Statistica Sinica **17**(2007), 713-724

A TEST FOR CONSTANCY OF ISOTONIC REGRESSIONS USING THE L_2-NORM

Ana Colubi, J. Santos Domínguez-Menchero and Gil González-Rodríguez

Universidad de Oviedo

Abstract: In this paper we present a nonparametric procedure for testing the constancy of an isotonic regression. We introduce a family of statistics based on the L_2-norm of the difference of an isotonic estimate of the regression and the estimate under the (null) hypothesis that it is a constant. We propose to choose as isotonic estimate a tail-smoothed version of the usual least squares isotonic regression estimate. We write the selected statistic in terms of a certain functional in order to analyze its asymptotic distribution from the continuity properties of the functional. Finally, we show some simulations to compare the stated procedure with the well-known parametric $\overline{\chi}^2_{01}$.

Key words and phrases: Isotonic regression, L_2-norm, testing constant regression.

1. Introduction

Consider the regression model

$$Y(x_{i,n}) = m(x_{i,n}) + \varepsilon(x_{i,n}) \quad 1 \leq i \leq n,$$

where the design $\{x_{1,n}, \ldots, x_{n,n}\} \subset A$ is assumed to be fixed, $x_{i,n} < x_{j,n}$ ($1 \leq i < j \leq n$) and $A \subseteq \mathbb{R}$ is an interval. For $x \in A$, the random variable $Y(x)$ has finite expectation $m(x)$ and finite variance $\sigma^2(x)$. The errors $\varepsilon(x_{i,n})$ form a triangular array of row-wise independent random variables. The regression function m is assumed to be isotonic and we are interested in testing

$$H_0 : m \text{ is constant} \quad \text{versus} \quad H_1 : m \text{ is isotonic and not constant.}$$

The likelihood ratio tests $\overline{\chi}^2_{01}$ and $\overline{E}^2_{01}$ can be applied to solve the related problem for finitely many populations if the conditional distributions are normal; the first one requires known conditional variances, and the second one a partial knowledge of the variance function (see, for instance, Barlow, Bartholomew, Bremmer and Brunk (1972) and Robertson, Wright and Dykstra (1988)). In this paper we present an asymptotic test for isotonic regression functions defined on a continuum.

For each $x \in A$, $D(x)$ denotes the distribution of $Y(x)$. We consider r independent observation $Y^1(x_{i,n}), \ldots, Y^r(x_{i,n})$ on each design point $x_{i,n}$ ($1 \leq i \leq n$). If T_n is any estimate of m under the isotonicity assumption, we define

$$\overline{DIP}_{T_n} = \min_{a \in \mathbb{R}} \sum_{i=1}^{n} \frac{1}{n} \left(T_n(x_{i,n}) - a \right)^2 .$$

The statistic $\overline{DIP}_{T_n}$ is connected to $DIP_{\hat{m}}$ (see Cuesta, Domínguez-Menchero and Matrán (1995)), defined as the minumum of $\sum_{i=1}^{n} \left(h(x_{i,n}) - \hat{m}(x_{i,n}) \right)^2 / n$ on the set of the isotonic functions h, where $\hat{m}(x_{i,n}) = \sum_{j=1}^{r} Y_{i,n}^{j} / r$. The $DIP_{\hat{m}}$ statistic is used in Domínguez-Menchero, González-Rodríguez and López-Palomo (2005) to test whether a general regression function is isotonic (null hypothesis) or not.

If f is a real function, the argmin of $\sum_{i=1}^{n} \left(h(x_{i,n}) - f(x_{i,n}) \right)^2 / n$ on the set of the isotonic functions is the well-known isotonic regression f_I, which can be computed by means of PAVA (see Ayer, Brunk, Ewing, Reid and Silverman (1955). The usual isotonic regression estimate is precisely $\hat{m}_I$ and it can be expressed in terms of the Greater Convex Minorant, GCM, of the Cumulative Cum Diagram, CSD, as follows.

Let $B[0,1]$ (resp. $F[0,1]$) be the set of bounded (resp. convex) functions on $[0,1]$. We define $\text{GCM} : B[0,1] \to F[0,1]$ so that $\text{GCM}(x)$ is the greatest convex minorant of x. If $x \in F[0,1]$, we denote by $S(x)$ the function which associates each $t \in [0,1]$ with the left-hand slope of x at the point t (defined by continuity at the point $t = 0$). The CSD of a real function f is denoted by CSD_f, and it is defined at each $t \in [0,1]$ by linear interpolation from the values $\text{CSD}_f(0) = 0$ and $\text{CSD}_f(i/n) = \sum_{k=1}^{i} f(x_{k,n})/n$ with $1 \leq i \leq n$. Thus, we see that $\hat{m}_I(x_{i,n})$ coincides with $S(\text{GCM}(\text{CSD}_{\hat{m}}))$ computed at the point i/n (see, for instance, Robertson, Wright and Dykstra (1988)).

In this paper we choose as an estimate of m under isotonicity a *tail-smoothed* modification of $\hat{m}_I$. Concretely, given $[a,b] \subseteq [0,1]$, we define

$$\hat{m}^{a,b}(x_{i,n}) = \begin{cases} \frac{\sum_{k=1}^{\lceil na \rceil} \hat{m}(x_{k,n})}{\lceil na \rceil} & \text{if } i \in \{1, \ldots \lceil na \rceil\} \\ \hat{m}(x_{i,n}) & \text{if } i \in \{\lceil na \rceil + 1, \ldots, \lceil nb \rceil\} \\ \frac{\sum_{k=\lceil nb \rceil + 1}^{n} \hat{m}(x_{k,n})}{n - \lceil nb \rceil} & \text{if } i \in \{\lceil nb \rceil + 1, \ldots, n\} \end{cases},$$

(where $\lceil x \rceil$ is the least integer greater than or equal to x) and we consider its isotonization $T_n = \hat{m}_I^{a,b}$. Thus, the statistic that we propose is $\overline{DIP}_{\hat{m}_I^{a,b}} = \sum_{i=1}^{n} \left(\hat{m}_I^{a,b} - \overline{Y}_n \right)^2 / n$, where $\overline{Y}_n = \sum_{i=1}^{n} \hat{m}(x_{i,n})/n$. This statistic will be called the *smoothed* $\overline{DIP}$ and it depends on the parameters $a, b \in [0,1]$ that determine the smoothing tail points. It should be recalled that, although $\hat{m}_I$ is a consistent estimate of the isotonic regression m, its behaviour is not generally good at the tails (see, for instance, Brunk (1970), or Cuesta, Domínguez and Matrán (1995)). Therefore, a reasonable selection for $a, b \subset [0,1]$ can provide robustness to the procedure. On the other hand, some simulation studies (see Section 5) have shown us that different reasonable values for a, b lead to similar conclusions and, consequently, it seems that the determination of such parameters is usually easy in practice.

In Section 3 we express the *smoothed* $\overline{Dip}$ as a composition of a functional H with a certain process in order to find its asymptotic distribution by means of the Continuous Mapping Theorem. This functional H has been analyzed by Groeneboom and Pyke (1983) in another context. We prove that it is not continuous with respect to the supremum metric when it is defined on the spaces considered in the mentioned paper, and we find a useful subspace on which continuity is attained.

The paper is organized as follows. Section 2 is devoted to the study of the functional H. In Section 3 we establish the relationship between the tail-smoothed statistic and H, and we obtain the asymptotic results that we need to state the testing procedure. Finally, in Section 4 we show some simulation studies concerning the power of the tests. To be precise, we compare the practical behaviour of the parametric $\overline{\chi}^2_{01}$ test with the one stated in this paper by considering different smoothing degrees. The Appendix provides the proofs.

2. The functional H.

In order to define the functional H on a space useful in our framework, we consider a subset of the space of bounded functions $B[0,1]$ delimited by some functions that control the evolution of the slopes of the elements $x \in B[0,1]$ close to 0 and 1. Define

$$f_\delta(t) = \begin{cases} t^{\frac{1}{2}+\delta} & \text{if } t \in [0, \frac{1}{2}] \\ (1-t)^{\frac{1}{2}+\delta} & \text{if } t \in (\frac{1}{2}, 1] \end{cases} \qquad \text{for all } \delta \in [0, \frac{1}{2}],$$

$$B^*[0,1] = \{x : [0,1] \to \mathbb{R} \mid \text{ there exists } \delta \in (0, \frac{1}{2}] \text{ so that } |x| \leq f_\delta/\delta\},$$

and $H(x)$ as the L_2-norm of the left-hand slope of the greatest convex minorant of x, that is,

$$H(x) = \Big(\int_{[0,1]} [S(\text{GCM}(x))]^2(t)dt \Big)^{\frac{1}{2}} \text{ for all } x \in B^*[0,1].$$

The space $B^*[0,1]$ is a subset of the one considered in Groeneboom and Pyke (1983) to work with H. The condition defining $B^*[0,1]$ could be weakened, although it contains enough functions for our purposes. We endow $B^*[0,1]$ with the supremum norm $\|\cdot\|_\infty$. We show that H is not continuous on $B^*[0,1]$, although we find a useful subspace on which continuity holds. We base the study of H on some properties of S and GCM.

The function GCM is uniformly continuous. On the other hand, if the restriction of GCM to $B^*[0,1]$ is denoted by GCM*, it is easy to check that the range of GCM* is the next subspace of the convex functions $F[0,1]$:

$$F^*[0,1] = \{x \in F[0,1] | \text{ there exists } \delta \in (0,1/2] \text{ so that } |x| \leq f_\delta/\delta\}.$$

It should be noted that $F^*[0,1]$ is a subset of $C[0,1]$, the set of continuous functions on $[0,1]$.

The restriction of S to $F^*[0,1]$ is represented by S^*. Basic properties in convex analysis assure that for all $x \in F^*[0,1]$, $S^*(x)$ is an isotonic function from $[0,1]$ to $\overline{\mathbb{R}}$, left-hand continuous on $(0,1]$, and right-hand continuous at 0. Moreover, we have the following.

Proposition 1. *Let $x \in F^*[0,1]$, then $\int_{[0,1]}[S^*(x)]^2(t)dt < \infty$.*

Thus, if we denote Lebesgue measure on $[0,1]$ by λ, we have that the range of S^* is contained in

$$L_2^{\uparrow}[0,1] = \{z : [0,1] \to \overline{\mathbb{R}} \mid z \text{ isotonic, left-hand continuous on (0,1],}$$
$$\text{right-hand continuous at 0 and } \lambda\text{-square integrable}\}.$$

Given the relation between S^* and H, we consider the L_2-norm on the range of S^* in order to analyze its continuity. Thus, S^* is defined from $(F^*[0,1], \|\cdot\|_\infty)$ to $(L_2^{\uparrow}[0,1], \|\cdot\|_2)$. In Example 1 we show that S^* is not continuous, and we illustrate the cause of this lack of continuity.

Example 1. Let $x \in F^*[0,1]$ be such that $x(t) = 0$ for all $t \in [0,1]$, then $S^*(x)(t) = 0$ for all $t \in [0,1]$. On the other hand, consider

$$x_n^\alpha(t) = \begin{cases} -n^\alpha t & \text{if} \quad t \in [0, \frac{1}{n}], \\ n^{\alpha-1} & \text{if} \quad t \in (\frac{1}{n}, 1-\frac{1}{n}), \\ n^\alpha(t-1) & \text{if} \quad t \in [1-\frac{1}{n}, 1]. \end{cases}$$

We can easily verify that $x_n^\alpha \overset{n\to\infty}{\longrightarrow} x$ uniformly for all $\alpha \in [0,1)$. Nevertheless,

$$\left(\int_{[0,1]} \left[S^*(x_n^\alpha)(t) - S^*(x)(t)\right]^2 dt\right)^{\frac{1}{2}} = 2^{\frac{1}{2}} n^{\alpha-\frac{1}{2}}$$

for all $\alpha \in [0,1)$, and hence $S^*(x_n^\alpha) \overset{n\to\infty}{\longrightarrow} S^*(x)$ in L_2-norm if, and only if, $\alpha \in [0,1/2)$. Thereby S^* is not continuous. In this example we note that the slope of the functions x_n^α close to the extremes of the interval $[0,1]$ goes to infinity too fast, which makes convergence fail.

In Proposition 2 we provide conditions that guarantee the "sequential continuity" of S^*.

Proposition 2. *Let $x, \{x_n\}_{n\in\mathbb{N}} \subset F^*[0,1]$ so that $x_n \stackrel{n\to\infty}{\longrightarrow} x$ uniformly. If there exists $\tilde{\delta} \in (0,1/2]$ with $|x| \le f_{\tilde{\delta}}/\tilde{\delta}$ and $|x_n| \le f_{\tilde{\delta}}/\tilde{\delta}$ for all $n \in \mathbb{N}$, then $S^*(x_n)$ converges to $S^*(x)$ in L_2-norm.*

In the next example we prove that the restriction of H to the space $B^*[0,1]$ is not continuous (and, hence, is not continuous on the space considered in Groeneboom and Pyke (1983)).

Example 2. Let $x, \{x_n^\alpha\}_{n\in\mathbb{N}} \subset B^*[0,1]$, the functions defined in the Example 1. Then $x_n^\alpha \stackrel{n\to\infty}{\longrightarrow} x$ uniformly for all $\alpha \in [0,1)$, but $H(x_n^\alpha) = 2^{1/2} n^{\alpha-1/2}$ and $H(x) = 0$. Hence $H(x_n^\alpha)$ converges to $H(x)$ if, and only if, $\alpha \in [0,1/2)$.

On this basis, we suggest restricting H in order to get a continuous function. Let $\alpha \in (0,1/2]$, $k \in (0,1/2]$ and take

$$B^*[0,1]^{\alpha,k} = \left\{ x \in B[0,1] \text{ so that } |x\varphi_{[0,\alpha)\cup(1-\alpha,1]}| \le f_k/k \right\},$$

where φ_B denotes the characteristic function of a set $B \subset \mathbb{R}$. Let H^* be the restriction of H from the space $(B^*[0,1]^{\alpha,k}, \|\cdot\|_\infty)$ to $(\mathbb{R}, |\cdot|)$.

Theorem 1. *Let $\alpha \in (0,1/2]$ and $k \in (0,1/2]$, then we have $B^*[0,1]^{\alpha,k} \subset B^*[0,1]$ and the mapping H^* is continuous.*

3. Asymptotic Results

In order to express the *smoothed* $\overline{Dip}$ as a function of H, we define $\Pi_{[a,b]} : C[0,1] \to C[0,1]$ as the *trimming mapping* that associates each $x \in C[0,1]$ with the function equal to x on $[a,b]$, taking on the value 0 at 0 and 1, and defined by linear interpolation at the remaining points. The trimming mapping takes functions in $B^*[0,1]$ to $B^*[0,1]^{\alpha,k}$ for fixed $\alpha \in (0,1/2]$ and $k \in (0,1/2]$. Define $Z_n(t) = \mathrm{CSD}_{\hat{m}}(t) - t\,\mathrm{CSD}_{\hat{m}}(1)$ for all $t \in [0,1]$.

Proposition 3. *If $[a,b] \subseteq [0,1]$, then we have $\overline{DIP}_n^{a,b} = H^*(\Pi_{[\lceil na\rceil/n),\lceil nb\rceil/n]}(Z_n))$.*

We assume that the empirical distribution function of the design points converges to some continuous distribution function F uniformly on A in order to find the asymptotic distribution of Z_n. Additionally, to make the test operative under H_1, we assume F has a bounded positive density, bounded away from 0.

Proposition 4. *If the conditional variance function σ^2 is continuous and bounded, and the family of distribution functions $\{D(x)\}_{x\in A}$ satisfies*

$$\sup_{x\in A} \int (y - m(x))^2 \varphi_{[-\tau,\tau]^c}(y - m(x)) D(x)(dy) \stackrel{\tau\to\infty}{\longrightarrow} 0,$$

then $(nr/h(1))^{1/2}\,[Z_n - E[Z_n]] \xrightarrow{\mathcal{L}} B \circ U - I \cdot B(1)$, *where* B *is a Brownian motion on* $[0,1]$, I *denotes the identity function and*

$$h(t) = \int_{[0,t]} \sigma^2(F^{-1}(\tau))\lambda(d\tau), \quad U(t) = h(t)/h(1) \quad \text{for all } t \in [0,1].$$

As a result of Propositions 3 and 4 we get the next theorem.

Theorem 2. *Let* $[a,b] \subset (0,1)$. *Under the conditions of Proposition* 4, *if the regression function* m *is constant,* $(nr/h(1))^{1/2}\,\overline{DIP}_n^{a,b} \xrightarrow{\mathcal{L}} H[\Pi_{[a,b]}(B \circ U - I \cdot B(1))]$, *where* B *is a Brownian motion on* $[0,1]$.

The limit process in Theorem 2 is linked to the results in Groeneboom and Pyke (1983). They showed that the standarized functional H, evaluated on a sequence of truncated Brownian bridges, converges very slowly to a Gaussian distribution, though the convergence rate improves substantially when using a square root transformation. Although the level of smoothed/truncation is fixed here, the convergence rate is similar to the one in Groeneboom and Pyke (1983) if the smoothed level is low, and it is similar to that of the Central Limit Theorem if the smoothed level is high.

Theorem 2 provides the basis for the testing procedure.

Theorem 3. *Under the conditions in Theorem* 2, *for testing* H_0, m *is constant, at asymptotic level* $\alpha \in [0,1]$ *against the alternative* H_1, m *is isotonic and not constant, reject* H_0 *if* $(nr/h(1))^{1/2}\,\overline{DIP}_n^{a,b} > c_\alpha$, *where* c_α *is the* $100(1-\alpha)$*-fractile of the distribution of* $H[\Pi_{[a,b]}(B \circ U - I \cdot B(1))]$.

Remark 1. The preceding results require complete knowledge of the variance function. In practice, this function is unknown, although it is usually assumed to be a constant θ^2. In this case, in Proposition 4 we have that $h(t) = \theta^2 t$ and $U(t) = t$. Consequently, if S_n^2 is a consistent estimate of θ^2 (see, for instance, Gasser, Sroka and Jennen-Steinmetz (1986), Hall, Kay and Titterington (1990), or the simple variance estimate, that is consistent under H_0), we can substitute $h(1)$ by S_n^2 in Theorem 3 to get the testing procedure in this set-up on the basis of Slutsky's Theorem.

Remark 2. If there are different number of observations $r_{i,n}$ on the design points $x_{i,n}$, we can substitute the empirical weights $1/n$ by the usual normalized weights $r_{i,n}(\sum_{j=1}^n r_{j,n})^{-1}$ whenever they appear. Let $r_{\max}(n)$ (resp. $r_{\min}(n)$) be the maximum (resp. minimum) of the number of observation at each design point. If $r_{\max}(n)/r_{\min}(n)$ converges to 1 then, under the conditions in Theorem 2, we can assure that the asymptotic distribution of $(nr_{\max}(n)/h(1))^{1/2}\,\overline{DIP}_n^{a,b}$ is $H[\Pi_{[a,b]}(B \circ U - I \cdot B(1))]$.

Remark 3. The testing procedure can be easily extended to random design models if the independent variable X is continuous. In a random design, F is the distribution function of X and $r = 1$ *a.s.* for all $n \in \mathbb{N}$. To get the asymptotic results we just need to apply those for the fixed design on the conditioned sample by the ordered statistic of X and the Dominated Convergence Theorem (see, for instance, Stute (1993)). The function F is unknown, although F_n converges uniformly to F a.s., and from Proposition 4 we have that $h_n(t) = \int_{[0,t]} \sigma(F_n^{-1}(\tau))\lambda(d\tau)$ converges uniformly a.s. to $h(t)$ and $U_n(t) = h_n(t)/h_n(1)$ converges uniformly a.s. to $U(t)$. Thus, arguments in the proof of the Proposition 4 allow us to conclude that in practice H_0 should be rejected whenever $(nr/h_n(1))^{1/2}\,\overline{DIP}_n^{a,b} > c_{n,\alpha}$, where $c_{n,\alpha}$ is the $100(1-\alpha)$-fractile of the distribution of $H[\Pi_{[a,b]}(B \circ U_n - I \cdot B(1))]$.

4. Simulation Studies

In this section we compare the $\overline{\chi}^2_{01}$ test and the one introduced here by simulating several models. We fixed $A = [0,1]$, $x_{i,n} = i/(n+1)$, the sample size at $n = 500$, and the number of observation per design point at $r = 5$.

We considered two regression models m_1 (non-continuous) and m_2 (continuous), both depending on a parameter δ which indicates the "deviation" from the null hypothesis. As well, we considered two distributions for the random errors ξ, namely a standard normal and an uniform distribution on $[-\sqrt{3},\sqrt{3}]$. Thus, the conditional distribution $D_1(x)$ was the distribution function of $\delta s_1(x) + \xi$, where

$$s_1(x) = \begin{cases} -1 & \text{if } x \le \frac{1}{2}, \\ 1 & \text{if } x > \frac{1}{2}. \end{cases}$$

On the other hand, $D_2(x)$ was the distribution function of $\delta s_2(x) + \xi$, where

$$s_2(x) = \begin{cases} (x - \frac{5}{8}) \Big/ \sqrt{\frac{37}{192}} & \text{if } x \le \frac{1}{2}, \\ (2x - \frac{9}{8}) \Big/ \sqrt{\frac{37}{192}} & \text{if } x > \frac{1}{2}. \end{cases}$$

We have that $\sigma_i^2(x) = 1$ and $m_i(x) = \delta s_i(x)$ for all $x \in [0,1]$ $i = 1,2$. Obviously if $\delta = 0$ the regression function is constant, and if $\delta > 0$, then H_0 is not true and $\inf_{t\in A} \int_{[0,1]} (m_i(x) - t)^2 = \delta$.

The tests were done at a nominal significance level $\alpha = 0.05$. By following the approximate procedure based on two cumulants in Robertson, Wright and Dykstra (1988) related to the $\overline{\chi}^2_{01}$ test under the described conditions, we checked that the null hypothesis was to be rejected whenever the value of the statistic was greater than 13.7304.

In connection with the *smoothed* $\overline{Dip}$, we chose two natural smoothing degrees over the two tails in order to look at practical behaviour. Concretely, we smoothed at 5% ($a = 0.05, b = 0.95$), and at 10% ($a = 0.10, b = 0.90$). In both cases we have $U = I$, consequently $h = I$ and the asymptotic distribution of $(nr/h(1))^{1/2}\,\overline{DIP}_n^{a,b}$ under the null hypothesis is $H[\Pi_{[a,b]}(B^0)]$, where B^0 denotes a Brownian bridge on $[0,1]$. We approximated the asymptotic distributions with $n = 100,000$, and we found that the null hypothesis was to be asymptotically rejected whenever the value of the statistic was greater than 3.252 (resp. 3.034) for $a = 1-b = 0.05$ (resp. $a = 1-b = 0.10$).

Table 4.1. Model 1. Empirical power functions (×100) for Normal (a) and Uniform (b) cases.

(a)

δ	$\overline{\chi}^2_{01}$	$\overline{DIP}_n^{0.05,0.95}$	$\overline{DIP}_n^{0.10,0.90}$
0.00	4.95	4.03	4.33
0.02	14.47	15.02	17.03
0.04	36.44	40.63	44.93
0.06	67.23	73.21	77.28
0.08	90.28	93.51	95.11
0.10	98.64	99.26	99.51
0.12	99.90	99.97	99.98
0.14	100.00	100.00	100.00

(b)

δ	$\overline{\chi}^2_{01}$	$\overline{DIP}_n^{0.05,0.95}$	$\overline{DIP}_n^{0.10,0.90}$
0.00	4.65	4.01	4.32
0.02	14.19	14.99	16.90
0.04	36.34	40.60	44.96
0.06	67.52	73.45	77.55
0.08	90.49	93.57	95.12
0.10	98.64	99.29	99.52
0.12	99.91	99.97	99.99
0.14	100.00	100.00	100.00

Table 4.2. Model 2. Empirical power functions (×100) for Normal (a) and Uniform (b) cases.

(a)

δ	$\overline{\chi}^2_{01}$	$\overline{DIP}_n^{0.05,0.95}$	$\overline{DIP}_n^{0.10,0.90}$
0.00	4.99	4.15	4.38
0.02	18.57	19.39	20.96
0.04	47.54	52.08	55.27
0.06	79.40	83.68	85.82
0.08	95.91	97.38	97.94
0.10	99.63	99.82	99.89
0.12	99.98	99.99	100.00
0.14	100.00	100.00	100.00

(b)

δ	$\overline{\chi}^2_{01}$	$\overline{DIP}_n^{0.05,0.95}$	$\overline{DIP}_n^{0.10,0.90}$
0.00	4.67	4.02	4.31
0.02	18.18	19.45	31.03
0.04	47.36	51.93	54.99
0.06	79.58	83.79	85.95
0.08	95.95	97.42	97.96
0.10	99.63	99.82	99.87
0.12	99.99	99.99	99.99
0.14	100.00	100.00	100.00

In Tables 4.1 and 4.2 we show the percentages of rejections of 100,000 replications of the test for different values of δ for Models 1 and 2, respectively. The empirical significance levels of the $\overline{\chi}^2_{01}$ show that it is a conservative test. This conservative behaviour is also observed for the *smoothed* $\overline{Dip}$, although it could be probably improved by considering bootstrap techniques. In connection with the power of the tests, we can see that the power of our test improves on that of the $\overline{\chi}^2_{01}$ test (up to 25% in some cases) irrespective of the smoothing degree chosen, and even though our procedures are slightly more conservative than $\overline{\chi}^2_{01}$.

5. Concluding Remarks

In this paper we developed a new asymptotic procedure to test constancy for isotonic regressions based on a nonparametric tail-smoothed measure of the degree of increase of the isotonic regressions. This study offers a complementary view to the classical one in the sense it requires neither further assumptions on the conditional distributions nor any parametric model for the regression. In addition, simulation studies showed good power versus that of the classical tests in this context.

Acknowledgement

This research has been partially supported by the Spanish DGI/MCT Grant No. BFM2001-3494. The authors are deeply grateful to the anonymous referees whose remarks and suggestions have contributed to the improvement of the paper.

Appendix

Proof of Proposition 1. To prove the square integrability of $S^*(x)$ we locate a sequence of uniformly square integrable functions on $L_2^{\uparrow}[0,1]$ converging $\lambda-a.s.$ to $S^*(x)$, so that the mean convergence criterion would lead to the result. For all $n \in \mathbb{N}$, consider the piece-wise linear function $x^n : [0,1] \to \mathbb{R}$ defined by $x^n(i/2^n) = x(i/2^n)$ for all $i \in \{0,\ldots,2^n\}$, and by linear interpolation on $[(i-1)/2^n, i/2^n]$ for all $i \in \{1,\ldots,2^n\}$.

Note that $x^n \to x$ as $n \to \infty$, and given that $x \in F^*[0,1] \subset C[0,1]$, the convergence is uniform. Taking into account the convexity of the functions considered, it is easy to check that if x is differentiable at t_0, then $\{S(x_n)(t_0)\}_{n\in\mathbb{N}}$ converges to $S(x)(t_0)$. Since the functions in $F[0,1]$ are differentiable in $[0,1]$, except at most at a denumerable set, we get that $S^*(x^n) \overset{n\to\infty}{\longrightarrow} S^*(x)\ \lambda - a.s.$.

Regarding the uniform integrability of the sequence $\{[S^*(x^n)]^2\}_{n\in\mathbb{N}}$, we check that $\lim_{a\to\infty}\sup_{n\in\mathbb{N}}\int_{\{[S^*(x^n)]^2>a\}}[S^*(x^n)]^2(t)dt = 0$. Equivalently, if $x \in F^*[0,1]$, there exists $\delta \in (0,1/2]$ so that $|x| \le f_\delta/\delta$ and

$$\limsup_{k\to\infty}\sup_{n\in\mathbb{N}}\int_{\left\{|S^*(x^n)|>\delta^{-1}\left(2^k\right)^{1/2-\delta}\right\}}[S^*(x^n)]^2(t)dt = 0.$$

Obviously, we can assume that $\delta \in (0,1/2)$. It is easy to check that $\{|S^*(x^n)| > \delta^{-1}\left(2^k\right)^{1/2-\delta}\} \subset [0,1/2^k] \cup [1-1/2^k,1]$.

The integrals involved are monotone for piece-wise linear functions and, by symmetral reasoning, it is possible to verify that

$$\int_{\left[0,\frac{1}{2^k}\right]} [S^*((x^n))]^2(t)dt \leq \int_{\left[0,\frac{1}{2^k}\right]} [S^*(-\frac{f_\delta^n}{\delta})]^2(t)dt + \int_{\left[1-\frac{1}{2^k},1\right]} [S^*(-\frac{f_\delta^n}{\delta})]^2(t)dt.$$

The last expression can be bounded by $2\delta^{-2}\left[2^{-2k\delta}+2^{-2k\delta}(1/2+\delta)^2(2\delta)^{-1}\right]$ and this converges to 0 as $k\to\infty$.

The same reasoning can be applied in connection with the integral over $[1-1/2^k,1]$, and this completes the proof.

Proof of Proposition 2. It is enough to verify that $\{[S^*(x_n)]^2\}_n$ is uniformly integrable. This can be proven by following similar steps to those in Proposition 1, taking into account that the monotonicity result can be extended to general functions in $F^*[0,1]$.

Proof of Theorem 1. The first assertion is obvious. Regarding the continuity of H^*, consider $x,\{x_n\}_{n\in\mathbb{N}}\subset B^*[0,1]^{\alpha,k}$ such that $x_n\to x$ uniformly as $n\to\infty$. The boundedness of x and the uniform convergence of the sequence $\{x_n\}_{n\in\mathbb{N}}$ to x allow us to find a common $\tilde{\delta}\in(0,1/2]$ in such a way that $|x|\leq f_{\tilde{\delta}}/\tilde{\delta}$ and $|x_n|\leq f_{\tilde{\delta}}/\tilde{\delta}$ for all $n\in\mathbb{N}$. Furthermore, the convexity of $f_{\tilde{\delta}}/\tilde{\delta}$ guarantees that $|\text{GCM}(x)|\leq f_{\tilde{\delta}}/\tilde{\delta}$ and $|\text{GCM}(x_n)|\leq f_{\tilde{\delta}}/\tilde{\delta}$ for all $n\in\mathbb{N}$. Hence, the result can be deduced from the uniform continuity of GCM and the sequential continuity of S^* established in Proposition 2.

Proof of Proposition 3. First of all, note that since $\overline{Y}_n=\text{CSD}_{\hat{m}^{a,b}}(1)$, then $\overline{DIP}_n^{a,b}=H^*(\text{CSD}_{\hat{m}^{a,b}}-I\cdot\text{CSD}_{\hat{m}^{a,b}}(1))$. On the other hand, it is easy to check that $\text{CSD}_{\hat{m}^{a,b}}-I\cdot\text{CSD}_{\hat{m}^{a,b}}(1)=\Pi_{[\lceil na\rceil/n,\lceil nb\rceil/n)}(Z_n)$ and this suffices.

Proof of Proposition 4. First it should be note that $(nr/h(1))^{1/2}[Z_n-E[Z_n]]=(nr/h(1))^{1/2}\,\text{CSD}_{\hat{m}-m}-I\cdot(nr/h(1))^{1/2}\,\text{CSD}_{\hat{m}-m}(1)$. Since

$$\begin{array}{rl} z:(C[0,1],\|\cdot\|_\infty) & \to (C[0,1],\|\cdot\|_\infty) \\ x & \rightsquigarrow z(x)=x-I\cdot x(1) \end{array}$$

is continuous and $U(1)=h(1)/h(1)=1$, if we prove that $(nr/h(1))^{1/2}\,\text{CSD}_{\hat{m}-m}$ converges in law to $B\circ U$, the Continuous Mapping Theorem completes the proof.

To verify this convergence we apply a Prohorov generalization of the Donsker Theorem (see, for instance, Billingsley (1968)) on the centered variables $\xi_{i,n}=r^{1/2}(\hat{m}(x_{i,n})-m(x_{i,n}))$. We need to introduce the following notation: $\widetilde{s}_{i,n}^2=\text{Var}\,[\xi_{1,n}]+\ldots+\text{Var}\,[\xi_{i,n}]$, $\widetilde{S}_{i,n}=\xi_{1,n}+\ldots\xi_{i,n}$ and $\widetilde{u}_{i,n}=\widetilde{s}_{i,n}^2/\widetilde{s}_{n,n}^2$ for all $i=1,\ldots,n$ and $n\in\mathbb{N}$. Let $\mathbb{X}_n$ be the random function that is linear on each interval $[\widetilde{u}_{n,i-1},\widetilde{u}_{i,n}]$ and taking the values $\mathbb{X}_n(\widetilde{u}_{i,n})=\widetilde{S}_{i,n}/\widetilde{s}_{n,n}$ at the division points.

If the family of random variables $\{\xi_{i,n}\}_{i=1,\dots,n}$, $n \in \mathbb{N}$, satisfies the Lindeberg condition, then $\mathbb{X}_n \xrightarrow{\mathcal{L}} B$.

On the other hand, let $\widetilde{U}_n : [0,1] \to [0,1]$ be linear on each interval $[(i-1)/n, i/n]$ and take on the values $\widetilde{U}_n(i/n) = \widetilde{u}_{i,n}$ at the division points. Since $\mathrm{Var}\,[\xi_{i,n}] = \sigma^2(x_{i,n})$, it is easy to check that

$$\left(\frac{nr}{h(1)}\right)^{\frac{1}{2}} \mathrm{C{\small SD}}_{\hat{m}-m} = \left(\frac{s^2_{n,n}/n}{h(1)}\right)^{\frac{1}{2}} \mathbb{X}_n \circ \widetilde{U}_n.$$

The condition imposed on the design points guarantees the uniform convergence of $\widetilde{U}_n$ to U. In addition, since $\widetilde{U}_n$ is continuous and strictly increasing for all $n \in \mathbb{N}$, then $\mathbb{X}_n \circ \widetilde{U}_n \xrightarrow{\mathcal{L}} B \circ U$ (see, Brunk (1970)). Slutzky's Lemma concludes the proof.

Next we are going to verify that the family of random variables $\{\xi_{i,n}\}_{i=1,\dots,n}$, $n \in \mathbb{N}$, satisfies the Lindeberg condition. First, we prove that the family of supporting random variables $Q_{nij} = r^{-1/2}\left(Y^j(x_{i,n}) - m(x_{i,n})\right)$ with $n \in \mathbb{N}$, $i \in 1, \dots, n$ and $j \in \{1, \dots, r\}$ satisfies the Lindeberg condition, and later we connect the conditions related to both families (the original and the supporting one).

Let $\epsilon > 0$. For each $i \in \{1, \dots, n\}$, define $s_n^2 = \sum_{i=1}^n \sum_{j=1}^r \mathrm{Var}\,[Q_{nij}]$, and denote by $D_{\epsilon,n}(x_{i,n})$ the integration support for the Lindeberg condition for the family Q_{nij}, that is, $\left\{y \in \mathbb{R} \,\middle|\, r^{-1}\left(y - m(x_{i,n})\right)^2 > \epsilon^2 s_n^2\right\}$. It is easy to verify that given $\phi \in \mathbb{R}^+$, there exists $n_0 \in \mathbb{N}$ so that, for all $n \geq n_0$, $D_{\epsilon,n}(x_{i,n}) \subseteq \{y \in \mathbb{R} \,|\, |y - m(x_{i,n})| > \phi\}$ for all $i \in \{1, \dots, n\}$. Consequently,

$$\begin{aligned} &\frac{1}{s_n^2} \sum_{i=1}^n \sum_{j=1}^r \int_{D_{\epsilon,n}(x_{i,n})} r^{-1} D(x_{i,n})(dy) \\ &\quad \leq \frac{n}{s_n^2} \sup_{x \in A} \int_{\{y \in \mathbb{R} \,|\, |y-m(x)|>\phi\}} (y - m(x))^2 D(x)(dy). \end{aligned}$$

Thus, for all $\phi \in \mathbb{R}^+$ it is verified that

$$\begin{aligned} &\limsup_{n\to\infty} \frac{1}{s_n^2} \sum_{i=1}^n \sum_{j=1}^r \int_{D_{\epsilon,n}(x_{i,n})} r^{-1}\left(y - m(x_{i,n})\right)^2 D(x_{i,n})(dy) \\ &\quad \leq h(1)^{-1} \sup_{x \in A} \int_{\{y \in \mathbb{R} \,|\, |y-m(x)|>\phi\}} (y - m(x))^2 D(x)(dy). \end{aligned}$$

The conditions imposed on the family of distributions D guarantee that this last expression converges to 0 as ϕ tends to infinity. Consequently the sequence of random variables Q_{nij} verifies the Lindeberg condition and

$$\sum_{i=1}^n \xi_{i,n} \Big(\sum_{i=1}^n \mathrm{Var}\,[\xi_{i,n}]\Big)^{-\frac{1}{2}} = \sum_{i=1}^n \sum_{j=1}^r Q_{nij} \Big(\sum_{i=1}^n \sum_{j=1}^r \mathrm{Var}\,[Q_{nij}]\Big)^{-\frac{1}{2}} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1).$$

On the other hand, by taking into account Tchebychev's inequality it is easy to verify that for all $\epsilon > 0$,

$$\max_{i\in\{1,\ldots,n\}} P\Big[\Big|\xi_{i,n}\Big(\sum_{j=1}^{n}\mathrm{Var}\,[\xi_{j,n}]\Big)^{-1/2}\Big| > \epsilon\Big] \stackrel{n\to\infty}{\longrightarrow} 0.$$

The last two convergences are equivalent to the Lindeberg condition for the $\xi_{i,n}$, and this finishes the proof.

References

Ayer, M., Brunk, H. D., Ewing, W. T., Reid, W. T. and Silverman, E. (1955). An empirical distribution function for sampling with incomplete information. *Ann. Math. Statist.* **26**, 641-647.

Barlow, R. E., Bartholomew, D. J., Bremner, J. M. and Brunk, H. D. (1972). *Statistical Inference under Order Restrictions.* Wiley, New York.

Billingsley, P. (1968). *Convergence of Probability Measures.* Wiley, New York.

Brunk, H. D. (1970). Estimation of isotonic regression. *Nonparametric Techniques in Statistical Inference*, 177-195. Cambridge University Press.

Cuesta, J. A., Domínguez, J. S. and Matrán, C. (1995). Consistency of L_p-best monotone approximations. *J. Statist. Plann. Inference* **47**, 295-318.

Domínguez-Menchero, J. S., González-Rodríguez, G. and López-Palomo, M. J. (2005). An L_2 point of view in testing monotone regression, *J. Nonparametr. Stat.* **17**, 135-153.

Gasser, T., Sroka, L. and Jennen-Steinmetz, C. (1986). Residual variance and residual pattern in nonlinear regression. *Biometrika* **73**, 625-633.

Groeneboom, P. and Pyke, R. (1983). Asymptotic normality of statistics based on the convex minorants of empirical distribution functions. *Ann. Probab.* **11**, 328-345.

Hall, P., Kay, J. W. and Titterington, D. M. (1990). Asymptotically optimal difference-based estimation of variance in nonparametric regression. *Biometrika* **77**, 521-528.

Robertson, T., Wright, F. T. and Dykstra, R. L. (1988). *Order Restricted Statistical Inference.* Wiley, New York.

Stute, W. (1993). U-Functions of concomitants of order statistics. *Probab. Math. Statist.* **14**, 143-155.

Departamento de Estadística e I.O. y D.M. Facultad de Ciencias. Universidad de Oviedo. C Calvo Sotelo s/n. Oviedo 33007. Spain.

E-mail: colubi@uniovi.es

Departamento de Estadística e I.O. y D.M. Facultad de Ciencias. Universidad de Oviedo. C Calvo Sotelo s/n. Oviedo 33007. Spain.

E-mail: jsdm@uniovi.es

Departamento de Estadística e I.O. y D.M. Facultad de Ciencias. Universidad de Oviedo. C Calvo Sotelo s/n. Oviedo 33007. Spain.

E-mail: gil@uniovi.es

(Received September 2004; accepted September 2005)