Revista Colombiana de Estadística
Diciembre 2009, volumen 32, no. 2, pp. 289 a 299

Cramér-Chernoff Theorem for L_1-norm in Kernel Density Estimator for Two Independent Samples

Teorema de Cramér-Chernoff para la norma L_1 del estimador núcleo para dos muestras independientes

Pablo Martínez-Camblor[1,a], Norberto Corral[2,b], Teresa López[2,c]

[1]CIBER de Epidemiología y Salud Pública (CIBERESP), Subdirección Salud Pública de Gipuzkoa, Donostia, Spain

[2]Estadística e Investigación Operativa y Didáctica de la Matemática, Universidad de Oviedo, Asturias, Spain

Resumen

In this paper a Chernoff type theorem for the L_1 distance between kernel estimators from two independent and identically distributed random samples is developed. The harmonic mean is used to correct the distance for inequal sample sizes case. Moreover, the proved result is used to compute the Bahadur slope of a test based on L_1 distance and to compare it with the classical nonparametric Mann-Whitney test by using the Bahadur relative efficiency.

***Palabras clave*:** Kernel estimator, Large deviation, Bahadur slope.

Abstract

En este trabajo se desarrolla un teorema de tipo Chernoff para la distancia L_1 entre estimadores núcleo procedentes de muestras aleatorias independientes e idénticamente distribuidas. Se usa la media armónica para corregir esta distancia en el caso de muestras de distintos tamaños. Además, se usa el resultado demostrado para el cálculo de la pendiente de Bahadur de un test para la comparación de densidades basado en la distancia L_1 y se compara con el clásico test de Mann-Whitney a partir de la eficiencia relativa de Bahadur.

***Key words*:** estimador núcleo, grandes muestras, pendiente de Bahadur.

[a]Investigador postdosctoral. E-mail: pmcamblor@hotmail.com
[b]Catedrático. E-mail: norbert@uniovi.es
[c]Profesora titular. E-mail: teresa@uniovi.es

1. Introduction

Let $X = \{x_1, \ldots, x_n\}$ be an independent random sample from a random variable X which is absolutely continuous with probability density function f. The kernel density estimator (KDE) of f introduced by Rosenblatt (1956) and Parzen (1962) is defined for each $t \in \mathbb{R}$ as

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

where K is a kernel function which is often chosen to be a continuous symmetric density with finite variance, and $\{h_n\}_{n\in\mathbb{N}}$ is a sequence of positive real numbers.

The kernel estimator and its properties have been widely studied by several authors. Silverman (1978) proved its uniform consistence and Konakov (1978) derived the asymptotic distribution for the L_∞-norm. Devroye & Wagner (1979) proved the L_1 convergence between the kernel density estimator and its target, and Devroye & Gyorfi (1985) carry out a widely study about the kernel density estimators from the L_1 approach. Berlinet et al. (1995) proved the asymptotic normality for the L_1-norm for the histogram density estimator, Hórvath (1991) demonstrated the asymptotic normality of the L_p norm between the kernel density estimator and the underlying density function, Martínez-Camblor & Corral (2008) proved the same result under weaker assumptions. Large deviation approaches were also considered. Louani (1998, 2000, 2005) studied Chernoff type theorems for the L_1 distance using for goodness of fit test. Cao & Lugosi (2005) studied the propierties of several goodness of fit tests based on the kernel density estimate. Osmoukhina (2001) applied these techniques in a symmetry test and Beirlant et al. (2001) studied the large deviations of divergence measures.

Let $X = \{x_1, \ldots, x_{n_1}\}$ and $Y = \{y_1, \ldots, y_{n_2}\}$ be two independent random samples from a random variable with density function f. The L_1 distance between two estimators f_{n_1} and f_{n_2} of f is defined by

$$D(f_{n_1}, f_{n_2}) = \|f_{n_1} - f_{n_2}\|_{L_1} = \int |f_{n_1}(X,t) - f_{n_2}(Y,t)|\, dt$$

The main objective of this paper is to obtain a Chernoff type theorem for the L_1 distance, $D(f_{n_1}, f_{n_2})$, between two kernel density estimators when both are computed from samples which are taken from the same population. That is, our purpose is to investigate the expression

$$\mathcal{P}\left\{\int |f_{n_1}(X,t) - f_{n_2}(Y,t)|\, dt > \lambda\right\} \tag{1}$$

where λ is close to zero.

We are interested in this result because it can be applied in two sample problems and let us to compare different tests by using the Bahadur relative slope (BRS). In Section 2 and, in a similar way than Louani (2000), we prove the large deviation theorem. In Section 3 we consider several alternative hypothesis and, from the previous results, the kernel density estimator based test with the classical nonparametric one of Mann-Whitney are compared in the Bahadur relative efficiency (BRE) sense.

2. Results and Proofs

In order to prove the main result, we will consider the following assumptions:

($C1$) $X = \{x_1, \dots, x_{n_1}\}$ and $Y = \{y_1, \dots, y_{n_2}\}$ are independent random samples from a continuous random variable.

($C2$) $\text{lím}_n\, n_1/n_2 = 1$, where $n \to \infty$ means that $n_1 \to \infty$ and $n_2 \to \infty$.

($C3$) The used kernel function, K, is a continuous, symmetric density function.

($C4$) $\text{lím}_{n_1} h_{n_1} = 0$, $\text{lím}_{n_2} h_{n_2} = 0$, $\text{lím}_{n_1} n_1 h_{n_1} = \infty$ and $\text{lím}_{n_2} n_2 h_{n_2} = \infty$

Nota 1. In practice, the used bandwith, h_n, is chosen to minimize certain error criterion (for example, the mean integrated squared error; MISE). Since the kernel function is a symmetric and differentiable density function having finite variance, these conditions are satisfied for most commonly used kernel functions like Gaussian, Epanechnikov, Triangular, among others. As consequence, the assumptions ($C3$) and ($C4$) are mild conditions.

Lema 1. *Let $X = \{x_1, \dots, x_n\}$ and $Y = \{y_1, \dots, y_n\}$ be two independent random samples from a random variable with density function f. For each interval $B \in \mathcal{B}$ (Borel σ-field) we define*

$$Z_{B,n,i} = \int_B \left(\frac{1}{h_{n_1}} K\left(\frac{x_i - t}{h_{n_1}}\right) - \frac{1}{h_{n_2}} K\left(\frac{y_i - t}{h_{n_2}}\right)\right) dt \quad \text{with } 1 \le i \le n$$

Then, under conditions ($C1$) ($C2$) and ($C4$), we have that

(1)
$$\text{lím}_n Z_{B,n,i} = \begin{cases} 1 & \text{if } (x_i, y_i) \in (\mathring{B} \times \overline{B}^c) \\ 1/2 & \text{if } (x_i, y_i) \in (\partial B \times \overline{B}^c) \bigcup (\mathring{B} \times \partial B) \\ -1/2 & \text{if } (x_i, y_i) \in (\overline{B}^c \times \partial B) \bigcup (\partial B \times \mathring{B}) \\ -1 & \text{if } (x_i, y_i) \in (\overline{B}^c \times \mathring{B}) \\ 0 & \text{otherwise} \end{cases}$$

where $\mathring{B}$ and ∂B denote the interior and the boundary of B, respectively, and $\overline{B} = B \cup \partial B$.

(2) *For all $\lambda > 0$, $\quad lim_n \frac{1}{n} \log \mathcal{P}\left\{\frac{1}{n}\sum_{i=1}^n Z_{B,n,i} > \lambda\right\} = inf_{\{t>0\}}\left\{-\lambda t + q_a(t)\right\}$ where $a = \int_B f(t)dt$ and $q_a(t) = \log(a^2 + (1-a)^2 + 2a(1-a)\cosh(t))$*

Demostración. For $1 \leq i \leq n$, we have that

$$\lim_n \int_B \left(\frac{1}{h_{n_1}} K\left(\frac{x_i - t}{h_{n_1}}\right) - \frac{1}{h_{n_2}} K\left(\frac{y_i - t}{h_{n_2}}\right)\right) dt = \lim_{n_1} \int_B \frac{1}{h_{n_1}} K\left(\frac{x_i - t}{h_{n_1}}\right) dt - \lim_{n_2} \int_B \frac{1}{h_{n_2}} K\left(\frac{y_i - t}{h_{n_2}}\right) dt$$

Taking $x_i = t + h_{n_1} u$ and $y_i = t + h_{n_2} v$ we have that

$$\begin{aligned} \lim_n \int_B \left(\frac{1}{h_{n_1}} K\left(\frac{x_i - t}{h_{n_1}}\right) - \frac{1}{h_{n_2}} K\left(\frac{y_i - t}{h_{n_2}}\right)\right) dt &= \\ \lim_{n_1} \int_{\mathbb{R}} I_{B_{n_1}}(u)\, K(u) du - \lim_{n_2} \int_{\mathbb{R}} I_{B_{n_2}}(v)\, K(v) dv &= \\ \int_{\mathbb{R}} [\lim_{n_1} I_{B_{n_1}}(u)]\, K(u) du - \int_{\mathbb{R}} [\lim_{n_2} I_{B_{n_2}}(v)]\, K(v) dv & \end{aligned}$$

where $B_{n_1} = \{u / x_i - u h_{n_1} \in B\}$ and $B_{n_2} = \{v / y_i - v h_{n_1} \in B\}$.

By taking into account that

$$\lim_{n_1} B_{n_1} = \begin{cases} \mathbb{R} & \text{if } x_i \in \overset{\circ}{B} \\ \mathbb{R}^+ & \text{if } x_i = \sup(B) \\ \mathbb{R}^- & \text{if } x_i = \inf(B) \\ \phi & \text{if } x_i \in \overline{B}^c \end{cases}$$

the result (1) is easily concluded.

Therefore, from the first part of this Lemma, for $1 \leq i \leq n$ there exists

$$Z_{B,i} = \lim_n Z_{B,n,i}$$

random sample from the same random variable, Z_B, whose moment generating function is given by

$$\begin{aligned} g_{Z_B}(t) &= \iint e^{tZ_B} f(x) f(y) dx dy \\ &= \int_B \int_{B^c} e^t f(x) f(y) dx dy \int_{B^c} \int_B e^{-t} f(x) f(y) dx dy \\ &\qquad + \int_B \int_B f(x) f(y) dx dy + \int_{B^c} \int_{B^c} f(x) f(y) dx dy \\ &= a^2 + (1-a)^2 + 2a(1-a)\cosh(t) \end{aligned}$$

with $a = \int_B f(t) dt$.

Applying the Cramér-Chernoff Theorem (Van der Vaart 1998) the proof is completed. □

Observación 1. This result can be immediately extended to the case in which B is a countable union of intervals.

From the above result we can derive a Chernoff type theorem. This result lets us to compare the test based on the L_1-norm of the kernel density estimator with different tests by using their respective Bahadur relative slope (BRS).

Teorema 1. *If conditions* $(C1)$, $(C2)$, $(C3)$ *and* $(C4)$ *are fulfilled and* λ *is a nonnegative constant close to zero, then*

$$\lim_n \frac{n_1+n_2}{2n_1n_2}\log \mathcal{P}\left\{\int |f_{n_1}(X,t)-f_{n_2}(Y,t)|\,dt>\lambda\right\} = -\frac{\lambda^2}{4}\,(1+o(1)) \ a.s.$$

Demostración. To prove this theorem we will assume that $n_1 = n_2 = n$. Now, we define the function,

$$Q_a(\lambda) = -\frac{\lambda}{2} \quad \operatorname{arccosh}\left(\frac{(1-2a+2a^2)\lambda^2 + 2\sqrt{\lambda^2(1-2a)^2+16a^2(a-1)^2}}{2a(a-1)(\lambda^2-4)}\right)$$
$$+\log\left((1-a)^2+a^2+\frac{(1-2a+2a^2)\lambda^2+2\sqrt{\lambda^2(1-2a)^2+16a^2(a-1)^2}}{(4-\lambda^2)}\right)$$

It is easy to check that $Q_a(\lambda) = \inf_{t>0}\left\{-\frac{\lambda}{2}t + q_a(t)\right\}$, and by using its Taylor expansion with λ in a neighborhood of zero, we have that

$$\sup_{a\in(0,1)} Q_a(\lambda) = Q_{1/2}(\lambda) = -\frac{\lambda^2}{4}\,(1+o(1))$$

On the other hand, by using the Scheffé Theorem

$$\int |f_{n_1}(X,t)-f_{n_2}(Y,t)|\,dt = 2\sup_{B\in\mathcal{B}}\left|\int_B (f_{n_1}(X,t)-f_{n_2}(Y,t))\,dt\right|$$

Moreover, for every $B \in \mathcal{B}$

$$\mathcal{P}\left\{2\int_B (f_{n_1}(X,t)-f_{n_2}(Y,t))\,dt>\lambda\right\} =$$
$$\mathcal{P}\left\{\frac{1}{n}\sum_{i=1}^{n}\int_B\left(\frac{1}{h_{n_1}}K\left(\frac{x_i-t}{h_{n_1}}\right)-\frac{1}{h_{n_2}}K\left(\frac{y_i-t}{h_{n_2}}\right)\right)dt>\frac{\lambda}{2}\right\}$$

from the properties of $Q_a(\lambda)$, the previous Lemma 1 and by taking $a = \int_B f(t)dt$, we conclude that

$$\lim_n \frac{1}{n}\log\mathcal{P}\left\{2\int_B (f_{n_1}(X,t)-f_{n_2}(Y,t))\,dt>\lambda\right\} = Q_a(\lambda)$$

Now, taken an arbitrary $B_0 \in \mathcal{B}$ such that $a = 1/2$,

$$Q_{1/2}(\lambda) = -\frac{\lambda}{2}\operatorname{arccosh}\left(\frac{4+\lambda^2}{4-\lambda^2}\right)+\log\left(\frac{1}{2}\left(1+\frac{4+\lambda^2}{4-\lambda^2}\right)\right)$$

and applying the Taylor expansion of $Q_{1/2}(\lambda)$ we get,

$$\begin{aligned}
&\liminf_n \frac{1}{n} \log\left(\mathcal{P}\left\{\int |f_n(X,t)-f_n(Y,t)|\,dt > \lambda\right\}\right)\\
&\qquad= \liminf_n \frac{1}{n} \log\left(\mathcal{P}\left\{\sup_{B\in\mathcal{B}}\left|\int (f_n(X,t)-f_n(Y,t))\,dt\right| > \frac{\lambda}{2}\right\}\right)\\
&\qquad\geq \liminf_n \frac{1}{n} \log\left(\mathcal{P}\left\{\left|\int_{B_0} (f_n(X,t)-f_n(Y,t))\,dt\right| > \frac{\lambda}{2}\right\}\right)\\
&\qquad\geq -\frac{\lambda^2}{4}\,(1+o(1))
\end{aligned} \tag{2}$$

To prove the upper bound, we know that for $\delta > 0$ and any density function K we can find a kernel L in the form

$$L = \sum_{j=1}^{N_\delta} \alpha_j I_{R_j}$$

satisfying (C3) and such that

$$\int |K - L| < \delta$$

where N_δ only depends on δ, α_j's are nonnegative finite constants and R_j's are disjoint open finite intervals.

Hence, if we define

$$L_n(Z,t) = \frac{1}{n}\sum_{i=1}^{n} \frac{1}{h_n} L\left(\frac{z_i - t}{h_n}\right)$$

we have the inequality

$$\int |f_n(X,t)-f_n(Y,t)|\,dt < 2\delta + \int |L_n(X,t)-L_n(Y,t)|\,dt$$

Following the proof of Theorem 3.1 (Devroye 1987) we obtain that for $\epsilon > 0$ and, if $\Lambda_n(X,.)$ and $\Lambda_n(Y,.)$ are the empirical probability measures associated to the samples X and Y, respectively,

$$\int |f_n(X,t)-f_n(Y,t)|\,dt < \epsilon + \sum_{B\in\Delta_{r,l}} |\Lambda_n(X,B)-\Lambda_n(Y,B)|$$

where $\Pi_{r,l}$ is a partition of $(-r,r)$, into intervals of length h_n/l for some $l > 0$, and $\Delta_{r,l} = \Pi_{r,l} \cup \{(-\infty,-r)\cup(r,\infty)\}$ is a partition of $\mathbb{R}$. As consequence

$$\mathcal{P}\left\{\int |f_n(X,t)-f_n(Y,t)|\,dt > \lambda\right\} \leq \mathcal{P}\left\{\sum_{B\in\Delta_{r,l}} |\Lambda_n(X,B)-\Lambda_n(Y,B)| > \lambda-\epsilon\right\}$$

If $\Im_{r,l}$ is the set of all posible sets given by unions of elements of the partition $\Delta_{r,l}$, using a similar argument as in the Scheffé Theorem, we have

$$\begin{aligned}\mathcal{P}\left\{\int |f_n(X,t)-f_n(Y,t)|\,dt>\lambda\right\} &\le \mathcal{P}\left\{\sup_{B\in\Im_{r,l}} |\Lambda_n(X,B)-\Lambda_n(Y,B)|>\frac{1}{2}(\lambda-\epsilon)\right\}\\ &\le \sum_{B\in\Im_{r,l}} \mathcal{P}\left\{|\Lambda_n(X,B)-\Lambda_n(Y,B)|>\frac{1}{2}(\lambda-\epsilon)\right\}\\ &\le \sum_{B\in\Im_{r,l}} \exp\{nQ_{\mathcal{P}\{B\}}(\lambda-\epsilon)\}\\ &\le Car(\Im_{r,l})\exp\left\{-\frac{n(\lambda-\epsilon)^2}{4}(1+o(1))\right\}\end{aligned}$$

As the cardinality of $\Im_{r,l}$, $Car(\Im_{r,l})$, is at most $2^{(2+2rl/h_n)}$ if we choose l such that $\log(Car(\Im_{r,l}))=o(n)$ and taking into account that the previous inequality is true for every ϵ, it is straightforward to deduce that

$$\operatorname*{lím\,sup}_n \frac{1}{n}\log\left(\mathcal{P}\left\{\int |f_n(X,t)-f_n(Y,t)|\,dt>\lambda\right\}\right) \le -\frac{\lambda^2}{4}(1+o(1)) \qquad (3)$$

Hence, from (2) and (3) the proof is completed when $n_1=n_2$. Under condition $(C2)$ this result can be generalized for any n_1 and n_2 as follows:

Let be

$$D_{m,m}=\int\left|\frac{1}{m}\sum_{i=1}^{m}\frac{1}{h_m}K\left(\frac{x_i-t}{h_m}\right)-\frac{1}{m}\sum_{i=1}^{m}\frac{1}{h_M}K\left(\frac{y_i-t}{h_M}\right)\right|dt$$

where $m=\text{mín}\{n_1,n_2\}$ and $M=\text{máx}\{n_1,n_2\}$. By comparing $D_{m,m}$ and D_{n_1,n_2} we obtain the following inequalities:

$$D_{m,m}-2\frac{M-m}{m}\le D_{n_1,n_2}\le D_{m,m}+2\frac{M-m}{m}$$

Under condition $(C2)$ we have that, for all $\epsilon>0$, there exists n^1 such that for each $m>n^1$ where $2(M-m)/m<\epsilon$. Therefore, for all $n_1,n_2>n^1$, we obtain that

$$\begin{aligned}\frac{n_1+n_2}{2n_1n_2}\log\mathcal{P}\{D_{n_1,n_2}>\lambda\} &= \lim_{\epsilon\to 0}\frac{n_1+n_2}{2n_1n_2}\log\mathcal{P}\{D_{n_1,n_2}>\lambda+\epsilon\}\\ &\le \frac{1}{m}\log\mathcal{P}\left\{D_{m,m}+2\frac{M-m}{m}>\lambda+\epsilon\right\}\le\frac{1}{m}\log\mathcal{P}\{D_{m,m}>\lambda\} \qquad (4)\end{aligned}$$

On the other hand,

$$\begin{aligned}\frac{n_1+n_2}{2n_1n_2}\log\mathcal{P}\{D_{n_1,n_2}>\lambda\} &\ge \frac{1}{M}\log\mathcal{P}\left\{D_{m,m}-2\frac{M-m}{m}>\lambda\right\}\\ &= \frac{1}{M}\log\mathcal{P}\left\{D_{m,m}>\lambda+2\frac{M-m}{m}\right\}\ge\frac{1}{M}\log\mathcal{P}\{D_{m,m}>\lambda+\epsilon\}\end{aligned}$$

then

$$\frac{n_1+n_2}{2n_1n_2}\log\mathcal{P}\left\{D_{n_1,n_2}>\lambda\right\}\geq\frac{1}{M}\log\mathcal{P}\left\{D_{m,m}>\lambda\right\}\tag{5}$$

Taking into account inequalities (4) and (5), the proof is straightforward. □

Observación 2. A weaker upper bound can be derived, easily, from the triangular inequality and the result of Louani (2000). Directly,

$$\mathcal{P}\left\{\int|f_n(X,t)-f_n(Y,t)|\,dt>\lambda\right\}\leq 2\mathcal{P}\left\{\int|f_n(X,t)-f(t)|\,dt>\frac{\lambda}{2}\right\}\leq e^{-\frac{n\lambda^2}{8}}$$

3. Bahadur Relative Efficiency

In order to show the application of the previous result, we use it to calculate the Bahadur slope (BS) of a test, based on the D_{n_1,n_2} distance, to determine whether or not two continuous random variables have the same distribution. On the other hand, we study its Bahadur relative efficiency (BRE) with respect to the classical nonparametric Mann-Whitney test.

Let $X=\{x_1,\ldots,x_{n_1}\}$ and $Y=\{y_1,\ldots,y_{n_2}\}$ be two independent random samples from two continuous distributions F_1 and F_2 with densities f_1 and f_2, respectively. From the above result one can compute the Bahadur slope (Bahadur & Zabell 1979) for a test based on $L_1(f_{n_1},f_{n_2})$ statistic and an arbitrary alternative hypothesis.

From Devroye (1983) we obtain,

$$\int|f_{n_1}-f_{n_2}|\xrightarrow{\mathcal{P}}\int|f_1-f_2|$$

so the Bahadur slope of $L_1(f_{n_1},f_{n_2})$ is

$$BS_{L_1}=\frac{1}{2}\left(\int|f_1-f_2|\right)^2(1+o(1))$$

Under the same conditions, it follows from the asymptotic distribution of the statistic that the Bahadur slope of the Mann-Whitney (M_W) test is

$$BS_{M_W}=\frac{3}{4}\left((1-2\int F_1f_2)^2+(1-2\int F_2f_1)^2)\right)$$

Hence, the Bahadur relative efficiency between the test based on L_1 and the Mann-Whitney one is

$$BRE_{L_1/M_W}=\frac{2\left(\int|f_1-f_2|\right)^2}{3\left((1-2\int F_1f_2)^2+(1-2\int F_2f_1)^2)\right)}\tag{6}$$

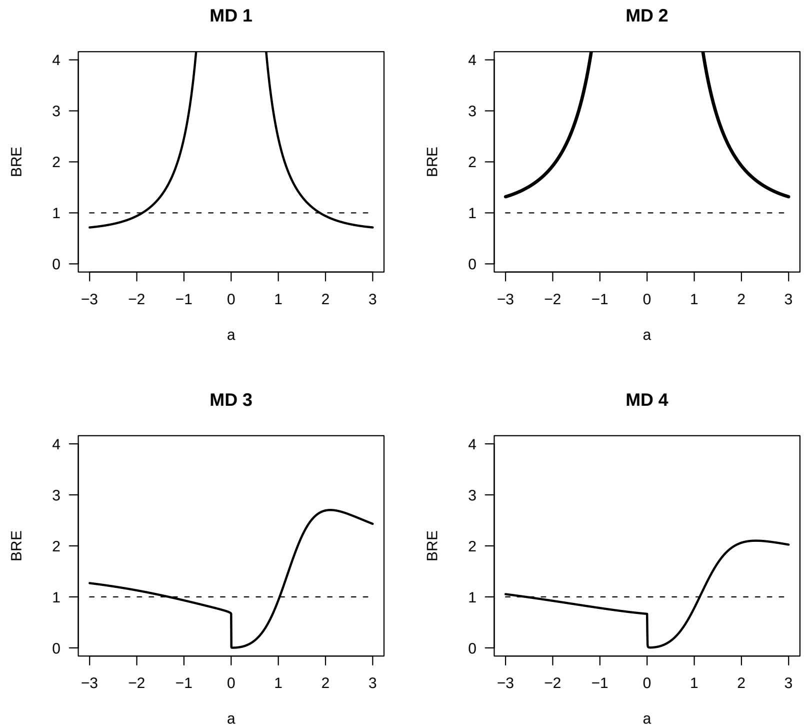

Figura 1: Bahadur relative efficiency (BRE) between test based on $L_1(f_{n_1}, f_{n_2})$ and the Mann-Whitney one for the four different models.

In order to get some particular illustration of the relative efficiency for both considered tests, we have computed (6) in four different situations (see Figure 1). A sample is drawn from the standard normal density $\varphi_{0,1}(t)$ and the other one, $f_2(t)$, follows one of the following densities,

MD 1. $f_2(t) = \varphi_{a,1}(t)$

MD 2. $f_2(t) = \varphi_{a,3}(t)$

MD 3. $f_2(t) = \chi^2_3(at)$

MD 4. $f_2(t) = \chi^2_4(at)$

here, $\varphi_{\mu,\sigma}(t)$ is the normal density function with mean μ and standard deviation σ, $\chi^2_k(t)$ is the density function of a χ^2 distribution with k degrees of freedom and a takes values within $(-3, 3)$.

Figure 1 reveals that the M_W test is more efficient (in the Bahadur sense) than the L_1 test whenever the difference among the densities is mainly in location and large while the L_1 test is more efficient when the main difference is in the shape and neither function uniform dominates to the other. These conclusions are strongly consistent with the obtained ones in other studies, which consider two sample tests based on kernel density estimator as Cao & Van Keilegom (2006) or Martínez-Camblor (2008).

[Recibido: octubre de 2008 — Aceptado: noviembre de 2009]

Referencias

Bahadur, R. R. & Zabell, S. L. (1979), 'Large Deviations of the Sample Mean in General Vector Space.', *Annals of Probability* **7**(4), 587–621.

Beirlant, L., Devroye, L., Györfi, L. & Vajda, A. (2001), 'Large Deviations of Divergence Measures on Partitions', *Journal of Statistical Planning and Inference* **93**, 1–16.

Berlinet, A., Devroye, L. & Györfi, L. (1995), 'Asymptotic Normality of the L_1-error in the Histogram Density Estimation', *Statistics* **26**, 329–343.

Cao, R. & Lugosi, G. (2005), 'Goodness-of-fit Tests Based on teh Kernel Density Estimate', *Scandinavian Journal of Statistics* **32**, 599–617.

Cao, R. & Van Keilegom, I. (2006), 'Empirical Likelihood Tests for Two-Sample Problems via Nonparametric Density Estimation', *Canadian Journal of Statistics* **34**, 61–77.

Devroye, L. (1983), 'The Equivalence of Weak, Strong and Complete Convergence in L_1 for Kernel Density Estimates', *Annals of Statistics* **11**, 896–904.

Devroye, L. (1987), *A course in Density Estimation*, Birkhauser, Boston, United States.

Devroye, L. & Gyorfi, L. (1985), *Nonparametric Density Estimation. The L_1-View*, Wiley, New York, United States.

Devroye, L. & Wagner, T. J. (1979), 'The L_1 Convergence of Kernel Density Estimates', *Annals of Statistics* **7**, 1136–1139.

Hórvath, L. (1991), 'On L_p-Norms of Multivariate Density Estimations', *Annals of Statistics* **19**(4), 1933–1949.

Konakov, V. (1978), 'Complete Asymptotic Expansions for the Maximun Deviation of the Empirical Density Function', *Theory Probability Applied* **28**, 495–509.

Louani, D. (1998), 'Large Deviations Limit Theorems for the Kernel Density Estimator', *Scandinavian Journal of Statistics* **25**(1), 243–253.

Louani, D. (2000), 'Large Deviation for L_1-Distance in Kernel Density Estimation', *Journal of Statistical Planning and Inference* **90**, 177–182.

Louani, D. (2005), 'Uniform L_1-Distance Large Deviations for Nonparametric Density Estimation', *Test* **14**, 1–24.

Martínez-Camblor, P. (2008), 'Tests de hipótesis para contrastar la igualdad entre k poblaciones', *Revista Colombiana de Estadística* **31**(1), 1–18.

Martínez-Camblor, P. & Corral, N. (2008), 'Weaker Conditions for Asymptotic Approximation to L_P-norms of the Kernel Estimators', *InterSTAT Journal* **june**, 1–18.

Osmoukhina, A. V. (2001), 'Large Deviations Probabilities for a Test of Symmetry Based on Kernel Density Estimator', *Statistics and Probability Letters* **54**(4), 363–371.

Parzen, E. (1962), 'On Estimation of a Probability Density Function and Mode', *Annals of Mathematical Statistics* **33**, 1065–1076.

Rosenblatt, M. (1956), 'Remarks on Some Nonparametric Estimates of a Density Functions', *Annals of Mathematical Statistics* **27**, 832–837.

Silverman, B. W. (1978), 'Weak and strong uniform consistency of the kernel estimate of a density and its derivates', *Annals of Statistics* **6**(1), 177–184.

Van der Vaart, A. W. (1998), *Asymptotic Statistics*, Cambridge University Press, London, England.