

Test de hipótesis para contrastar la igualdad entre k -poblaciones

Hypothesis Test to Compare the Equality Among k -populations

PABLO MARTÍNEZ-CAMBLOR^a

FUNDACIÓN CAUBET-CIMERA ILLES BALEARS, MALLORCA, ESPAÑA

Resumen

Este trabajo estudia las ventajas y limitaciones de un test para contrastar la igualdad de las distribuciones de origen de k -muestras independientes. El estadístico propuesto, denominado LG_k , está basado en una medida que generaliza la norma L_1 entre funciones de densidad y que permite comparar simultáneamente k densidades. Desde esta medida y a partir de la estimación *kernel*, se desarrolla un test para contrastes de igualdad entre k poblaciones independientes (LG_k). A partir de un “amplio” estudio de simulación, se estudia la potencia del test propuesto y se compara con algunos de los test no paramétricos ya existentes, considerando ocho estadísticos diferentes. También se analiza el tema de la elección del tamaño del parámetro ventana y se realizan algunas propuestas relativas a este problema.

Palabras clave: estimación *kernel*, medida L_1 , selección del parámetro ventana, *bootstrap*.

Abstract

In this paper we study a test to contrast the equality among the origin distributions of k -independent samples. The proposed statistic, denoted as LG_k , is based in a measure which generalizes the L_1 -norm among density functions and it allows us to compare k -different densities. From this measure and the *kernel* density estimation, a k -sample test for independent populations is developed. We make a wide simulation study for the proposed test and we compare its power with other nonparametric k -sample test, by considering a total of eight different statistics. We also analyze the topic of the bandwidth selection and make the same proposals about this problem.

Key words: Kernel density estimation, L_1 Measure, Bandwidth selection, Bootstrap.

^aPrograma de epidemiología e investigación clínica. E-mail: martinez@caubet-cimera.es

1. Introducción

Las L_p medidas son sin duda las más populares cuando se pretenden comparar dos densidades. Dadas dos funciones f y g con $0 < p < \infty$, estas quedan definidas por

$$L_p(f, g) = \int |f - g|^p$$

En este trabajo, se presenta una medida entre funciones de densidad que, en cierto sentido, extiende la medida L_1 ($p = 1$) al caso de k funciones y, a partir de ella, se desarrolla un test de hipótesis para la comparación de k -muestras independientes.

Note que si f y g verifican que $\int g = \int f$, se tiene que

$$0 = \int (f - g) = \int (f - g)^+ - \int (f - g)^- \implies \int (f - g)^+ = \int (f - g)^-$$

Luego si $p = 1$, se obtiene la igualdad

$$\begin{aligned} \int |f - g| &= \int (f - g)^+ + \int (f - g)^- = 2 \int (f - g)^- = \\ &= 2 \int (f - g) I_{\{g \geq f\}} = 2 \left(1 - \int \min\{f, g\} \right) \end{aligned}$$

La función mínimo está definida para más de dos funciones; así pues, se puede extender la medida L_1 sin más que definir la medida

$$LG_k(f_1, \dots, f_k) = k \left(1 - \int \min\{f_1, \dots, f_k\} \right) \quad (1)$$

Esta medida puede utilizarse para contrastar la igualdad de k -poblaciones sin más que sustituir las funciones de densidad por algún estimador y, posteriormente calcular o, en su defecto, aproximar la distribución del estadístico resultante.

En la segunda sección de este trabajo, y dada $X = \{X_1, \dots, X_k\}$, donde cada X_i representa una muestra aleatoria simple de tamaño n_i , se estudia la consistencia del estadístico definido por

$$LG_k(X) = LG_k(f_{n_1}, \dots, f_{n_k}) \quad (2)$$

donde para cada $i \in 1, \dots, k$, f_{n_i} es el conocido estimador núcleo para la función de densidad referido a la muestra X_i .

La Estimación Núcleo para la Función de Densidad (ENFD) fue introducida por Rosenblatt (1956) y, para una muestra aleatoria simple, X , procedente de una variable aleatoria absolutamente continua queda definida por

$$f_n(X, t) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x_i - t}{h}\right)$$

donde K es conocida como función núcleo, usualmente elegida entre las funciones de densidad simétricas con media cero y varianza finita, y $h = h_{(n)}$ es una sucesión de números reales positivos denominados parámetro ventana.

La normalidad asintótica de $L_p(f_n, f)$ ha sido estudiada por Horvath (1991), si bien las hipótesis necesarias no garantizan la normalidad asintótica del estadístico para un amplio número de densidades. El caso $p = 1$ ha sido ampliamente estudiado por Devroye & Györfi (1985).

El problema general de comparar k poblaciones sin asumir ningún modelo paramétrico ha sido estudiado por un amplio número de autores; existen versiones para k muestras de muchos de los test clásicos de comparación de dos muestras. De la misma manera que el test de Kruskal-Wallis (Kruskal & Wallis 1952) generaliza el test de Mann-Whitney, Kiefer (1959) y Conover (1965) proponen generalizaciones para el test de Kolmogorov-Smirnov; Scholz & Stephens (1987), para el test de Anderson-Darling; y Kiefer (1959), para el de Cramér-Von Mises. Otros tests para comparar k -muestras han sido propuestos por Lewis (1972) o, más recientemente, por Zhang & Wu (2007).

En la sección 3 se proponen dos formas de aproximar la distribución del estadístico definido en (2) mediante remuestreo, obteniendo un nuevo test para comparación de k poblaciones. En la sección 4 se hace un amplio estudio de simulación en el que se comparan los resultados obtenidos por el nuevo test desarrollado con algunos de los existentes.

2. Propiedades

En este apartado se probará la consistencia de la medida LG_k cuando se comparan k ENFD procedentes de muestras independientes con idéntica distribución de origen. Para ello, se exigen algunas condiciones sobre los parámetros involucrados en la estimación y sobre la función de densidad real, aunque todas las hipótesis necesarias son muy laxas y la consistencia demostrada es válida para la mayoría de las situaciones que se presentan en la práctica.

Proposición 1. *Dadas k muestras aleatorias procedentes de poblaciones con idéntica distribución, si $h_i \rightarrow_n 0 \quad \forall i \in 1, \dots, k$, la función núcleo K , es una función de densidad continua y acotada, y la función de densidad real es continua, entonces, se tiene*

$$LG_k(X) = k \left(1 - \int \min\{f_{n_1}, \dots, f_{n_k}\} \right) \rightarrow 0 \quad c.s. \quad (3)$$

Demostración. En primer lugar, si se define $g_{n,k} = \min\{f_{n_1}, \dots, f_{n_k}\}$, se tiene por un lado la igualdad

$$\int \min\{f_{n_1}, \dots, f_{n_k}\} = \int g_{n,k}$$

y, por otro,

$$\begin{aligned} g_{n,k} &= f_{n_k} + (f_{n_1} - f_{n_k})I_{\{f_{n_1}=g_{n,k}\}} + (f_{n_2} - f_{n_k})I_{\{f_{n_2}=g_{n,k}\}} + \cdots \\ &= f_{n_k} + \sum_{i=1}^{k-1} (f_{n_i} - f_{n_k})I_{\{f_{n_i}=g_{n,k}\}} \end{aligned}$$

Obviamente,

$$\int g_{n,k} = 1 + \sum_{i=1}^{k-1} \int (f_{n_i} - f_{n_k})I_{\{f_{n_i}=g_{n,k}\}}$$

Para cada $i \in 1, \dots, k$ se tiene

$$0 \leq \int (f_{n_k} - f_{n_i})I_{\{f_{n_i}=g_{n,k}\}} \leq \int |f_{n_k} - f_{n_i}|$$

Desde las propiedades exigidas a h_{n_i} y a K , sin más que aplicar el Teorema de la Convergencia Dominada, se tiene que $\int |f_{n_k} - f_{n_i}| \rightarrow_n 0$ c.s. para cada $i \in 1, \dots, k$. Así

$$\sum_{i=1}^{k-1} \int (f_{n_i} - f_{n_k})I_{\{f_{n_i}=g_{n,k}\}} \rightarrow_n 0 \quad c.s.$$

luego

$$LG_k(X) = k \left(1 - \int \min\{f_{n_1}, \dots, f_{n_k}\} \right) = k \left(1 - \int g_{n,k} \right) \rightarrow 0 \quad c.s.$$

concluyendo así la demostración del resultado. \square

3. Método *bootstrap*

En esta sección se proponen dos planes de remuestreo que permiten utilizar la medida LG_k para realizar contrastes de igualdad entre densidades. En este contexto, en el que se supone que las poblaciones de origen son continuas, el método que mejor resultados obtiene es el *bootstrap* suavizado (Hall et al. 1989), basado en generar muestras desde la Función de Distribución Empírica Suavizada (FDES).

La FDES es introducida por Nadaraya (1964) y, dada una muestra aleatoria $X = \{x_1, \dots, x_n\}$, queda definida de la siguiente manera:

$$\tilde{F}_n(X, t) = \frac{1}{n} \sum_{i=1}^n \tilde{K} \left(\frac{t - x_i}{h} \right) \quad (4)$$

donde \tilde{K} es una función de distribución, usualmente elegida de modo que su derivada sea una función de densidad simétrica, de media nula y varianza finita;

$h = h_{(n)}$ es una sucesión de números reales positivos, conocidos como parámetros ventana que, en general, no son de la misma naturaleza que los utilizados en la ENFD.

Así las cosas, dadas k muestras aleatorias simples $X = \{X_1, \dots, X_k\}$ de tamaños n_1, \dots, n_k , respectivamente, se pretende construir una región crítica para el contraste:

$$\begin{cases} H_0 : f_i = f_j, \quad \forall i, j \in 1, \dots, k; \\ H_1 : \text{no } H_0. \end{cases} \quad (5)$$

Asumiendo que la hipótesis nula es cierta, si las muestras son independientes, es razonable plantear el siguiente plan de remuestreo:

A_1 . Desde la muestra conjunta X , se calcula la FDES definida en (5).

B_1 . Se generan muestras *bootstrap* X^B desde la distribución anterior y con los mismos tamaños muestrales que las de origen.

C_1 . Se aproxima la distribución de $LG_k(X)$ a partir de la de $LG_k(X^B)$.

Sin embargo, asumiendo que la hipótesis nula es cierta, también se tiene que si $n = \sum_{i=1}^k n_i$,

$$\widetilde{PF}_n(X, t) = \frac{1}{n} \sum_{i=1}^k n_i \widetilde{F}_{n_i}(X_i, t) \quad (6)$$

es un estimador adecuado para F ; luego también es razonable el siguiente plan de remuestreo:

A_2 . Desde la muestra conjunta X , se calcula el estimador dado en (6).

B_2 . Se generan muestras *bootstrap* X^B desde la distribución anterior y con los mismos tamaños muestrales que las de origen.

C_2 . Se aproxima la distribución de $LG_k(X)$ a partir de la de $LG_k(X^B)$.

Estudiando la construcción de ambos métodos de remuestreo, se llega rápidamente a la conclusión de que la única diferencia entre ambos radica en el parámetro ventana utilizado que, para tamaños de muestra medios, y cuando se supone la hipótesis nula cierta, son muy próximos; luego la diferencia en la práctica es casi inapreciable, si bien el segundo método no asume independencia entre las muestras, por lo que se puede utilizar para comparar densidades en muestras relacionadas. Este problema, aunque no es muy usual en la práctica, si se presenta en ocasiones: el estudio de la calidad de pruebas diagnósticas o la comparación de la evolución de las distribuciones de los ingresos de un país a lo largo del tiempo son algunos ejemplos.

4. Simulaciones

En esta sección se estudia la aplicación práctica de los métodos propuestos anteriormente. Para ello, se proponen varios modelos dentro de dos “ambientes” generales: simetría y asimetría de las distribuciones; en cada uno de ellos, se compara la potencia del test anteriormente propuesto (LG_k) con las versiones para k -muestras de tests clásicos, como la versión k -muestras del test de Kolmogorov-Smirnov, propuesta por Kiefer (1959); la del test de Anderson-Darling, dada por Scholz & Stephens (1987); la del test de Cramér-Von Mises (Kiefer 1959), cuyas expresiones vienen dadas respectivamente por

$$K_S = \sup_{t \in \mathbb{R}} \sum_{i=1}^k n_i (F_{n_i}(t) - F_n(t))^2$$

$$A_D = \sum_{i=1}^k n_i \int \frac{(F_{n_i}(t) - F_n(t))^2}{F_n(t)(1 - F_n(t))} dF_n(t)$$

$$C_M = \sum_{i=1}^k n_i \int (F_{n_i}(t) - F_n(t))^2 dF_n(t)$$

siendo F_{n_i} y F_n la Función de Distribución Empírica (FDE) de la i -ésima muestra y de la muestra conjunta, respectivamente. Además, se da la potencia obtenida por el test de Kruskal & Wallis (1952) y la obtenida por los tres tests basados en la función de verosimilitud y propuestos por Zhang & Wu (2007), cuyas expresiones son

$$Z_K = \sup_{t \in \mathbb{R}} \sum_{i=1}^k n_i Z_{n,i}(t)$$

$$Z_A = \sum_{i=1}^k n_i \int \frac{Z_{n,i}(t)}{F_n(t)(1 - F_n(t))} dF_n(t)$$

$$Z_C = \frac{1}{n} \sum_{i=1}^k \sum_{j=1}^{n_i} \log \left(\frac{n_i}{j - 0.5} - 1 \right) \log \left(\frac{n}{R_{ij} - 0.5} - 1 \right)$$

donde $n = n_1 + \dots + n_k$, R_{ij} es el rango en la muestra conjunta de x_{ij} y

$$Z_{n,i}(t) = F_{n_i}(t) \log \left(\frac{F_{n_i}(t)}{F_n(t)} \right) + (1 - F_{n_i}(t)) \log \left(\frac{1 - F_{n_i}(t)}{1 - F_n(t)} \right)$$

En todos los casos, la potencia se estima desde 2500 simulaciones y, para cada problema, la significación se estima a partir de 199 réplicas del método de Monte Carlo (mismo número que el usado en el clásico Anderson et al. (1994)) para los estadísticos basados en la función de distribución empírica y de 199 muestras *bootstrap* según el primer procedimiento de remuestreo descrito en la sección anterior para el estadístico LG_k . Para las simulaciones se han utilizado rutinas programadas en lenguaje R (R Development Core Team 2007) y se han considerado distintos tamaños muestrales, indicados en la parte izquierda de las tablas 1, 2, 3, 4, 5 y 6.

4.1. Simetría

En el primer tipo de problemas tratados, las densidades que se manejan son “simétricas”. Se tienen dos muestras procedentes de poblaciones con densidades f_1 y f_2 y tamaños muestrales n_1 y n_2 , procedentes de una distribución normal estándar mientras que la tercera muestra, de tamaño n_3 , procede de una variable Z , densidad f_3 , cuya distribución queda determinada por

MD 0-I: $Z \equiv N(0, 1)$ (Hipótesis nula)

MD 1-I: $Z \equiv (1 - a)N(0, 1) + aN(0, 2)$

MD 2-I: $Z \equiv (1 - a)N(0, 1) + aN(1, 1)$

MD 3-I: $Z \equiv (1 - a)N(0, 1) + aN(1, 2)$

MD 4-I: $Z \equiv (1 - a)N(0, 1) + aN(3, 3)$

MD 5-I: $Z \equiv a/2N(-2, 1/2) + (1 - a)N(0, 1) + a/2N(2, 1)$

MD 6-I: $Z \equiv (1 - a)N(0, 1) + a\chi_2^2$

donde, $n = n_1 + n_2 + n_3$ y $a = 6n^{-1/2}$.

Los resultados obtenidos para $\alpha = 0.05$ se muestran en la tabla 1; para $\alpha = 0.01$, en la tabla 2.

La significación (modelo MD 0-I) observada está dentro de lo esperado. Cuando el test propuesto se compara con las adaptaciones de los tests clásicos basados en la FDE, se observa que este es más potente para tamaños muestrales pequeños e iguales en todos los modelos propuestos, a excepción del modelo MD 2-I, donde el test más potente es el Kruskal-Wallis (K_W), si bien, dependiendo de los tamaños muestrales, las potencias observadas en los tests basados en la razón de verosimilitud (tests Z_A , Z_K y Z_C) son muy próximas (las mayores diferencias se dan en los modelos MD 1-I, MD 5-I y MD 6-I para $n_1 = n_2 = n_3 = 25$), esto es, siempre que las distribuciones difieren en algo más que un parámetro de posición. Los principales competidores del test propuesto son los basados en la razón de verosimilitud estudiados por Zhang & Wu (2007). Estos tests, a diferencia de los clásicos (K_S , C_V , K_W , A_D), detectan diferencias entre las distribuciones cuando estas no se diferencian en parámetros de posición. En este tipo de problemas, tienen la ventaja sobre el test LG_k de aumentar su potencia al aumentar el tamaño muestral; además, no se ven muy afectados cuando los tamaños muestrales son diferentes.

Considerando el rendimiento del test basado en el estadístico LG_k , destaca en primer lugar la gran pérdida de potencia que este sufre cuando los tamaños muestrales son distintos. Estas diferencias se acentúan cuanto mayor es la asimetría entre los mismos. En segundo lugar, destaca la pérdida de potencia general al aumentar los tamaños muestrales. Cabe la posibilidad de que estos problemas sean el efecto de una inadecuada elección de los parámetros ventana (para cada muestra, $i \in 1, 2, 3$, se ha elegido $h_i = \hat{\sigma}_i n_i^{-0.2}$, siendo $\hat{\sigma}_i$ la raíz cuadrada de

la varianza muestral en la muestra i). Es sabido que este parámetro tiene una gran influencia en el resultado de la estimación de la densidad y también en los resultados de los tests basados en este tipo de técnicas (Cao & Van Keilegom 2006). Esta posibilidad se explora más adelante.

TABLA 1: Proporción de rechazos observados en 2500 réplicas para $\alpha = 0.05$.

n_1	n_2	n_3	MD	K_S	A_D	C_M	K_W	Z_A	Z_K	Z_C	LG_k
25	25	25	0-I	0.056	0.050	0.049	0.042	0.051	0.049	0.052	0.052
			1-I	0.078	0.066	0.085	0.052	0.256	0.333	0.316	0.475
			2-I	0.424	0.348	0.363	0.650	0.532	0.633	0.635	0.466
			3-I	0.085	0.082	0.084	0.309	0.582	0.624	0.601	0.631
			4-I	0.202	0.172	0.191	0.825	0.987	0.981	0.980	0.991
			5-I	0.311	0.265	0.317	0.062	0.748	0.833	0.823	0.920
			6-I	0.069	0.057	0.068	0.260	0.630	0.748	0.722	0.820
25	50	50	0-I	0.056	0.053	0.050	0.051	0.051	0.053	0.052	0.050
			1-I	0.184	0.167	0.180	0.043	0.322	0.408	0.365	0.381
			2-I	0.199	0.157	0.144	0.674	0.561	0.650	0.655	0.400
			3-I	0.053	0.047	0.056	0.328	0.713	0.719	0.704	0.542
			4-I	0.085	0.084	0.074	0.845	0.999	0.996	0.995	0.991
			5-I	0.543	0.501	0.531	0.055	0.818	0.890	0.878	0.898
			6-I	0.277	0.202	0.208	0.900	0.913	0.950	0.948	0.686
25	50	75	0-I	0.049	0.049	0.044	0.048	0.046	0.043	0.045	0.040
			1-I	0.301	0.257	0.268	0.042	0.336	0.392	0.320	0.313
			2-I	0.059	0.034	0.030	0.668	0.571	0.648	0.652	0.345
			3-I	0.071	0.062	0.063	0.316	0.741	0.724	0.679	0.468
			4-I	0.044	0.036	0.039	0.861	0.999	0.998	0.996	0.985
			5-I	0.700	0.614	0.655	0.044	0.849	0.897	0.878	0.882
			6-I	0.086	0.048	0.047	0.919	0.936	0.954	0.952	0.610
50	50	50	0-I	0.056	0.053	0.050	0.051	0.051	0.053	0.052	0.050
			1-I	0.123	0.091	0.127	0.664	0.342	0.487	0.458	0.461
			2-I	0.219	0.195	0.189	0.633	0.537	0.633	0.649	0.458
			3-I	0.053	0.044	0.058	0.278	0.667	0.740	0.723	0.606
			4-I	0.077	0.076	0.076	0.813	0.997	0.997	0.996	0.994
			5-I	0.390	0.323	0.381	0.061	0.822	0.910	0.908	0.933
			6-I	0.301	0.250	0.265	0.874	0.904	0.952	0.947	0.759
50	100	100	0-I	0.056	0.055	0.052	0.056	0.055	0.054	0.059	0.051
			1-I	0.141	0.120	0.126	0.006	0.536	0.638	0.648	0.420
			2-I	0.050	0.036	0.036	0.678	0.596	0.673	0.670	0.343
			3-I	0.063	0.067	0.068	0.310	0.821	0.858	0.840	0.511
			4-I	0.062	0.059	0.060	0.854	1.000	1.000	0.999	0.992
			5-I	0.486	0.455	0.481	0.042	0.893	0.946	0.944	0.894
			6-I	0.134	0.101	0.093	0.896	0.958	0.980	0.976	0.701
50	100	150	0-I	0.061	0.053	0.051	0.046	0.048	0.051	0.050	0.052
			1-I	0.414	0.320	0.356	0.048	0.461	0.561	0.512	0.273
			2-I	0.104	0.085	0.077	0.654	0.578	0.658	0.656	0.374
			3-I	0.104	0.092	0.101	0.300	0.860	0.869	0.842	0.426
			4-I	0.063	0.062	0.063	0.875	1.000	1.000	0.999	0.983
			5-I	0.635	0.563	0.616	0.041	0.916	0.956	0.953	0.872
			6-I	0.053	0.040	0.039	0.920	0.978	0.984	0.982	0.585

4.2. Asimetría

La segunda familia de modelos propuesta pretende ser análoga a la primera en un “ambiente” de asimetría. Se generan dos muestras de tamaños respectivos n_1 y n_2 desde poblaciones cuyas densidades, f_1, f_2 , se corresponden con una χ_3^2 (distribución marcadamente asimétrica). Una tercera muestra de tamaño n_3 y

TABLA 2: Proporción de rechazos observados en 2500 réplicas para $\alpha = 0.01$.

n_1	n_2	n_3	MD	K_S	A_D	C_M	\bar{K}_W	Z_A	Z_K	Z_C	LG_k
25	25	25	0-I	0.014	0.014	0.011	0.010	0.000	0.008	0.009	0.010
			1-I	0.012	0.011	0.018	0.000	0.078	0.093	0.089	0.211
			2-I	0.235	0.219	0.174	0.394	0.280	0.387	0.357	0.238
			3-I	0.028	0.031	0.025	0.010	0.313	0.309	0.281	0.338
			4-I	0.105	0.102	0.094	0.626	0.935	0.906	0.895	0.942
			5-I	0.084	0.044	0.083	0.014	0.413	0.517	0.488	0.721
			6-I	0.017	0.016	0.015	0.107	0.335	0.426	0.389	0.570
25	50	50	0-I	0.008	0.015	0.008	0.008	0.009	0.008	0.008	0.010
			1-I	0.039	0.044	0.035	0.006	0.102	0.134	0.108	0.145
			2-I	0.076	0.066	0.035	0.420	0.300	0.403	0.412	0.173
			3-I	0.008	0.016	0.008	0.137	0.430	0.410	0.394	0.263
			4-I	0.023	0.023	0.012	0.666	0.988	0.963	0.959	0.916
			5-I	0.242	0.204	0.204	0.012	0.529	0.629	0.599	0.676
			6-I	0.128	0.110	0.078	0.074	0.722	0.818	0.810	0.373
25	50	75	0-I	0.008	0.022	0.008	0.008	0.006	0.006	0.005	0.007
			1-I	0.069	0.123	0.056	0.004	0.104	0.113	0.087	0.104
			2-I	0.013	0.013	0.003	0.436	0.310	0.391	0.393	0.127
			3-I	0.010	0.034	0.012	0.133	0.478	0.381	0.342	0.192
			4-I	0.007	0.020	0.007	0.692	0.994	0.970	0.959	0.871
			5-I	0.362	0.447	0.292	0.009	0.582	0.643	0.592	0.631
			6-I	0.024	0.022	0.008	0.783	0.766	0.839	0.829	0.264
50	50	50	0-I	0.013	0.021	0.011	0.001	0.014	0.009	0.009	0.009
			1-I	0.024	0.027	0.022	0.000	0.123	0.189	0.180	0.224
			2-I	0.100	0.120	0.076	0.385	0.267	0.388	0.392	0.217
			3-I	0.009	0.016	0.012	0.117	0.411	0.441	0.416	0.350
			4-I	0.026	0.038	0.024	0.625	0.976	0.982	0.968	0.962
			5-I	0.129	0.123	0.138	0.015	0.528	0.694	0.676	0.781
			6-I	0.161	0.180	0.124	0.711	0.719	0.833	0.826	0.508
50	100	100	0-I	0.012	0.020	0.012	0.000	0.013	0.012	0.011	0.012
			1-I	0.053	0.076	0.078	0.003	0.284	0.383	0.396	0.185
			2-I	0.006	0.020	0.003	0.434	0.326	0.426	0.412	0.129
			3-I	0.010	0.022	0.016	0.139	0.605	0.610	0.575	0.248
			4-I	0.012	0.019	0.012	0.681	0.998	0.998	0.996	0.937
			5-I	0.242	0.243	0.210	0.009	0.650	0.788	0.776	0.699
			6-I	0.040	0.042	0.023	0.747	0.844	0.898	0.892	0.404
50	100	150	0-I	0.010	0.035	0.012	0.008	0.011	0.011	0.012	0.010
			1-I	0.127	0.206	0.101	0.008	0.179	0.258	0.214	0.100
			2-I	0.030	0.032	0.013	0.406	0.304	0.402	0.409	0.154
			3-I	0.018	0.056	0.019	0.131	0.640	0.588	0.543	0.168
			4-I	0.011	0.033	0.009	0.699	0.999	0.997	0.998	0.885
			5-I	0.372	0.468	0.314	0.008	0.701	0.807	0.780	0.623
			6-I	0.011	0.021	0.007	0.770	0.870	0.917	0.910	0.278

densidad f_3 se genera desde una variable aleatoria Z cuya distribución se hace variar entre los siguientes modelos:

MD 0-II: $Z \equiv \chi_3^2$ (Hipótesis nula)

MD 1-II: $Z \equiv (1 - a)\chi_3^2 + aN(3, 1)$

MD 2-II: $Z \equiv (1 - a)\chi_3^2 + a\chi_5^2$

MD 3-II: $Z \equiv (1 - a)\chi_3^2 + a/2N(3, 1) + a/2\chi_5^2$

MD 4-II: $Z \equiv (1 - a)\chi_3^2 + a\chi_7^2$

MD 5-II: $Z \equiv a/2N(2, 1/4) + (1 - a)\chi_3^2 + a/2N(3, 1/2)$

MD 6-II: $Z \equiv (1 - a)\chi_3^2 + aLN(2, 1)$

donde nuevamente $n = n_1 + n_2 + n_3$ y $a = 6n^{-1/2}$.¹

En las figuras 1 y 2 se pueden observar las diferencias entre las densidades involucradas en los distintos modelos tipo I y tipo II, respectivamente.

Los resultados obtenidos para $\alpha = 0.05$ se muestran en la tabla 3; para $\alpha = 0.01$, en la tabla 4.

A pesar de los conocidos problemas que existen para estimar este tipo de densidades, algunas propuestas que mejoran los resultados de la estimación núcleo para estimar densidades con soporte distinto de \mathbb{R} y marcadamente asimétricas son dadas, por ejemplo, en Wand & Jones (1995). El test basado en el estadístico LG_k obtiene resultados altamente “competitivos”.

Al igual que en los modelos de tipo I, no existen problemas con el tamaño del test (modelo MD 0-II). En todos los modelos, a excepción del modelo MD 5-II (muy variable al cambiar el parámetro a), nuevamente se detecta pérdida de potencia en el test LG_k al aumentar los tamaños muestrales y cuando estos son distintos. Destaca sin embargo, a diferencia de lo observado previamente, que este fenómeno también ocurre en el resto de los tests considerados. Así las cosas, el estadístico LG_k resulta ser el más potente para los modelos MD 1-II, MD 5-II y MD 6-II cuando los tamaños muestrales de las tres muestras son iguales a 25 y, a pesar de la pérdida de potencia observada, sigue siendo el más potente para los modelos MD 1-II y MD 5-II en el caso más desfavorable, esto es, para muestras de tamaños 50, 100, 150. De nuevo, los resultados obtenidos por este test son los mejores cuando las densidades tienen formas distintas, mientras que cuando las diferencias se deben principalmente a un parámetro de posición (modelos MD 2-II y MD 4-II), el test de Kruskal-Wallis (K_W) es el que mejor resultados obtiene. Otro detalle por destacar es que, a diferencia de lo ocurrido en los modelos de tipo I (con simetría), las potencias obtenidas por las versiones de los tests clásicos estudiadas son muy competitivas en casi todos los modelos, si bien, en ocasiones, sus resultados dependen fuertemente de los tamaños muestrales.

¹La expresión LN representa la Distribución Logaritmo Normal de parámetros dos y uno.

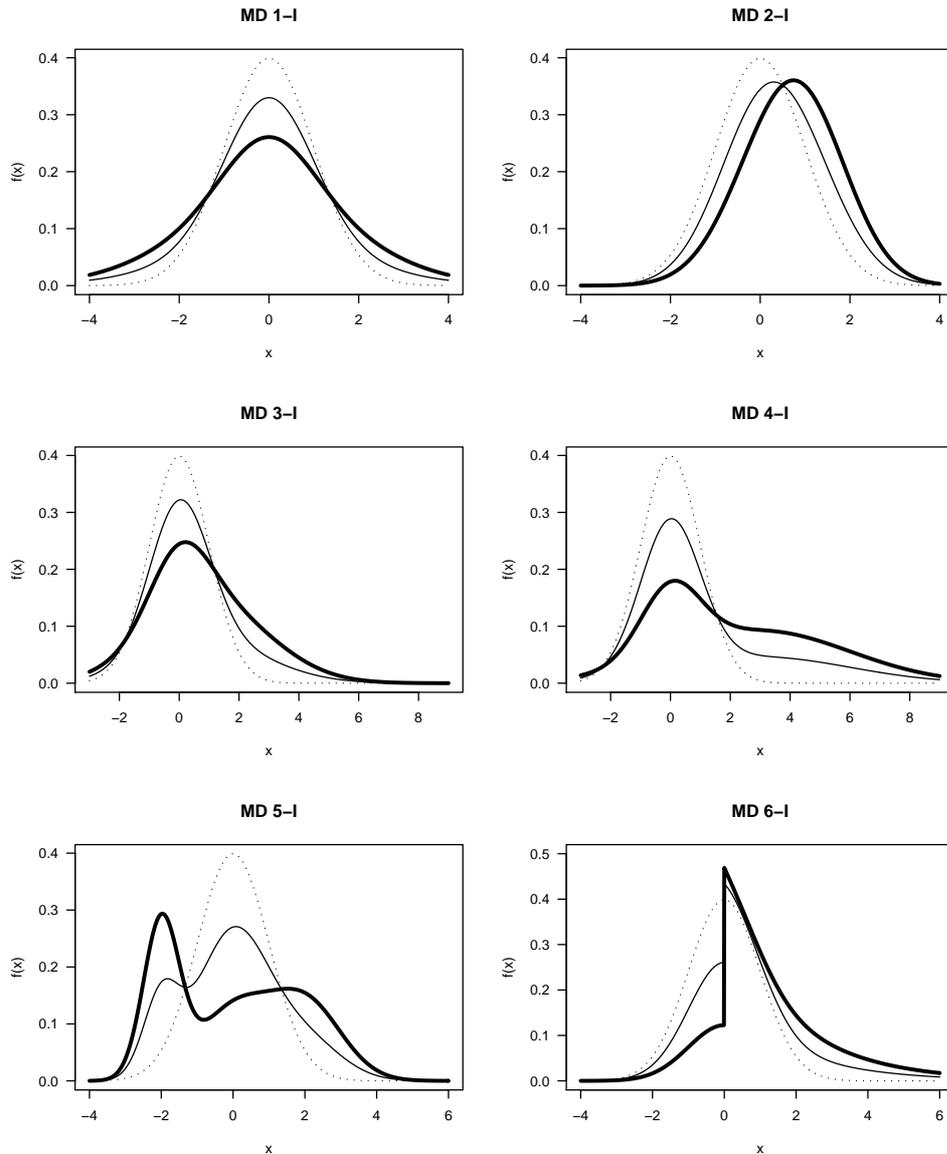


FIGURA 1: Representación gráfica de los modelos de tipo I (simetría) para $n = 300$ (línea fina) y $n = 75$ (línea gruesa). En línea punteada la hipótesis nula.

4.3. Efecto de la ventana

En el contexto de las *estimaciones suaves*, el parámetro ventana representa un papel muy importante y, por tanto, debe tratarse con mucho cuidado la elección del mismo. En todas las simulaciones descritas en las tablas anteriores, la ventana utilizada es de la forma $h_i = \hat{\sigma}_i n_i^{-1/5}$ que, asintóticamente, coincide con la ventana

óptima en el sentido del ECMI (error cuadrático medio integrado) cuando se supone que la distribución de origen es gaussiana (Silverman 1986). Con este criterio, el test LG_k no hace ningún tipo de ponderación sobre los tamaños muestrales; en consecuencia, la potencia decrece mucho cuando estos son diferentes (empíricamente hemos comprobado que al fijar una hipótesis alternativa, la potencia observada en muestras de tamaños 25, 50 y 75, respectivamente, es igual a la de

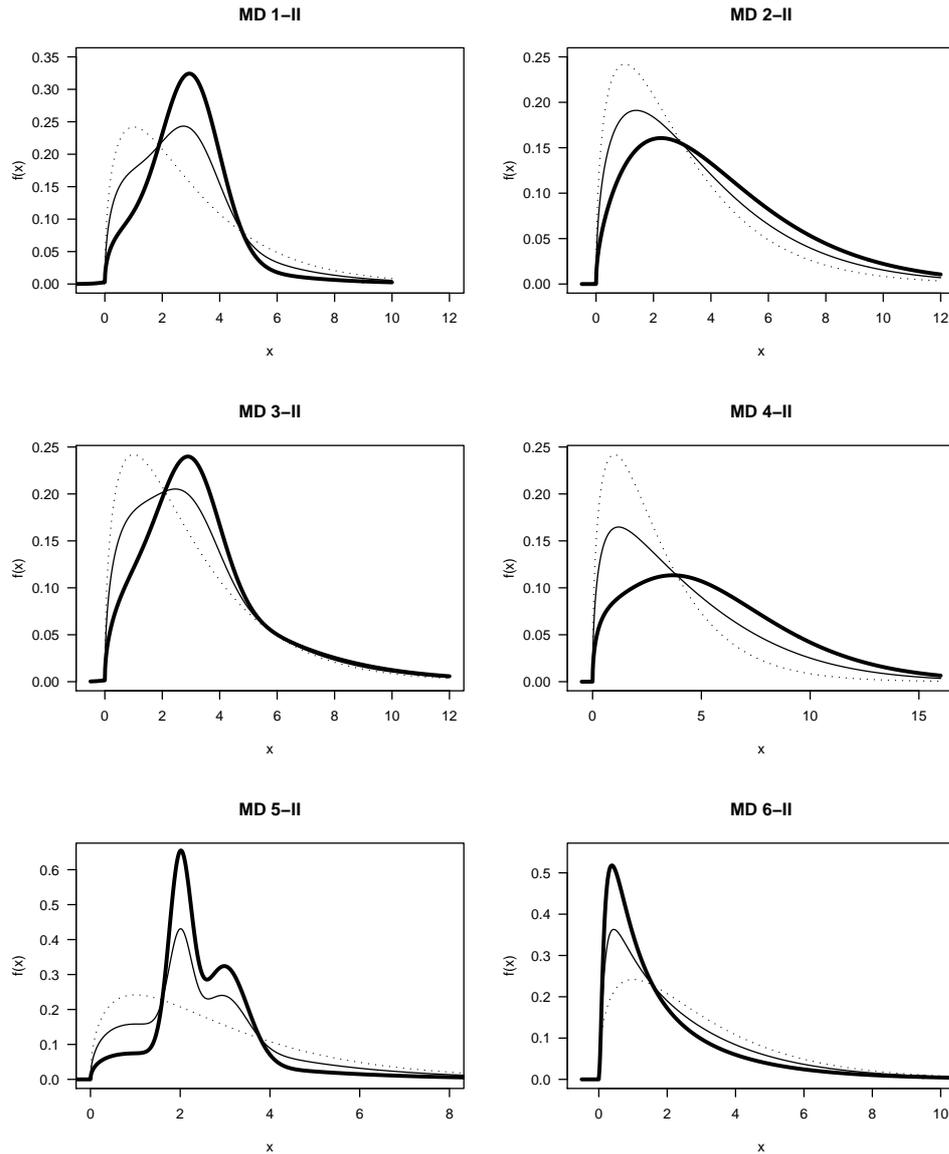


FIGURA 2: Representación gráfica de los modelos de tipo II (asimetría) para $n = 300$ (línea fina) y $n = 75$ (línea gruesa). En línea punteada la hipótesis nula.

TABLA 3: Proporción de rechazos observados en 2500 réplicas para $\alpha = 0.05$.

n_1	n_2	n_3	MD	K_S	A_D	C_M	K_W	Z_A	Z_K	Z_C	LG_k
25	25	25	0-II	0.062	0.054	0.050	0.053	0.050	0.052	0.052	0.052
			1-II	0.409	0.328	0.329	0.094	0.270	0.257	0.256	0.444
			2-II	0.355	0.292	0.291	0.453	0.336	0.404	0.413	0.340
			3-II	0.372	0.313	0.315	0.223	0.220	0.222	0.228	0.166
			4-II	0.545	0.401	0.482	0.868	0.780	0.850	0.861	0.824
			5-II	0.495	0.390	0.379	0.274	0.576	0.618	0.608	0.662
			6-II	0.454	0.355	0.374	0.822	0.854	0.905	0.908	0.945
25	50	50	0-II	0.048	0.044	0.050	0.049	0.051	0.053	0.053	0.051
			1-II	0.206	0.167	0.151	0.114	0.264	0.269	0.279	0.368
			2-II	0.179	0.135	0.124	0.458	0.342	0.401	0.404	0.281
			3-II	0.083	0.071	0.066	0.199	0.143	0.143	0.152	0.131
			4-II	0.274	0.204	0.198	0.905	0.834	0.875	0.887	0.816
			5-II	0.283	0.216	0.209	0.371	0.928	0.876	0.861	0.955
			6-II	0.230	0.178	0.162	0.838	0.918	0.845	0.938	0.941
25	50	75	0-II	0.049	0.050	0.048	0.050	0.053	0.054	0.054	0.052
			1-II	0.076	0.043	0.045	0.120	0.269	0.273	0.296	0.309
			2-II	0.065	0.037	0.037	0.485	0.359	0.416	0.423	0.233
			3-II	0.064	0.040	0.038	0.267	0.220	0.219	0.228	0.130
			4-II	0.090	0.044	0.044	0.908	0.848	0.889	0.889	0.744
			5-II	0.091	0.048	0.050	0.426	0.983	0.929	0.902	0.993
			6-II	0.077	0.038	0.039	0.876	0.943	0.961	0.956	0.882
50	50	50	0-II	0.060	0.054	0.054	0.051	0.058	0.053	0.052	0.058
			1-II	0.236	0.203	0.205	0.103	0.225	0.218	0.227	0.382
			2-II	0.217	0.184	0.190	0.446	0.314	0.378	0.390	0.358
			3-II	0.201	0.153	0.146	0.259	0.218	0.227	0.238	0.151
			4-II	0.268	0.221	0.238	0.842	0.776	0.840	0.851	0.837
			5-II	0.280	0.233	0.250	0.362	0.925	0.905	0.885	0.979
			6-II	0.243	0.206	0.211	0.811	0.924	0.956	0.951	0.977
50	100	100	0-II	0.051	0.046	0.047	0.044	0.047	0.048	0.046	0.049
			1-II	0.112	0.093	0.089	0.125	0.262	0.233	0.244	0.356
			2-II	0.115	0.092	0.084	0.482	0.348	0.418	0.420	0.303
			3-II	0.118	0.094	0.086	0.243	0.202	0.198	0.202	0.164
			4-II	0.127	0.098	0.088	0.892	0.851	0.893	0.893	0.815
			5-II	0.128	0.102	0.098	0.410	0.997	0.984	0.984	0.999
			6-II	0.120	0.098	0.093	0.841	0.968	0.977	0.974	0.951
50	100	150	0-II	0.049	0.043	0.042	0.047	0.046	0.046	0.045	0.047
			1-II	0.050	0.034	0.033	0.117	0.250	0.240	0.254	0.320
			2-II	0.050	0.036	0.035	0.506	0.374	0.423	0.428	0.243
			3-II	0.050	0.039	0.036	0.267	0.206	0.213	0.218	0.140
			4-II	0.056	0.041	0.039	0.900	0.855	0.892	0.898	0.728
			5-II	0.052	0.032	0.036	0.449	1.000	0.995	0.993	1.000
			6-II	0.052	0.032	0.032	0.834	0.980	0.986	0.983	0.866

muestras de tamaños 25, 25 y 25, esto es, no se utiliza la información contenida en la mitad de la muestra).

Tratando de corregir este problema, se han realizado nuevas pruebas de simulación sobre los modelos ya propuestos, en las cuales la ventana viene determinada por $h_i = S\hat{\sigma}_i m^{-1/5}$, siendo m la media armónica de los tamaños muestrales y S es un factor de escala que se hace oscilar entre 0.5 y 12 (para muestras iguales y $S = 1$ el procedimiento coincide con el de apartados anteriores). En la toma de

TABLA 4: Proporción de rechazos observados en 2500 réplicas para $\alpha = 0.01$.

n_1	n_2	n_3	MD	K_S	A_D	C_M	K_W	Z_A	Z_K	Z_C	LG_k
25	25	25	0-II	0.012	0.015	0.012	0.011	0.013	0.011	0.012	0.011
			1-II	0.202	0.196	0.133	0.026	0.088	0.078	0.077	0.190
			2-II	0.170	0.169	0.123	0.221	0.141	0.194	0.197	0.144
			3-II	0.178	0.186	0.120	0.078	0.081	0.076	0.076	0.050
			4-II	0.352	0.282	0.274	0.688	0.527	0.647	0.658	0.568
			5-II	0.251	0.246	0.138	0.105	0.296	0.326	0.300	0.360
			6-II	0.261	0.228	0.196	0.625	0.620	0.729	0.724	0.793
25	50	50	0-II	0.011	0.013	0.010	0.012	0.023	0.010	0.010	0.007
			1-II	0.083	0.077	0.044	0.033	0.091	0.081	0.082	0.139
			2-II	0.060	0.054	0.035	0.232	0.242	0.196	0.204	0.098
			3-II	0.022	0.024	0.010	0.066	0.045	0.040	0.048	0.036
			4-II	0.121	0.105	0.068	0.744	0.620	0.706	0.711	0.554
			5-II	0.116	0.106	0.058	0.170	0.738	0.591	0.566	0.802
			6-II	0.097	0.086	0.047	0.650	0.746	0.810	0.792	0.706
25	50	75	0-II	0.009	0.027	0.005	0.014	0.012	0.011	0.012	0.016
			1-II	0.017	0.015	0.007	0.031	0.088	0.095	0.100	0.088
			2-II	0.011	0.012	0.006	0.251	0.151	0.198	0.204	0.061
			3-II	0.013	0.016	0.006	0.103	0.081	0.077	0.077	0.031
			4-II	0.018	0.016	0.008	0.750	0.634	0.709	0.697	0.440
			5-II	0.024	0.018	0.006	0.200	0.888	0.680	0.650	0.938
			6-II	0.014	0.016	0.003	0.694	0.798	0.831	0.820	0.574
50	50	50	0-II	0.011	0.018	0.013	0.010	0.010	0.011	0.014	0.014
			1-II	0.115	0.134	0.088	0.026	0.077	0.062	0.068	0.171
			2-II	0.096	0.122	0.068	0.211	0.132	0.175	0.188	0.152
			3-II	0.074	0.069	0.036	0.091	0.082	0.082	0.086	0.037
			4-II	0.135	0.147	0.106	0.664	0.548	0.647	0.666	0.639
			5-II	0.144	0.162	0.107	0.163	0.784	0.676	0.656	0.908
			6-II	0.112	0.135	0.081	0.622	0.757	0.841	0.829	0.873
50	100	100	0-II	0.010	0.016	0.009	0.008	0.010	0.009	0.010	0.009
			1-II	0.040	0.036	0.016	0.033	0.088	0.082	0.081	0.128
			2-II	0.030	0.040	0.019	0.251	0.144	0.207	0.207	0.113
			3-II	0.033	0.040	0.020	0.090	0.073	0.068	0.069	0.037
			4-II	0.042	0.043	0.024	0.724	0.618	0.717	0.729	0.555
			5-II	0.041	0.044	0.020	0.206	0.981	0.887	0.857	0.994
			6-II	0.042	0.039	0.018	0.668	0.872	0.898	0.897	0.718
50	100	150	0-II	0.008	0.017	0.011	0.009	0.009	0.009	0.008	0.009
			1-II	0.007	0.017	0.003	0.035	0.090	0.082	0.084	0.101
			2-II	0.010	0.020	0.006	0.265	0.164	0.211	0.212	0.080
			3-II	0.008	0.022	0.005	0.110	0.080	0.076	0.084	0.031
			4-II	0.011	0.024	0.006	0.730	0.664	0.721	0.728	0.440
			5-II	0.011	0.017	0.009	0.224	0.998	0.929	0.902	0.999
			6-II	0.007	0.017	0.003	0.682	0.898	0.923	0.905	0.500

decisiones se propone, para cada problema, calcular el valor medio de LG_k en los distintos h_n considerados (en este caso 6 valores distintos).

Las potencias obtenidas cuando se utiliza este criterio, denotado por MdH , se muestran en las tablas 5 y 6. El proceso seguido es análogo al utilizado anteriormente, si bien el número de simulaciones para cada problema fue 1000.

TABLA 5: Proporción de rechazos observados en 1000 réplicas para $\alpha = 0.05$.

			S							
n_1	n_2	n_3	MD	1/2	1	3	6	9	12	MdH
25	25	25	0-I	0.051	0.036	0.044	0.049	0.046	0.046	0.046
			1-I	0.277	0.428	0.593	0.597	0.593	0.593	0.579
			2-I	0.337	0.462	0.284	0.118	0.091	0.082	0.226
			3-I	0.395	0.627	0.703	0.665	0.649	0.644	0.651
			4-I	0.926	0.986	0.995	0.992	0.984	0.983	0.992
			5-I	0.844	0.910	0.939	0.946	0.944	0.943	0.954
			6-I	0.612	0.738	0.548	0.389	0.348	0.329	0.594
25	50	75	0-I	0.059	0.060	0.053	0.053	0.056	0.057	0.059
			1-I	0.207	0.384	0.629	0.633	0.627	0.624	0.604
			2-I	0.279	0.376	0.279	0.154	0.127	0.123	0.227
			3-I	0.325	0.545	0.758	0.739	0.723	0.717	0.724
			4-I	0.910	0.988	0.999	0.999	0.998	0.998	0.997
			5-I	0.815	0.910	0.964	0.969	0.971	0.971	0.970
			6-I	0.501	0.663	0.604	0.506	0.480	0.472	0.628
50	50	50	0-I	0.046	0.043	0.048	0.053	0.049	0.049	0.043
			1-I	0.258	0.454	0.709	0.728	0.723	0.724	0.696
			2-I	0.314	0.438	0.380	0.174	0.134	0.123	0.305
			3-I	0.402	0.628	0.853	0.835	0.821	0.817	0.823
			4-I	0.911	0.994	1.000	1.000	1.000	1.000	1.000
			5-I	0.823	0.932	0.972	0.975	0.975	0.975	0.978
			6-I	0.587	0.743	0.769	0.684	0.650	0.638	0.788
50	100	150	0-I	0.063	0.063	0.048	0.040	0.045	0.045	0.050
			1-I	0.214	0.382	0.751	0.763	0.755	0.755	0.714
			2-I	0.245	0.353	0.336	0.203	0.164	0.146	0.276
			3-I	0.288	0.560	0.866	0.867	0.859	0.851	0.831
			4-I	0.907	0.997	1.000	1.000	1.000	1.000	1.000
			5-I	0.759	0.904	0.984	0.987	0.989	0.989	0.982
			6-I	0.480	0.683	0.792	0.732	0.699	0.682	0.804

En la tabla 5 se puede observar que en los modelos MD 1-I, MD 3-I, MD 4-I y MD 5-I, la potencia aumenta al aumentar S , mientras que en los modelos MD 2-I y MD 6-I, las potencias más elevadas se obtienen para $S = 1$. Destaca también que las potencias para tamaños de muestra diferentes, a pesar de seguir siendo inferiores que cuando estos son iguales, aumenta sensiblemente respecto a los observados en la tabla 1. Se tiene, por tanto, que con una elección adecuada de la ventana, el test LG_k obtiene las mejores potencias en el escenario más desfavorable (muestras de tamaños 50, 100, 150), en los modelos MD 1-I, MD 3-I, MD 4-I y MD 5-I.

Los resultados que se obtienen en los modelos de tipo II (asimetría) son una prueba de la complejidad que conlleva la búsqueda de un parámetro ventana óptimo o, en su defecto, razonablemente bueno. En estos modelos, los mejores resultados se obtienen para S no muy elevado. De nuevo se pone de manifiesto su importancia, ya que la potencia desciende considerablemente cuando el parámetro ventana es demasiado grande.

El criterio de toma de decisiones propuesto “suaviza” de alguna manera los resultados obtenidos en la rejilla h 's elegida, mejora los resultados descritos en la tabla 1 para todos los modelos, salvo para los modelos MD 2-I y MD 6-I, pero

TABLA 6: Proporción de rechazos observados en 1000 réplicas para $\alpha = 0.05$.

			S							
n_1	n_2	n_3	MD	1/2	1	3	6	9	12	MdH
25	25	25	0-II	0.029	0.045	0.048	0.047	0.046	0.047	0.049
			1-II	0.397	0.407	0.370	0.374	0.373	0.371	0.426
			2-II	0.294	0.303	0.180	0.139	0.135	0.134	0.208
			3-II	0.191	0.135	0.059	0.056	0.053	0.053	0.077
			4-II	0.668	0.787	0.653	0.549	0.524	0.411	0.698
			5-II	0.370	0.650	0.690	0.662	0.658	0.656	0.661
			6-II	0.899	0.946	0.895	0.881	0.887	0.887	0.918
25	50	75	0-II	0.055	0.061	0.062	0.063	0.063	0.062	0.061
			1-II	0.275	0.268	0.181	0.174	0.174	0.176	0.227
			2-II	0.229	0.273	0.195	0.158	0.144	0.141	0.200
			3-II	0.160	0.130	0.055	0.043	0.044	0.044	0.063
			4-II	0.666	0.836	0.761	0.656	0.627	0.615	0.778
			5-II	0.996	0.998	0.997	0.997	0.997	0.997	0.998
			6-II	0.925	0.953	0.918	0.898	0.896	0.895	0.939
50	50	50	0-II	0.056	0.059	0.063	0.060	0.058	0.060	0.054
			1-II	0.388	0.398	0.288	0.266	0.272	0.272	0.345
			2-II	0.214	0.271	0.172	0.140	0.132	0.128	0.193
			3-II	0.218	0.183	0.074	0.058	0.058	0.058	0.099
			4-II	0.728	0.836	0.705	0.578	0.541	0.531	0.747
			5-II	0.941	0.982	0.987	0.987	0.987	0.987	0.986
			6-II	0.945	0.979	0.947	0.935	0.932	0.931	0.968
50	100	150	0-II	0.057	0.055	0.061	0.063	0.063	0.061	0.056
			1-II	0.278	0.260	0.140	0.115	0.112	0.100	0.177
			2-II	0.244	0.316	0.215	0.165	0.155	0.152	0.231
			3-II	0.188	0.154	0.069	0.057	0.051	0.050	0.090
			4-II	0.661	0.800	0.666	0.549	0.522	0.506	0.703
			5-II	1.000	1.000	1.000	1.000	1.000	1.000	1.000
			6-II	0.958	0.961	0.936	0.930	0.929	0.927	0.962

empeora los observados en la tabla 3. Así este criterio, si bien da una posible solución objetiva en la toma de decisiones, dista mucho de solucionar el problema real de forma óptima.

5. Conclusiones

En este trabajo se ha presentado y estudiado una medida que permite comparar la igualdad entre k poblaciones a partir de sus funciones de densidad y que, en cierto sentido, generaliza la medida L_1 . Se proponen dos planes de remuestreo, ambos basados en lanzar muestras desde la Función de Distribución Empírica Suavizada, aunque realizando esta estimación por procedimientos diferentes que, a pesar de ser muy similares, piden diferentes hipótesis a los datos. Se realiza un amplio estudio de simulación comparando el test basado en el estadístico propuesto con siete tests basados en la Función de Distribución Empírica. Se comprueba que, en general, el test en estudio es más potente cuando los tamaños muestrales son

iguales y pequeños, y las distribuciones difieren en la forma y no solamente en un parámetro de posición.

En el primer tipo de poblaciones estudiadas (simetría), y con el criterio propuesto en la subsección 4.3, se tiene que el test basado en el estadístico LG_k obtiene mejores resultados para muestras iguales y pequeñas, destacando la pérdida de potencia al aumentar el tamaño muestral.

Para el segundo tipo de poblaciones (asimetría), los resultados cambian, aunque a pesar de los problemas de la estimación núcleo en este tipo de distribuciones, los resultados obtenidos son muy positivos ya que la potencia obtenida es la mejor en los modelos MD 1-II y MD 5-II, siendo muy competitivos para el resto de los modelos, sobre todo para muestras pequeñas e iguales.

En la sección 4.3 se proponen algunos cambios que mejoran considerablemente los resultados y se pone de manifiesto la relevancia de la elección del parámetro ventana. Una elección adecuada del mismo puede hacer que la potencia del test crezca de forma interesante. Así mismo, de las tablas 5 y 6 se deduce que este problema no es sencillo ya que, por un lado, depende del tamaño de muestra (presumiblemente el orden de convergencia óptimo no debe coincidir con el utilizado para la estimación de la densidad) y, por otro, las distribuciones de origen representan un papel crucial. A pesar de estos inconvenientes, parece probado que, de los tests considerados, el basado en el estadístico LG_k , obtiene los mejores resultados cuando las distribuciones se diferencian únicamente en la forma y tienen la misma posición, sobre todo cuando se considera la versión propuesta en la subsección 4.3.

Finalmente, se propone un criterio de toma de decisiones que, lejos de solucionar el problema de la elección del parámetro ventana, ofrece una solución suavizada entre un conjunto de posibles resultados.

[Recibido: junio de 2007 — Aceptado: agosto de 2007]

Referencias

- Anderson, N. H., Hall, P. & Titterton, D. M. (1994), 'Two-Sample Test Statistics for Measuring Discrepancies Between Two Multivariate Probability Density Functions using Kernel-Based Density Estimates', *Journal of Multivariate Analysis* **50**, 41–54.
- Cao, R. & Van Keilegom, I. (2006), 'Empirical Likelihood Tests for Two-Sample Problems via Nonparametric Density Estimation', *Canad. J. Statist.* **34**, 61–77.
- Conover, W. J. (1965), 'Several k -sample Kolmogorov-Smirnov tests', *Annals of Math. Statistics* **36**, 1019–1026.
- Devroye, L. & Györfi, L. (1985), *Nonparametric Density Estimation. The L_1 -View*, Wiley, New York, United States.
- Hall, P., DiCiccio, J. T. & Romano, J. P. (1989), 'On Smoothing and the Bootstrap', *Annals of Statistics* **17**(2), 692–704.

- Horvath, L. (1991), 'On L_p -Norms of Multivariate Density Estimations', *Annals of Statistics* **19**(4), 1933–1949.
- Kiefer, J. (1959), ' k -Sample Analogues of the Kolmogorov-Smirnov, Cramér-Von Mises Test', *Ann. Math. Statist.* **30**, 420–447.
- Kruskal, W. H. & Wallis, W. A. (1952), 'Use of Ranks in One-Criterion Variance Analysis', *Journal of the American Statistical Association* **47**(260), 583–621.
- Lewis, J. L. (1972), 'A k -Sample Test Based on Range Intervals', *Biometrika* **59**(1), 155–160.
- Nadaraya, E. A. (1964), 'Some new Estimates for Distribution Functions', *Theory Prob. Appl.* **9**, 497–500.
- R Development Core Team (2007), *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.
*<http://www.R-project.org>
- Rosenblatt, M. (1956), 'Remarks on Some nonparametric Estimates of a Density Functions', *Annals Math. Statistics* **27**, 832–837.
- Scholz, F. W. & Stephens, M. A. (1987), ' k -Samples Anderson-Darling Test', *J. Amer. Statist. Assoc.* **82**, 918–924.
- Silverman, B. W. (1986), *Density Estimation for Statistics and Data Analysis*, Chapman & Hall, London, United Kingdom.
- Wand, M. P. & Jones, M. C. (1995), *Kernel Smoothing*, Chapman & Hall, London, United Kingdom.
- Zhang, J. & Wu, Y. (2007), ' k -Sample Tests Based on the Likelihood Ratio', *Comput. Stat. Data Anal.* **51**(9), 4682–4691.