibits some of the inconvenients encountered, and introduces a new family of quadratic statistics of the Cramér - von Mises type, in order to circumvent the difficulties.
AMS Classification numbers: 62G10, 62G20, 62G30, 60E20.

RESUMEN

Los procesos empricos transformados (TEPs) han sido utilizados por los autores para la construccin de pruebas de ajuste coherentes y selectivamente eficientes del tipo de Kolmogorov-Smirnov.
Para aplicar las mismas ideas, de manera directa, a la construccin de pruebas de ajuste del tipo de Cramér - von Mises con las mismas propiedades de coherencia y potencia selectiva, se requiere realizar clculos muy complicados.
En esta breve nota se describen algunas de las dificultades que se encuentran al intentar tal generalizacin, y se introduce una nueva familia de estadsticos del tipo de Cramér - von Mises que permite evitar esas dificultades.
Números de clasificación AMS: 62G10, 62G20, 62G30, 60E20.

1 Introduction

The design of tests suited to detect a specific kind of alternative is a common procedure in nonparametric statistics. Linear rank tests (see [7]) are typical examples. In applying them, the statistician chooses the scores in such a way that the power of the test is optimized for a specific family of alternatives. As for other fixed alternatives, the resulting tests may be inefficient or even unable to detect them.

Other well known nonparametric tests have the property of rejecting any fixed alternative for sufficiently large sample sizes. Kolmogorov - Smirnov and Cramér - von Mises tests, for instance, have this consistency property.

In a previous paper (see [3]), the authors have proposed a way of obtaining goodness-of-fit tests, having these two just mentioned desirable properties. They are both consistent against every fixed alternative and specially efficient against a given sequence of contiguous ones. The critical regions are of the Kolmogorov-Smirnov type, and *Transformed Empirical Processes* (TEPs) play there the role of the empirical process in the classical K-S tests. The families of tests in [3] depend on a functional parameter, that has to be adequately chosen in order to achieve the optimal efficiency against the given sequence of alternatives, maintaining the consistency. The resulting tests can be applied when the statistician requires consistency, and, in addition, is specially interested in avoiding an *error of type II* when the alternatives are of some specific kind.

Tests of the Cramér - von Mises type, based on quadratic functionals of the empirical process, can be modified as well, by replacing the empirical process by a TEP, to produce consistent and selectively efficient tests. After describing the TEPs and their asymptotic distributions (§2), we show in §3 and §4 that, when the modified Cramér - von Mises statistics are defined in the apparently simplest way, the optimum score functions and the asymptotic distributions of the test statistics may be quite difficult to obtain.

In order to overcome these difficulties, we introduce in §5 a particular family of statistics for which the optimum score functions are easily obtained. In fact, they are the same that optimize the behaviour of the modified Kolmogorov - Smirnov tests for the same TEPs, and under the same family of contiguous alternatives.

In addition, these test statistics are asymptotically distribution free both under the null hypothesis and under the alternatives. The asymptotic distribution of the test statistic under the null hypothesis, and under the privileged alternatives, depends only on the weight function, and the *size* but not the *shape* of the alternatives. Furthermore, we show elsewhere ([4]) that the shape of the weight function has a little effect on the resulting power.

As a consequence, tables of critical levels and asymptotic powers of the tests can be constructed by simulation. Such tables are provided in the last section (§6).

Some theoretical comments on the distribution of the test statistic, particularly for the case of a constant weight function, are also contained in [4]. A particular application to the derivation of normality tests is developed in [5];

though the general form of the test statistics may look rather complicated, for this latter case, the test statistic is simply a quadratic form evaluated on a vector of sums of polynomials in the sample points.

2 The Transformed Empirical Processes and their asymptotic distributions.

Let $\{X_1, X_2, \ldots, X_n\}$ denote a sample of independent real random variables with distribution function F, and let us consider a sequence of probability distributions $F^{(n)}$ *contiguous* to a given probability distribution F_0 (see [8],[9]), and such that

i. $F^{(n)}$ has density f_n with respect to F_0,

ii. the functions k_n defined by $\sqrt{f_n} = 1 + \frac{\delta k_n}{2\sqrt{n}}$ are absolutely and uniformly bounded by K such that $\int K^2 dF_0 << \infty$,

iii. there exists k such that $\int k^2 dF_0 = 1$, and $\lim_{n\to\infty} \int (k_n - k)^2 dF_0 = 0$.

Families of Transformed Empirical Processes depending on a functional parameter (*score function*) have been introduced by one of the authors in [2], with the purpose of designing goodness-of-fit tests of the Kolmogorov-Smirnov type for the null hypothesis $\mathcal{H}_0$: "$F = F_0$". Those tests share the following two properties: (a) they are consistent against any fixed alternative "$F \neq F_0$" and (b) they are specially sensitive against the particular sequence of alternatives $\mathcal{H}_n$: "$F = F^{(n)}$".

In this article we focus our attention in introducing new tests that also share properties (a) and (b), based on a quadratic statistic of the Anderson-Darling type.

Following [3], we define the Transformed Empirical Process (TEP) of the sample $\{X_1, X_2, \ldots, X_n\}$, associated to the distribution function F_0, the isometry $\mathcal{T}$ on $L_2 = L_2(\mathbf{R}, dF_0)$ with range equal to the orthogonal complement of the constant function 1, and the *score function* a with $\|a\|^2 = \int a^2(x)dF_0(x) = 1$ as

$$w_n^{(a,\mathcal{T})}(x) = \int \mathcal{T}(a1_x)db_n, \quad (1)$$

where 1_x is the indicator function of $(-\infty, x]$, $b_n(x) = \sqrt{n}(F_n(x) - F_0(x))$ is the empirical process and $F_n(x) = \frac{1}{n}\sum_{i=1}^n 1_{\{X_i \leq x\}}$ is the sample distribution function.

Let $\mathcal{J} = \{(-\infty, x] : x \in \mathbf{R}\}$ and let $w^{(V)}$ denote a Wiener process on $\mathbf{R}$ with covariance function

$$V(x) = \mathbf{E}\left(w^{(V)}(x)\right)^2 = \int_{-\infty}^{x} a^2(t)dF_0(t), \ \mathbf{E}w^{(V)}(x)w^{(V)}(y) = V(x \wedge y). \quad (2)$$

It is shown in [3] that, under suitable conditions, the TEP $w_n^{(a,\mathcal{T})}(x), x \in \mathbf{R}$ converges in law to

$$w^{(V)}(x) + \delta \int k\mathcal{T}(a1_x)dF_0, x \in \mathbf{R} \tag{3}$$

in the space of right continuous functions with left limits, as n tends to infinity. In particular, $w_n^{(a,\mathcal{T})}$ converges in law to $w^{(V)}$ under $\mathcal{H}_0$.

We refer to [3] for general conditions ensuring the convergence to (3).

3 The modified Cramér - von Mises statistics.

Classical tests of Cramér - von Mises type are based on the quadratic statistics

$$S_n = \int (b_n(x))^2 \psi(F_0(x))dF_0(x). \tag{4}$$

In this expression, ψ is a *weight function* to be adequately chosen. The *Cramér - von Mises test* corresponds to the selection $\psi(x) = 1$, and other selections of the weight function lead to other tests with similar properties. In particular, the *Anderson-Darling test* is based on the quadratic statistic $\text{AD}_n = \int (b_n(x))^2 \frac{dF_0(x)}{F_0(x)(1-F_0(x))}$ with $\psi(F_0(x)) = (\mathbf{E}(b_n(x))^2)^{-1}$.

It is well known that b_n is asymptotically distributed as a Brownian bridge $b^{(F_0)}$ associated to the probability F_0 (that is, $b^{(F_0)}$ is a centered Gaussian process with covariances $\mathbf{E}b^{(F_0)}(x)b^{(F_0)}(y) = F_0(x \wedge y) - F_0(x)F_0(y)$) and therefore S_n is asymptotically distributed as $\int (b(s))^2 \psi(s)ds$ where b denotes a standard Brownian bridge.

When the empirical process b_n is replaced by the TEP $w_n^{(a,\mathcal{T})}$, the analogue of (4) is

$$S_n^{(a,\mathcal{T})} = \int (w_n^{(a,\mathcal{T})}(x))^2 \psi(V(x))dV(x). \tag{5}$$

The variance $V(x) = \mathbf{Var} w_n^{(a,\mathcal{T})}(x)$ plays in (5) the same role as F_0 in (4) with the same purpose of simplifying the description of the asymptotic distributions. Now it is easily verified that $S_n^{(a,\mathcal{T})}$ is asymptotically distributed as $\int w^2(s)\psi(s)ds$, where w is a standard Wiener process, under $\mathcal{H}_0$.

When $\mathcal{H}_n$ applies instead, the asymptotic distribution of $S_n^{(a,\mathcal{T})}$ is that of

$$\int \left(w^{(V)}(x) + \delta \int k\mathcal{T}(a1_{(-\infty,x]})dF_0 \right)^2 \psi(V(x))dV(x)$$

$$= \int \left(w(V(x)) + \delta \int_{-\infty}^{x} a^{-1}(\mathcal{T}^{-1}k)dV \right)^2 \psi(V(x))dV(x)$$

$$= \int \left(w(s) + \delta \int_0^s h(r)dr \right)^2 \psi(s)ds,$$

with

$$h(V(x)) = (\mathcal{T}^{-1}k)(x)/a(x). \tag{6}$$

The function h must satisfy

$$\int_0^1 h^2(s)ds = \int_{-\infty}^{\infty} h^2(V(x))dV(x) = \int_{-\infty}^{\infty} (\mathcal{T}^{-1}k)^2 dF_0 = \int_{-\infty}^{\infty} k^2 dF_0 = 1.$$

A reasonable heuristic criterion to improve the efficiency of tests with critical regions $S_n^{(a,\mathcal{T})} > c$ (where c is a suitable constant) is to maximize the asymptotic bias

$$B = \mathbf{E}(S_n^{(a,\mathcal{T})}|\mathcal{H}_n) - \mathbf{E}(S_n^{(a,\mathcal{T})}|\mathcal{H}_0) = \delta^2 \int_0^1 \left(\int_0^s h(r)dr \right)^2 \psi(s)ds.$$

4 The optimization problem.

4.1 The general setting.

The criterion sketched in the previous section poses the problem of finding h in $L^2(0,1)$ with $\|h\| = 1$, such that $\int_0^1 \left(\int_0^s h(r)dr\right)^2 \psi(s)ds$ is maximum, for the given nonnegative weight function ψ. The quantity to be maximized can be written as

$$\int_0^1 h(r_1)dr_1 \int_0^1 h(r_2)dr_2 \int_{r_1 \vee r_2}^1 \psi(s)ds = \int_0^1 \int_0^1 K(r_1, r_2)h(r_1)h(r_2)dr_1 dr_2,$$

with kernel $K(r_1, r_2) = \int_{r_1 \vee r_2}^1 \psi(s)ds$. The maximum is the largest eigenvalue of the Fredholm operator

$$f \mapsto \int_0^1 K(\cdot, r)f(r)dr, \tag{7}$$

and it is attained al the corresponding normalized eigenfunction h.

In order to obtain the eigenfunctions h and eigenvalues λ of (7) we must solve

$$\lambda h(x) = \int_0^1 \int_{x \vee r}^1 \psi(s)ds h(r)dr = \int_x^1 \psi(s)ds \int_0^s h(r)dr.$$

The solutions satisfy the differential equation $\lambda h'(x) = -\psi(x) \int_0^x h(r)dr$ with boundary condition $h(1) = 0$. Therefore, the primitive $H(t) = \int_0^t h(s)ds$ a primitive of h satisfies the conditions

$$\lambda H''(t) = -\psi(t)H(t),\ H(0) = 0,\ H'(1) = 0. \tag{8}$$

The equations (8) characterize the function H, together with the condition $\|h\| = 1$ that can be written as

$$1 = \int_0^1 (H'(t))^2 dt = -\int_0^1 H(t)H''(t)dt = \lambda^{-1}\int_0^1 H^2(t)\psi(t)dt. \qquad (9)$$

Once solved (8), (9) in H, $h = H'$ is known and the differential equation

$$V'(x)h^2(V(x)) = (\mathcal{T}^{-1}k(x))^2$$

with the initial condition $V(0) = 0$ has to be solved in order to obtain the score function $a(x) = \sqrt{V'(x)}$. Notice that this is equivalent to the integral equation

$$\int_0^{V(x)} h^2(y)dy = \int_0^x (\mathcal{T}^{-1}k(t))^2 dt. \qquad (10)$$

The sign of a is determined from (6) as $\mathrm{sgn} a = \mathrm{sgn}(h \circ V)\mathrm{sgn}(\mathcal{T}^{-1}k)$.

4.2 The solutions in two particular cases.

The difficulties of the preceding approach are better evaluated by means of some examples: let us first obtain the optimum score function for the case $\psi(x) = 1$.

In this case, (8) implies that $H(t)$ is proportional to $\sin\sqrt{\lambda^{-1}}t$, $\sqrt{\lambda^{-1}} = \nu\pi - \frac{\pi}{2}$. The eigenfunctions are $h(t) = \sqrt{2}\cos(\nu\pi - \frac{\pi}{2})t$, so that the maximum of the eigenvalues $(\nu\pi - \frac{\pi}{2})^{-2}$ is obtained for $\nu = 1$ and the corresponding eigenfunction is $h(t) = \sqrt{2}\cos\frac{\pi}{2}t$

Now (10) reads $V(x) + \frac{1}{\pi}\sin\pi V(x) = \int_0^x (\mathcal{T}^{-1}k(y))^2 dy$ and can be solved in $V(x)$ because the function $z \mapsto z + \frac{1}{\pi}\sin\pi z$ is strictly increasing, but even in this simple case, such procedure does not give us a closed formula for V neither for the score function $a = \sqrt{V'}$.

Our second example is the analogue of the Anderson - Darling test, obtained with $\psi(x) = 1/x$. This particular selection of the weight function imitates the criterion applied for the definition of the classical Anderson - Darling statistic, that is, to choose the weight so as the expectation of the integrand is constant. In the present case, $\psi(V(x)) = \left(\mathbf{E}\left(w_n^{(a,\mathcal{T})}(x)\right)^2\right)^{-1} = 1/V(x)$ is obtained. The corresponding statistic is $\mathrm{AD}_n^{(a,\mathcal{T})} = \int (w_n^{(a,\mathcal{T})}(x))^2 dV(x)/V(x)$.

The differential equation in (8) and the initial condition $H(0) = 0$ lead to the series expansion $h(t) = H'(t) = c\sum_{n=0}^{\infty}(-1)^n(\lambda^{-1}t)^n/(n!^2) = cJ_0(2\sqrt{\lambda^{-1}t})$, where J_0 denotes as usual the Bessel function of the first kind of order 0 and

c is a constant to be determined. The additional condition $H'(1) = 0$ implies that $\zeta = 2\sqrt{\lambda^{-1}}$ must be one of the roots of J_0.

The normalization $\|h\| = 1$ gives $c = \zeta(2\int_0^\zeta zJ_0^2(z)dz)^{-1/2}$. Up to this point, h is determined up to the selection of the root ζ. In order to obtain the maximum λ, we choose ζ equal to the minimum root ζ_1 of J_0.

After integrating in the left hand side of (10), we get

$$\frac{V(x)(J_{-1}(\sqrt{V(x)}\zeta)J_1(\sqrt{V(x)}\zeta) - J_0^2(\sqrt{V(x)}\zeta))}{J_{-1}(\zeta)J_1(\zeta) - J_0^2(\zeta)} = \int_0^x (\mathcal{T}^{-1}k(y))^2 dy,$$

and it remains to replace ζ by ζ_1, solve in V and compute $a = \sqrt{V'}$.

We believe unnecessary to go further in order to show that the calculations involved in using these modified quadratic statistics make them rather unmanageable.

5 A new family of quadratic statistics.

Consider first the family of statistics

$$T_{n,x}^{(a)} = \int \left(\int c_{(x,y)}(z)dw_n^{(a)}(z)\right)^2 \frac{dV(y)}{\int c_{(x,y)}(z)dV(z)} \tag{11}$$

with

$$c_{(x,y)}(s) = \begin{cases} 1 & \text{if } x << s << y, y << x << s \text{ or } s << y << x, \\ 0 & \text{otherwise.} \end{cases} \tag{12}$$

In particular, $T_{n,-\infty}^{(a)}$ equals (5).

Then we integrate $T_{n,x}^{(a)}$ with respect to $dV(x)$ thus defining the new statistic

$$T_n^{(a)} = \iint \left(\int c_{(x,y)}(z)dw_n^{(a)}(z)\right)^2 \frac{dV(x)dV(y)}{\int c_{(x,y)}(z)dV(z)}. \tag{13}$$

Although (13) looks intricate because of the multiple integration, we show below that $T_n^{(a)}$ does not have the disadvantages sketched in regard with the examples in §4.2. On the contrary, the optimum score function and the asymptotic distributions under $\mathcal{H}_0$ and $\mathcal{H}_n$ are extremely simple.

5.1 Asymptotic behaviour of $T_n^{(a)}$ under fixed alternatives. Consistency of the tests.

From $w_n^{(a)} = \frac{1}{\sqrt{n}}\sum_{i=1}^n w_{\{X_i\}}^{(a)}$, with $w_{\{X_i\}}^{(a)}$ equal to the TEP corresponding to the sample $\{X_i\}$ of size one, the expectation and variance of $\int c_{(x,y)}(z)dw_n^{(a)}(z)$

are respectively

$$\mathbf{E}\int c_{(x,y)}(z)dw_n^{(a)}(z) = \sqrt{n}\,\mathbf{E}\int c_{(x,y)}(z)dw_{\{X_1\}}^{(a)}(z) \tag{14}$$

and

$$\mathbf{Var}\int c_{(x,y)}(z)dw_n^{(a)}(z) = \mathbf{Var}\int c_{(x,y)}(z)dw_{\{X_1\}}^{(a)}(z).$$

It is easily shown by applying the same arguments used in [3] that when $F \neq F_0$, then there exist x_0 and y_0 such that

$$\mathbf{E}\int c_{(x,y)}(z)dw_{\{X_1\}}^{(a)}(z) \neq 0 \text{ for } X \sim F \tag{15}$$

for $x = x_0$ and $y = y_0$. Then, the continuity of the left-hand member of (15) as a function of x, y implies that there are neighbourhoods I of x_0 and J of y_0 such that $|\mathbf{E}\int c_{(x,y)}(z)dw_{\{X_1\}}^{(a)}(z)|$ is greater than a certain constant $k > 0$ for $x \in I$ and $y \in J$.

PROPOSITION 1 *When the score function a is $F_0 - a.s.$ different from zero, and $X \sim F \neq F_0$, then* $\mathbf{P}\{\lim T_n^{(a)} = +\infty\} = 1$.

Proof. From the previous context and the assumption on a it follows that $V(I)$ and $V(J)$ do not vanish. Hence (14) tends to infinity as $n \to \infty$ so that the required conclusion is readily obtained. □

As a corollary, the test with critical region $T_n^{(a)} >$ constant is consistent for any $F \neq F_0$.

5.2 Asymptotic behaviour of $T_n^{(a)}$ under the sequence of contiguous alternatives.

The process (3) that has the limit distribution of $\{w_n^{(a)}(A) : A \in \mathcal{J}\}$ under the sequence of alternatives $\mathcal{H}_n$: "$F = F^{(\delta/\sqrt{n})}$", can also be written as $\{w^{(V)}(A) + \delta\int_A a(\mathcal{T}^{-1}k)dF_0$, since $\int k\mathcal{T}(1_A a)dF_0 = \int_A a\mathcal{T}^{-1}kdF_0$ because $\mathcal{T}$ is an isometry.

Therefore $T_n^{(a)}$ is asymptotically distributed under $\mathcal{H}_n$ as

$$\iint \left(\int c_{(x,y)}(z)(dw^{(V)}(z) + \delta a(\mathcal{T}^{-1}k)dF_0)\right)^2 \frac{dV(x)dV(y)}{\int c_{(x,y)}(z)dV(z)}$$

$$= \iint \left(\int c_{(x,y)}(z)(dw(V(z)) + \delta h(V(z))dV(z))\right)^2 \frac{dV(x)dV(y)}{\int c_{(x,y)}(z)dV(z)}, \tag{16}$$

where $h(V(z))a(z) = (\mathcal{T}^{-1}k)(z)$. Since $c_{(x,y)}(z) = c_{(V(x),V(y))}(V(z))$, then, with new variables $r = V(x)$, $s = V(y)$, $t = V(z)$, (16) reduces to

$$\iint \left(\int c_{(r,s)}(t)(dw(t) + \delta h(t)dt) \right)^2 \frac{dr\ ds}{\int c_{(r,s)}(t)dt}$$

$$= \int_0^1 \int_0^1 C(t,u)(dw(t) + \delta h(t)dt)(dw(u) + \delta h(u)du), \tag{17}$$

with

$$C(t,u) = \int_0^1 \int_0^1 c_{(r,s)}(t)c_{(r,s)}(u) \frac{dr\ ds}{\lambda(r,s)}, \quad \lambda(r,s) = \int c_{(r,s)}(t)dt. \tag{18}$$

The distribution of (17) depends only on the selected score function, through the function h. In particular, the asymptotic bias under the alternatives is

$$\Delta(a) = \delta^2 \int_0^1 \int_0^1 C(t,u)h(t)h(u)dt\ du. \tag{19}$$

The limit behaviour described by (17) suggests to reject the null hypothesis $\mathcal{H}_0$ when $T_n^{(a)}$ is greater than an adequate constant, and, in order to improve the sensitivity of the test with respect to the given sequence of contiguous alternatives, we propose to choose the score function a that maximizes the asymptotic bias $\Delta(a)$.

PROPOSITION 2 *The asymptotic bias $\Delta(a)$ given by* (19) *is maximum when the score function a is chosen equal to $\hat{a} = \mathcal{T}^{-1}k$, and its maximum value is $\delta^2/2$.*

Remark. The optimum score $\hat{a} = \mathcal{T}^{-1}k$ is the same that optimizes the power of tests of the Kolmogorov-Smirnov type (see [2, 3]).

Proof. Let us compute, for $t << u$,

$$C(t,u) = \int_0^t dr \int_u^1 \frac{ds}{s-r} + \int_0^t dr \int_0^r \frac{ds}{1+s-r}$$

$$+ \int_t^u dr \int_t^r \frac{ds}{1+s-r} + \int_u^1 dr \int_u^r \frac{ds}{1+s-r} = \gamma(|u-t|), \tag{20}$$

with $\gamma(y) = 1 + |y|\log(|y|) + (1-|y|)\log(1-|y|)$. The expression (20) also holds for $u << t$, since it depends symmetrically on t and u.

The function γ is symmetric with respect to 0 and 1/2, and this implies that $\int_0^1 C(t,u)du$ does not depend on t and equals $\int_0^1 \gamma(y)dy = 1/2$ so that, when h is the constant 1, and hence $a = \hat{a}$, we have $\Delta(\hat{a}) = 1/2$. On the other hand, by Cauchy-Schwartz Inequality,

$$\Delta(a) \leq \int_0^1 \int_0^1 h^2(t) C(t,u) dt\, du = \frac{1}{2} \int_0^1 h^2(t) dt = \frac{1}{2}$$

since the restriction $\|k\| = 1$ implies that h must satisfy

$$\int_0^1 h^2(s) ds = \int (h(V(x)))^2 dV(x) = \int (\mathcal{T}^{-1} k(x))^2 dF_0 = \|\mathcal{T}^{-1} k\|^2 = 1.$$

This ends the proof of our proposition.

6 Performing the test.

6.1 Computing the test statistic.

Let us abbreviate $T_n = T_n^{(\hat{a})}$. The same changes of variables made in §3 lead us to write our optimum test variable as

$$T_n = \int_0^1 \int_0^1 C(t,u) dw_n^{(\hat{a})}(V^{-1}(t)) dw_n^{(\hat{a})}(V^{-1}(u)),$$

$$= \frac{1}{n} \sum_{i,j=1}^n \int_0^1 \int_0^1 C(t,u) dw_{X_i}^{(\hat{a})}(V^{-1}(t)) dw_{X_j}^{(\hat{a})}(V^{-1}(u)).$$

It is easily verified that for any measurable g, $\int g(x) dw_X^{(a)}(x) = \mathcal{T}(ag)(X)$, and therefore, with the notations

$$\mathcal{T}_x g(x,y)|_{x=X} = \mathcal{T}(g(\bullet, y))(X),\ \mathcal{T}_y g(x,y)|_{y=Y} = \mathcal{T}(g(x,\bullet))(Y), \tag{21}$$

we are lead to the expression $T_n = \frac{1}{n} \sum_{i,j=1}^n S(X_i, X_j)$ with

$$S(X,Y) = \mathcal{T}_x \mathcal{T}_y \hat{a}(x) \hat{a}(y) C(V(x), V(y))|_{x=X, y=Y} \tag{22}$$

that points out that T_n is a second-order U-statistic.

6.2 Critical regions and power.

The critical region $T_n > \kappa(\alpha)$ with κ defined by

$$\mathbf{P} \left\{ \int_0^1 \int_0^1 C(t,u) dw(t) dw(u) > \kappa(\alpha) \right\} = \alpha$$

provides a test for $\mathcal{H}_0$ consistent under any fixed alternative, with asymptotic level equal to α. Its asymptotic power is

$$\pi(\delta) = \mathbf{P}\left\{\int_0^1 \int_0^1 C(t,u)(dw(t) + \delta dt)(dw(u) + \delta du) > \kappa(\alpha)\right\}. \tag{23}$$

The values of $\kappa(\alpha)$ and $\pi(\delta)$ indicated in tables 1 and 2, were obtained by simulations based on 8000 replications. The table indicates two other series of power values, with the purpose of comparison: the computed asymptotic power π^* of the modified Kolmogorov-Smirnov test introduced in [2] for optimum score function, and the asymptotic power $\hat{\pi}(\delta) = 1 - \Phi(\Phi^{-1}(1-\alpha/2) - \delta) + \Phi(\Phi^{-1}(\alpha/2) - \delta)$ of the two-sided test with critical region $|\Lambda| >$constant, where Λ denotes the typified logarithm of the likelihood ratio. It will be noticed that the performance of all three tests is very much the same.

α	1%	2.5%	5%	10%
$\kappa(\alpha)$	3.78	2.97	2.40	1.84

Table 1: Numerical approximation of the critical values $\kappa(\alpha)$ for sizes $\alpha = 1$, 2.5, 5 and 10%.

δ	0	0.2	0.4	0.6	0.8	1.0	1.2	1.4
$\pi(\delta)(\%)$	5.0	5.3	6.6	8.9	12.3	16.8	22.3	28.7
$\pi^*(\delta)(\%)$	5.0	5.4	6.7	8.9	12.0	16.2	21.3	27.3
$\hat{\pi}(\delta)(\%)$	5.0	5.5	6.9	9.2	12.6	17.0	22.4	28.8
δ	1.6	1.8	2.0	2.2	2.4	2.6	2.8	3.0
$\pi(\delta)(\%)$	35.7	43.2	52.0	59.9	67.2	74.3	80.2	85.0
$\pi^*(\delta)(\%)$	34.1	41.4	49.1	56.9	64.4	71.4	77.7	83.1
$\hat{\pi}(\delta)(\%)$	36.0	43.6	51.6	59.5	67.0	73.9	80.0	85.1

Table 2: Asymptotic powers $\pi(\delta)$ of the proposed test, $\pi^*(\delta)$ of the Modified K-S Test with optimum score function, and $\hat{\pi}(\delta)$ of the two-sided test based on the logarithm of the likelihood ratio, for a level of significance of 5%, as a function of δ.

6.3 Example: Goodness-of-fit to standard normal.

We finally indicate for completeness (see [3] for details) how to compute the statistic $S(X,Y)$ defined in (22) for the particular isometry

$$\mathcal{T}g = g - \int_{-\infty}^{\bullet} \frac{g(t)d\Phi(t)}{1-\Phi(t)},\ \mathcal{T}^{-1}h = h + \frac{1}{1-\Phi(\bullet)}\int_{-\infty}^{\bullet} h(t)d\Phi(t)$$

in two simple cases:

6.3.1 Case 1: test sensitive to changes in position.

We assume $F_0(x) = \Phi(x) = \int_{-\infty}^x \varphi(t)dt$, $\varphi(x) = \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-x^2/2}$, $f^{(\tau)}(x) = \frac{\varphi(x-\tau)}{\varphi(x)} = \mathrm{e}^{-\frac{\tau^2-2x\tau}{2}}$, hence $k(x) = \lim_{\tau\to 0^+} \frac{2}{\tau}(\mathrm{e}^{-\frac{\tau^2-2x\tau}{4}} - 1) = x$, $\|k\|^2 = \int_{-\infty}^{\infty} x^2\varphi(x)dx = 1$, $\hat{a}(x) = \mathcal{T}^{-1}(x) = x + \frac{1}{1-\Phi(x)}\int_{-\infty}^x t\varphi(t)dt = x - \frac{\varphi(x)}{1-\Phi(x)}$, and therefore $V(x) = \int_{-\infty}^x \left(t - \frac{\varphi(t)}{1-\Phi(t)}\right)^2 \varphi(t)dt = \int_{-\infty}^x t^2\varphi(t)dt + \frac{\varphi^2(x)}{1-\Phi(x)}$. Once we have the analytical expressions of C, $\hat{a}$ and V, S can be computed by means of a simple algorithm, involving numerical integration.

6.3.2 Case 2: test sensitive to changes in dispersion.

Let $F^{(\tau)}(x) = \Phi((1-\frac{\tau}{\sqrt{2}})x)$, so that $f^{(\tau)}(x) = (1-\frac{\tau}{\sqrt{2}})\varphi((1-\frac{\tau}{\sqrt{2}})x)/\varphi(x)$ $= (1-\frac{\tau}{\sqrt{2}})\mathrm{e}^{\frac{1}{2}x^2(1-(1-\frac{\tau}{\sqrt{2}})^2)}$, $k(x) = \lim_{\tau\to 0^+}\frac{2}{\tau}\left(\sqrt{1-\frac{\tau}{\sqrt{2}}\mathrm{e}^{\frac{1}{4}x^2(\sqrt{2}\tau-\tau^2/2)}} - 1\right) = \frac{1}{\sqrt{2}}(x^2-1)$, and $\|k\|^2 = \frac{1}{2}\int_{-\infty}^{\infty}(x^2-1)^2\varphi(x)dx = 1$.

The score function is $\hat{a}(x) = \frac{\mathcal{T}^{-1}(x^2-1)}{\sqrt{2}} = \frac{x^2-1}{\sqrt{2}} + \frac{1}{1-\Phi(x)}\int_{-\infty}^x \frac{t^2-1}{\sqrt{2}}\varphi(t)dt$ $= \frac{1}{\sqrt{2}}\left[x^2 - 1 - \frac{x\varphi(x)}{1-\Phi(x)}\right]$, and consequently, $V(x) = \frac{1}{2}\int_{-\infty}^x \left[t^2 - 1 - \frac{t\varphi(t)}{1-\Phi(t)}\right]^2 dt$ $= \int_{-\infty}^x (t^2-1)^2\varphi(t)dt - \frac{x^2\varphi^2(x)}{1-\Phi(x)}$. As in the previous case, numeric integration can be used to compute each evaluation of S.

References

[1] Anderson, T. W. and Darling, D. A., *A test of goodness of fit.* J. Amer. Statist. Assoc., **49** (1954), 765-769.

[2] Cabaña, A., *Transformations of the empirical process and Kolmogorov-Smirnov tests.* Ann. Statist. **24**(1996) 2020-2035.

[3] Cabaña, A. and Cabaña, E.M., Transformed Empirical Processes and Modified Kolmogorov-Smirnov Tests for multivariate distributions, Ann. Statist. **25** (1997) 2388-2409.

[4] Cabaña, A. and Cabaña, E.M. (2000), *Goodness-of-fit tests based on quadratic functionals of Transformed Empirical Processes*, Statistics, Gordon and Breach, Berlin, to appear.

[5] Cabaña, A. and Cabaña, E.M. (unpublished), *Normality tests based on Transformed Empirical Processes.*

[6] Groeneboom, P. and Wellner, J.A., *Information Bounds and Nonparametric Maximum Likelihood Estimation.*, DMV Seminars, Band 19, Birkhauser, New York, 1992.

[7] Hájek, J. and Sidák, Z. (1967), *Theory of Rank Tests.*, Academic Press, New York.

[8] Le Cam, L. and Yang, G.L. *Asymptotics in Statistics. Some basic concepts.* Springer-Verlag, New York, 1990.

[9] Oosterhoff, J; van Zwet, W. R. *A note on contiguity and Hellinger distance.* Contributions to Statistics, 157-166, Reidel, Dordrecht, 1979.

Alejandra Cabaña
acabana@ivic.ve
Instituto Venezolano de Investigaciones Científicas (IVIC),
Departamento de Matemática
y
Universidad Simón Bolívar,
Departamento de Matemática.
IVIC-Matematica, apartado 21827 Caracas 1020-A,
Venezuela.

Enrique M. Cabaña
ecabana@chaja.ccee.edu.uy
Facultades de Ciencias y de Ciencias Económicas y de Administración,
Universidad de la República.
CSIC - Juan D. Jackson 1303, 11200 Montevideo,
Uruguay.