QÜESTIIÓ, Vol 17, 1 pp. 3-32, 1993

ESTIMACIÓN DE LA FUNCIÓN DE DENSIDAD CON OBSERVACIONES OBTENIDAS EN INSTANTES ALEATORIOS

JOSÉ A. VILAR y JUAN M. VILAR*

Departamento de Matemáticas

Universidad de La Coruña

Sea $X(t)$ un proceso estacionario en tiempo continuo con función de densidad marginal univariante $f(x)$. A partir de un conjunto de n observaciones; $X(\tau_1), X(\tau_2), \ldots, X(\tau_n)$ recogidas en instantes muestrales, τ_i, espaciados irregularmente o aleatorios, se estudia la estimación no paramétrica de $f(x)$, utilizando un estimador recursivo tipo núcleo.

Asumiendo condiciones débiles de dependencia ($\alpha-$mixing) se obtiene la expresión del sesgo y varianza del estimador definido, así como propiedades de normalidad asintótica.

Estimation of the pobability density from random sampling.

Key words: estimación recursiva de la densidad, muestreo aleatorio, procesos mixing de parámetro continuo.

Clasificación A.M.S. (1991): 62G07.

*Juan Manuel Vilar Fernández. Facultad de Informática. Campus da Zapateira s/n. LA CORUÑA, 15071.

–Article rebut el març de 1992.

–Acceptat el desembre de 1992.

1. INTRODUCCIÓN. DEFINICIONES

En los estudios relativos a la estimación de curvas notables (densidad, distribución, regresión...) asociadas a un proceso estacionario real en tiempo continuo, $X(t)$, a partir de una muestra $X(\tau_1), X(\tau_2), \ldots, X(\tau_n)$, generalmente se supone que los instantes muestrales se han tomado equiespaciados. En muchos casos, este planteamiento no es correcto, ya que por distintas causas: instrumentos de medición sujetos a un margen de error, imposibilidad física de muestreo periódico, diseño del experimento, etc., los instantes de observación son aleatorios, lo que va a influir en los resultados de la estimación, pues éstos se ven afectados por la posible dependencia de los elementos muestrales que, en general, es función de la distancia temporal que hay entre ellos.

En este trabajo se estudian propiedades asintóticas de la estimación no paramétrica, recursiva, tipo núcleo de la función de densidad asociada a un proceso estacionario $X(t)$ en tiempo continuo a partir de un conjunto de observaciones tomadas en instantes aleatorios. Indicando la influencia de la aleatoriedad de las observaciones en el error cuadrático medio de la estimación. Se considerarán dos estructuras aleatorias concretas de recogida de datos que se exponen a continuación:

1.1. El Modelo

Sea el modelo estocástico (continuo-discreto) definido por el par (X, T), donde,

* **La primera componente $\boldsymbol{X}$** $= \{X(t): t \in \mathbb{R}\}$ es un proceso estocástico real en tiempo continuo, verificando:

1. Es estrictamente estacionario, con función de densidad $f(x)$ contínua y acotada.
2. Verifica la condición de dependencia "fuertemente mixing", (α−mixing), esto es, si denotamos por F_a^b la σ−álgebra generada por $\{X(t): a \leq t \leq b\}$, con $-\infty < a \leq b < +\infty$, se verifica para $t > 0$,

$$\text{Sup}\left\{|P(AB) - P(A)P(B)|: A \in F_{-\infty}^{0},\, B \in F_t^{+\infty}\right\} = \alpha(t) \downarrow 0$$

Esta condición de dependencia fue introducida por Rosenblatt y es de las más débiles, verificándose en múltiples situaciones.

* **La segunda componente $\boldsymbol{T} = \{\tau_k : k \in \mathbb{N}\}$** es una sucesión real, estrictamente creciente de instantes aleatorios, a la que dotaremos de dos posibles estructuras:

ESTRUCTURA OI (Observaciones Irregulares): el proceso $\boldsymbol{T} = \{\tau_k : k \in \mathbb{N}\}$ es de la forma $\tau_k = k/\beta + Z_k$, con $k = 0, 1, 2, \ldots$, $\beta > 0$, siendo Z_k variables aleatorias independientes e igualmente distribuidas, con función de densidad $g(x)$ simétrica y con soporte $[-1/2\beta,\ 1/2\beta]$.

La estructura OI es la combinación de un factor determinista (k/β) y un factor aleatorio (Z_k), siendo ésta la que lo diferencia del muestreo periódico.

ESTRUCTURA PR (Procesos de Renovación): el proceso $\boldsymbol{T} = \{\tau_k : k \in \mathbb{N}\}$ es de la forma $\tau_k = \sum_{i=1}^{k} t_i$, con $k = 1, 2, 3, \ldots$ y $\tau_0 = 0$, siendo la sucesión de tiempos intermedios, t_i, variables aleatorias i.i.d., con distribución $G(t)$ absolutamente contínua definida en $[0, \infty)$ y densidad $g(t)$, verificando que $E(t_i) = \frac{1}{\beta} < \infty$. Se denotará N_t al "número de observaciones realizadas en el intervalo $(0, t]$", siendo la "función de renovación" (renewal function): $H(t) = E(N_t)$, y la "densidad de renovación" (renewal density): $h(t) = \mathrm{d}H(t)/\,\mathrm{d}t$. Una interpretación de este modelo se basa en considerar que β representa "a largo plazo" el número medio de observaciones realizadas en la unidad de tiempo. De particular interés es el caso en que $G(t)$ es exponencial, entonces PR es un proceso de Poisson de media β.

Estas estructuras **OI** y **PR** son bastante usuales y marcan una progresiva aleatorización en la recogida de los datos. Han sido utilizadas, entre otros, por Masry (1983) aunque autores como Stoyanov-Robinson (1991) y Blum-Boyles (1981) han utilizado otros modelos.

Además en este trabajo se supone que el proceso en tiempo contínuo, $X(t)$, y el proceso en tiempo discreto, T, son independientes.

1.2. Definición del estimador

La estimación no paramétrica de la densidad ha sido ampliamente estudiada en los últimos años, tanto en un contexto de datos independientes (Silverman, 1986) como de dependencia (Gyorfi y otros, 1989), por ser un conjunto de técnicas intuitivas y de fácil cálculo que permiten obtener buenos resultados bajo condiciones muy generales. Siendo el estimador más estudiado y utilizado el tipo núcleo, introducido por Rosenblatt-Parzen y cuya definición es la siguiente:

$$(1) \qquad f_n(x) = \frac{1}{n}\sum_{i=1}^{n} K_n(x - X_i), \quad \text{con } K_n(u) = \frac{1}{h_n} K\left(\frac{u}{h_n}\right)$$

con $K(u)$ la función núcleo (generalmente una función de densidad) y h_n el parámetro ventana que indica el nivel de suavización que se introduce en la estimación.

Este estimador ha sido estudiado por Masry (1983) con muestreo aleatorio, aunque bajo este supuesto consideramos de interés la utilización de estimadores recursivos, ya que al ir obteniendo secuencialmente nuevos datos, las estimaciones son más fáciles de actualizar ganando en tiempo de computación y ahorro de memoria. Por ello, en este trabajo, se estudia el siguiente estimador recursivo, tipo núcleo:

$$(2) \qquad \hat{f}_n(x) = \left(\sum_{i=1}^{n} h_i^{\eta}\right)^{-1} \left[\sum_{j=1}^{n} h_j^{\eta} K_j \left(x - X(\tau_j)\right)\right], \quad \eta \in \mathbb{R}$$

que verifica la siguiente relación de recursividad:

$$\hat{f}_{n+1}(x) = H_{n+1}^{-1}(x) \left[H_n \hat{f}_n(x) + h_{n+1}^{\eta} K_{n+1} \left(x - X(\tau_{n+1})\right)\right], \ \text{con } H_n = \left(\sum_{i=1}^{n} h_i^{\eta}\right) \tag{3}$$

El estimador definido, aunque muy general, es un caso particular del introducido por Deheuvels (1974) y estudiado por Wolverton-Wagner (1969) para $\eta = 0$ en el supuesto de independencia y por Masry (1986) para observaciones regulares dependientes. El parámetro η influye, de forma inversa, en el sesgo y la varianza del estimador y normalmente se elige en el intervalo [0,1] que es donde se obtienen los mejores resultados (Wertz, 1985).

En el apartado 2, de este trabajo, se obtienen las expresiones del sesgo y varianza del estimador de la densidad definido en (2) obtenido a partir de una muestra $X(\tau_1), X(\tau_2), \ldots, X(\tau_n)$, teniendo los instantes muestrales una estructura OI o PR, e indicando la influencia de la aleatoriedad en la recogida muestral. Como consecuencia se obtiene la Media del Error Cuadrático Integrado, MECI, del estimador y su consistencia en media cuadrática. En el apartado 3 se dan las condiciones para obtener la normalidad asintótica del estimador. Las demostraciones de los resultados obtenidos pueden verse en el Apéndice final.

2. SESGO Y VARIANZA

En el resto del trabajo se supone que se verifican las siguientes hipótesis:

H.1. La función núcleo, $K(u)$, está acotada, tiene soporte compacto y es de orden s de forma que:

$$\int K(u)\,u^j = C_j = 0 \text{ para } j = 1,\ldots,s-1; \quad \int K(u) = C_0 = 1, \quad \int K(u)\,u^s = C_s < \infty$$

H.2. La sucesión de parámetros h_n verifica:

i) $h_n \longrightarrow 0$ y $nh_n \longrightarrow \infty$ cuando $n \longrightarrow \infty$

ii) $\displaystyle \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} \left(\frac{h_i}{h_n}\right)^{j+\eta} = \theta_{j+\eta} < \infty$ para $j \leq s+1$, $j \in \mathbb{N}$ y $\eta \in [0,1]$

Estas hipótesis son poco restrictivas, siendo usual utilizar funciones núcleo de orden dos o superior. La primera parte de la hipótesis H1 es clásica en los estudios de estimación no paramétrica, siendo la segunda parte de tipo técnico, verificándose para la elección usual de $h_n = Cn^{-\alpha}$, $\alpha \in (0,1)$, ya que entonces $\theta_j = (1-\alpha j)^{-1}$, para $0 < \alpha j < 1$.

Bajo estas hipótesis se verifica el siguiente Lema que será útil en la demostración de los resultados obtenidos. Su demostración puede verse en Masry (1986).

Lema 1

Sea $g(x)$ una función de L_1, entonces:

i) $\displaystyle \int_{-\infty}^{+\infty} K_n(x-u)g(u)\,\mathrm{d}u = g(x)$

ii) $\displaystyle \lim_{n\to\infty} \int_{-\infty}^{+\infty} h_n K_n(x-u)K_n(y-u)g(u)\,\mathrm{d}u = \delta_{x,y}\,g(x)D_k,$

con $D_k = \displaystyle\int K^2(u)\,\mathrm{d}u$

para todo x punto de continuidad de g.

A continuación se obtiene la expresión del sesgo del estimador definido en (1).

Teorema 1

Se verifican las hipótesis H1–H2 y además:

H.3. La función $f(x)$ es $(s+1)$ veces continuamente diferenciable con derivadas acotadas.

Entonces:

$$\text{Sesgo}\left(\hat{f}_n(x)\right) = \frac{f^{(s)}(x)}{s!} C_s H(\eta)(-h_n)^s + O\left(h_n^{s+1}\right) \tag{4}$$

siendo $H(\eta) = \theta_{\eta+s}/\theta_\eta$.

Comentarios:

El sesgo del estimador $\hat{f}_n$ no depende del tipo de instantes muestrales ni del mayor o menor grado de dependencia de los datos. Y sí depende de la función de densidad a estimar, $f(x)$, por medio de sus derivadas, de la función núcleo por la constante C_s, del parámetro $\eta \in [0,1]$, a través de la función $H(\eta)$, que para la elección del parámetro de suavización $h_n = Cn^{-\alpha}$, se obtiene que $H(\eta)$ es una función estrictamente creciente y, por tanto, el menor sesgo se obtiene para $\eta = 0$. Y, finalmente, el sesgo depende del parámetro de suvización, cuanto menor sea, menor es el sesgo pero mayor será la varianza como se verá a continuación.

Comparando el estimador recursivo $\hat{f}_n$ con el no recursivo f_n (dado en (1)) se obtiene que Sesgo $\left(\hat{f}_n(x)\right) = H(\eta)\,\text{Sesgo}\,(f_n(x))$, siendo la función $H(\eta)$ acotada inferiormente por 1, por tanto, el sesgo del estimador recursivo es siempre mayor que el del no recursivo (en la mayoría de las situaciones $H(\eta)$ está entre 3 y 4) aunque son del mismo orden (h_n^s bajo las hipótesis del teorema).

A continuación se estudia la covarianza del estimador $\hat{f}_n(x)$ que sí depende de la estructura de los instantes aleatorios en que se obtienen los datos muestrales y de la dependencia entre éstos.

En primer lugar se considera la estructura de OBSERVACIONES IRREGULARES (OI), obteniéndose:

Teorema 2

Sea el modelo estocástico $(X, T = \text{OI})$ definido en el apartado 1, se verifican las hipótesis H1–H2′ (H2′ es igual que H2, con $j = 0$ y $\eta \in [-1,1] \subset \mathbb{R}$) y además la sucesión de coeficientes mixing $\alpha(t)$ es tal que:

H.4. $\int_0^\infty (\alpha(t))^q < \infty, \quad$ para algún $0 < q < 1$

Entonces

$$\left|\mathrm{Cov}\left(\hat{f}_n(x), \hat{f}_n(y)\right)\right| \leq 20\beta \left(\frac{1}{nh_n^{1+q}}\right) (f(x)f(y))^{\frac{1-q}{2}} K_q^{1-q} V(\eta) \int_0^\infty (\alpha(t))^q \, \mathrm{d}t +$$

$$+ \left(\frac{1}{nh_n}\right) \delta_{x,y} V(\eta) f(x) D_k + o\left(\frac{1}{nh_n}\right) \tag{5}$$

siendo $K_q = \int K(u)^{2/(1-q)} \, \mathrm{d}u, \quad V(\eta) = \theta_{2\eta-1}/\theta_\eta^2, \; D_k = \int K^2(u) \, \mathrm{d}u$

Si H4 se verificase para algún $0 < q < 1/2$ y además,

H.5. La densidad conjunta $f(x, y, t)$ asociada al vector aleatorio $(X(\tau), X(\tau + t))$ existe y está acotada sobre el plano XY uniformemente cuando $|t| > 1/\beta$.

Entonces:

$$\lim_{n\to\infty} nh_n \left|\mathrm{Cov}\left(\hat{f}_n(x), \hat{f}_n(y)\right)\right| \leq V(\eta) \left(f(x)\delta_{x,y} + [f(x)f(y)]^{1/2}\right) D_k \tag{6}$$

(Esta expresión es válida para toda densidad $g(x)$ asociada a Z_k).

Finalmente, si además se tiene:

H.6. $p(x, y) := \int_0^\beta \gamma(t - \beta^{-1}) f(x, y, t) \, \mathrm{d}t \leq M < \infty$, siendo $\gamma(t)$ la función de densidad asociada a la variable $(Z_i - Z_j)$.

Entonces:

$$\lim_{n\to\infty} nh_n \left|\mathrm{Cov}\left(\hat{f}_n(x), \hat{f}_n(y)\right)\right| = V(\eta) f(x) \delta_{x,y} D_k \tag{7}$$

Comentarios:

1. En la cota para la covarianza del estimador $\hat{f}_n$ (expresión (5)) puede verse la influencia del muestreo aleatorio, ya que ésta es directamete proporcional a β. Como $1/\beta$ representa el promedio de tiempo transcurrido entre dos instantes muestrales consecutivos τ_i y τ_{i+1}, para grandes valores β habrá una fuerte dependencia entre $X(\tau_i)$ y $X(\tau_{i+1})$, por lo que será mayor la covarianza del estimador.

2. El orden de la cota de la covarianza es $(nh_n^{1+q})^{-1}$, mayor que en el supuesto de independencia que es $(nh_n)^{-1}$. Se puede conseguir que la cota sea de este orden exigiendo hipótesis más severas, como se indica en las expresiones (6) y (7), o bien, bajo condiciones de dependencia más restrictivas, como la de ser uniformemente mixing (φ−mixing) con coeficientes, $\varphi(t)$, verificando $\int_0^{+\infty} \varphi(t)^{1/2}\, dt < \infty$.

3. De las expresiones (5), (6) y (7) se deduce de forma directa una cota y expresiones asintóticas de la Varianza de $\hat{f}_n(x)$, en distintos supuestos con sólo hacer $x = y$.

4. Bajo la hipótesis H5 se anula la influencia del valor de β y de los coeficientes mixing, ello es debido a que esta hipótesis es rigurosa para grandes valores de β puesto que $f(x, y, t)$ es divergente cuando $t \longrightarrow 0$. Pero, aún en este supuesto, $B_n(x, y) = V(\eta)[f(x)f(y)]^{1/2} D_k (nh_n)^{-1}$, que aparece en (6), recoge la influencia del muestreo aleatorio a través ahora de la densidad $g(x)$, obteniendo la misma expresión en el término principal (de orden $(nh_n)^{-1}$) que en el caso de independencia de las observaciones. De esta forma, si $x = y$, la cota para la varianza del estimador $\hat{f}_n(x)$ duplica en términos absolutos a la del mismo estimador cuando la distancia entre los tiempos de muestreo no están sujetos a ningún tipo de irregularidad aleatoria. Recogiéndose la influencia de la dependencia en términos secundarios de menor orden.

5. La hipótesis H6 finalmente consigue eliminar toda la influencia de la densidad $g(x)$ y permite igualar la cota alcanzada por este mismo estimador en el supuesto de independencia y observaciones periódicas.

6. La diferencia en las expresiones 5–6–7 entre utilizar el estimador recursivo, $\hat{f}_n(x)$, o el no recursivo, $f_n(x)$, viene dada por el factor $V(\eta)$, que en el último caso vale 1.

   La función $V(\eta) = \theta_{2\eta-1}/\theta_\eta^2$, con $\eta \in [0, 1]$, para la elección clásica del parámetro de suavización $h_n = Cn^{-\alpha}$, tiene la forma: $V(\eta) = \dfrac{1 - 2\eta\alpha + \eta^2\alpha^2}{1 - 2\eta\alpha + \alpha}$

que es una función decreciente y acotada superiormente por 1. Por tanto, cuanto mayor es η, mayor es el sesgo pero menor la varianza del estimador recursivo, siendo para todo $\eta \in [0,1]$ menor la varianza del estimador recursivo que la del no recursivo.

Es conocido que una de las medidas de bondad de ajuste más utilizadas en la estimación no paramétrica de curvas de probabilidad es el Error Cuadrático Medio Integrado (MECI) (Hardle-Marron, 1986) que para la función de densidad viene definido como sigue:

$$
\begin{aligned}
\text{MECI}\left(\hat{f}_n\right) &= E\left\{\int \left(\hat{f}_n(x) - f(x)\right)^2 W(x)\,\mathrm{d}x\right\} = \\
&= E\left\{\int \left[\text{Sesgo}\left(\hat{f}_n(x)\right)\right]^2 W(x)\,\mathrm{d}x\right\} + \\
&+ E\left\{\int \text{Varianza}\left(\hat{f}_n(x)\right) W(x)\,\mathrm{d}x\right\}
\end{aligned}
$$

siendo $W(x)$ una función peso.

De esta expresión y de los teoremas 1 y 2 se sigue de forma inmediata el siguiente resultado:

Corolario 1

"Bajo las hipótesis H1–H5 se obtiene que:

$$
\begin{aligned}
\text{MECI}\left(\hat{f}_n\right) &= \frac{h_n^{2s} C_s^2 H(\eta)^2}{s!^2} \int \left(f^{(s)}(x)\right)^2 W(x)\,\mathrm{d}x + \\
&+ \frac{2}{nh_n} K_2 V(\eta) \int f(x) W(x)\,\mathrm{d}x + o\left(h_n^{2s} + \frac{1}{nh_n}\right) \text{''}
\end{aligned}
\tag{8}
$$

De la expresión (8) se deduce que el elemento de mayor influencia en el MECI es el parámetro de suavización, que actúa de balanza entre el sesgo y la varianza. Un efecto análogo tiene el parámetro η, pero con mucho menor peso, y como puede verse en Wertz (1985) el valor $\eta = 0$ minimiza el MECI, lo que hace que sea el más utilizado. En este caso el MECI del estimador recursivo es mayor que el del no recursivo, aunque no en mucha proporción, por lo que es recomendable la utilización del estimador recursivo en la situación en estudio.

Para la elección $h_n = Cn^{-\alpha}$ se obtiene que el α que minimiza el MECI es $\alpha = 1/(1+2s)$, siendo el MECI $= O\left(n^{-2s/(1+2s)}\right)$, resultado igual al obtenido para el caso de datos independientes (Silverman, 1986).

Sobre la covarianza del estimador $\hat{f}_n(x)$ cuando se considera la estructura de PROCESOS DE RENOVACIÓN (PR), se ha obtenido:

Teorema 3

Sea el modelo estocástico $(X, T = \text{PR})$ definido en el apartado 1, donde se ha llamado $h(t)$ a la "densidad renewal" que se supone acotada en $[0, \infty)$. Si se verifican las hipótesis H1–H2$'$ y H4, se obtiene:

$$\left|\text{Cov}\left(\hat{f}_n(x), \hat{f}_n(y)\right)\right| \leq 20\left(\frac{1}{nh_n^{1+q}}\right)(f(x)f(y))^{\frac{1-q}{2}} K_q^{1-q} V(\eta) \int_0^\infty (\alpha(t))^q\, h(t)\ \mathrm{d}t$$

$$+\left(\frac{1}{nh_n}\right)\delta_{x,y} V(\eta) f(x) K_2 + o\left(\frac{1}{nh_n}\right) \tag{9}$$

Si además se verifica:

H.7. $\int_0^\infty (1+t)(\alpha(t))^q < \infty$, para algún $0 < q < 1$ (más estricta que H4)

H.8. $p(u,v,s) := \int_0^\infty f(u,v,t+s) g(t)\ \mathrm{d}t \leq D < \infty$, para $u, v \in \mathbb{R}, s \geq 0$

Entonces

$$\lim_{n\to\infty} nh_n \left|\text{Cov}\left(\hat{f}_n(x), \hat{f}_n(y)\right)\right| = V(\eta) f(x) \delta_{x,y} D_k \tag{10}$$

Comentarios:

1. Se observa que en la cota de la covarianza obtenida en (9) influye el muestreo aleatorio por la densidad $h(t)$, asociada al modelo PR, y la dependencia de las observaciones ($\alpha-$mixing) por el factor $\int \alpha(t)^q h(t)\ \mathrm{d}t$. Además en el supuesto de un proceso de Poisson se obtiene que $h(t) = \beta$, para todo t, y, por tanto, la cota dada en (9) es igual a la obtenida en (5) bajo la hipótesis de observaciones irregulares.

2. Nuevamente se ha obtenido una cota de la covarianza de orden $\left(nh_n^{1+q}\right)^{-1}$, aunque bajo hipótesis adicionales H7–H8 se obtiene que la expresión de la covarianza tiene el mismo término principal que en el caso de independencia, que es del orden $(nh_n)^{-1}$. Influyendo el tipo de muestreo y la dependencia en términos secundarios de menor orden.

3. Al igual que en el teorema 2, la diferencia entre utilizar el estimador recursivo $\hat{f}_n(x)$ y el no recursivo $f_n(x)$ viene dada por el factor $V(\eta)$, siendo válidos los comentarios allí realizados.

4. Haciendo $x = y$ de las expresiones (9) y (10) se deduce la cota y forma de la varianza de $\hat{f}_n(x)$ cuando el muestreo es del tipo PR, lo que permite deducir el siguiente corolario:

Corolario 2

"Bajo las hipótesis de los Teoremas 1 y 3 se sigue que el MECI de $\hat{f}_n$ es el dado en (8)".

3. NORMALIDAD ASINTÓTICA

En este apartado se estudia el problema de encontrar sucesiones de números reales $\{a_n\}$, con $a_n \uparrow \infty$ cuando $n \uparrow \infty$, tales que $\{a(\hat{f}_n(x) - f(x))\}$ tenga distribución asintótica, que en nuestro caso será gaussiana. Esto, intuitivamente, parece posible tras haber comprobado que bajo hipótesis relativamente suaves se obtiene la convergencia en media cuadrática de $\hat{f}_n(x)$ a $f(x)$. Además, este tipo de resultados son interesantes ya que permiten calcular intervalos de confianza asintóticos en torno a $f(x)$ utilizando solamente la información que proporciona la muestra. El resultado que se ha obtenido es el siguiente:

Teorema 4

"Si se verifican las hipótesis del Teorema 2 o del Teorema 3 y las hipótesis adicionales:

H.9. Para algún $\gamma, 0 < \gamma < 1, nh_n^{3-2\gamma+4\eta} \longrightarrow \infty$

H.10. Existe una sucesión de enteros $\{q_n\} \uparrow \infty$ tal que:

i) $q_n = o\left(nh_n^{3-2\gamma+4\eta}\right)^{1/2}$

ii) $\left(\frac{n}{h_n}\right)^{1/2} \sum_{k=q_n, k\in\mathbb{Z}} (\alpha(k/\beta))^{1-\gamma} \longrightarrow 0$

Entonces:

$$(nh_n)^{1/2}\left\{\hat{f}_n(x) - \mathrm{E}\,\hat{f}_n(x)\right\} \xrightarrow{d} \mathrm{N}(0,\sigma_f) \tag{11}$$

siendo $\sigma_f^2 = \mathrm{V}(\eta)f(x)D_k$ (expresión obtenida en (7) y (10) para $x = y$)."

Teniendo en cuenta que $(nh_n)^{1/2}\left\{\hat{f}_n(x) - f(x)\right\}$ se puede descomponer como la suma de: $(nh_n)^{1/2}\left\{\hat{f}_n(x) - \mathrm{E}\,\hat{f}_n(x)\right\} + (nh_n)^{1/2}\left\{\mathrm{E}\,\hat{f}_n(x) - f(x)\right\}$, y, por tanto de los teoremas 1 y 4 se deduce de forma inmediata el siguiente corolario:

Corolario 3

"Si se verifican las hipótesis de los teoremas 1 y 4. Y la hipótesis:

H.11. $nh_n^{1+2s} \downarrow 0$

Entonces:

$$(nh_n)^{1/2}\left\{\hat{f}_n(x) - f(x)\right\} \xrightarrow{d} \mathrm{N}(0,\sigma_f) \tag{12}$$

Comentarios:

1. En la demostración del Teorema 4 se utiliza el denominado "método Bernstein" (ver Peligrad M., 1985) que consiste en descomponer la suma de las variables aleatorias que definen el estimador en sumas de grandes bloques separados por bloques más pequeños, probándose a continuación que la aportación de los bloques pequeños es asintóticamente nula, mientras que los grandes bloques tienden a ser independientes lo que permite aplicar el Teorema Central del Límite de Lindenberg-Feller para variables aleatorias independientes. Este procedimiento ha sido ampliamente utilizado entre otros autores por Rosenblatt (1970), Robinson (1983) ó Masry (1986).

2. Las hipótesis H9 y H10 se pueden debilitar a costa de exigir condiciones de dependencia más restrictivas (imponiendo velocidades de convergencia de tipo exponencial a los coeficientes $\alpha-$mixing), lo que además permite demostrar de forma más sencilla el teorema 4, utilizando un teorema central de Bradley (1981) para disposiciones triangulares de variables aleatorias fuertemente mixing.

APÉNDICE. (Demostración de los Teoremas)

Demostración del TEOREMA 1

Basta tener en cuenta que E $\left(\hat{f}_n(x)\right) = H_n^{-1}\sum_{j=1}^{n} h_j^\eta \left[\int_{-\infty}^{+\infty} \frac{1}{h_j} K\left(\frac{x-u}{h_j}\right) f(u)du\right]$ haciendo el cambio de variable $u = x - vh_j$ y un desarrollo de Taylor de orden s en $f(u)$ se obtiene que:

$$\mathrm{E}\left(\hat{f}_n(x)\right) = H_n^{-1}\sum_{j=1}^{n} h_j^\eta \left[\sum_{r=0}^{s} \frac{f^{(r)}(x)}{r!} C_r(-h_j)^r\right] + O\left(h_j^{s+1}\right)$$

De la aplicación del Lema de Toeplitz, el lema 1 y las hipótesis del teorema se sigue la conclusión de éste.

■

Las demostraciones de los Teoremas 2, 3 y 4 se han desarrollado siguiendo razonamientos análogos a los realizados por Masry (1986), quién, para datos muestrales observados de forma regular y en un contexto de dependencia, estudia el propio estimador $\hat{f}_n(x)$ para $\eta = 0$ y otro estimador dado por:

$$\tilde{f}_n(x) = \frac{1}{nh_n^{1/2}}\left[\sum_{j=1}^{n} h_j^{1/2} K_j\left(x - X_j\right)\right]$$

Acerca de éste último nótese que el estimador definido en (2), $\hat{f}_n(x)$, con $\eta = \frac{1}{2}$ es asintóticamente equivalente a una versión reescalada de $\tilde{f}_n(x)$. En efecto, por el apartado (ii) de la hipótesis H2 relativa a la sucesión de parámetros de suavización h_n, haciendo $\eta = 1/2$ y $j = 0$, se tiene la existencia de:

$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n}\left(\frac{h_i}{h_n}\right)^{1/2} = \theta_{1/2} < \infty$$

Por tanto y en virtud del Lema de Toeplitz:

$$\hat{f}_n(x) \cong \theta_{1/2}^{-1}\tilde{f}_n(x)$$

En consecuencia la varianza asintótica de $\tilde{f}_n(x)$ es del orden de $\theta_{1/2}^2$ por la varianza asintótica de $\hat{f}_n(x)$, con $\eta = 1/2$, y, en este sentido, el estimador definido en (2), $\hat{f}_n(x)$, generaliza también el estudiado por Masry $\left(\tilde{f}_n(x)\right)$.

A continuación se exponen, de forma esquemática, las citadas demostraciones, en las que será útil el siguiente resultado:

Lema 2 (Doe, 1973)

"Sea $\{X(t): t \in \mathbb{R}\}$ un proceso fuertemente mixing y sean ξ y η dos variables aleatorias medibles respecto a las $\sigma-$álgebras $F_{-\infty}^{0}$ y $F_{t}^{+\infty}$, respectivamente, con $t \geq 0$ arbitrario. Si para algún $\mu > 0$ se verifica: $\mathrm{E}\,|\xi|^{2+\mu} < \infty$ y $\mathrm{E}\,|\eta|^{2+\mu} < \infty$, entonces:

$$|\mathrm{Cov}\,(\xi, \eta)| \leq 10\,(\alpha(t))^{\mu/(2+\mu)} \left(\mathrm{E}\,|\xi|^{2+\mu}\,\mathrm{E}\,|\eta|^{2+\mu}\right)^{1/(2+\mu)}"$$

Demostración del TEOREMA 2

$$\mathrm{Cov}\left(\hat{f}_n(x), \hat{f}_n(y)\right) = T_{n,0}(x,y) + \mathcal{R}_n(x,y) \tag{13}$$

siendo

$$\mathcal{R}_n(x,y) = \sum_{j=1}^{n-1} \left(T_{n,j}(x,y) + T_{n,-j}(x,y)\right) \tag{14}$$

$$T_{n,j}(x,y) = \left(\sum_{k=1}^{n} h_k^{\eta}\right)^{-2} \sum_{i=1}^{n-|j|} h_{i+|j|}^{\eta} h_i^{\eta}\,\mathrm{Cov}\left\{K_{i+|j|}\left(x - X\left(\tau_{|j|}\right)\right) K_i\left(y - X(0)\right)\right\}$$

en virtud de la estacionariedad de $X(t)$.

De la hipótesis H2' y el Lema 1 se sigue que:

$$T_{n,0}(x,y) = \frac{1}{nh_n} V(\eta) f(x) \delta_{x,y} D_k + \frac{1}{n} \tag{15}$$

Dado que $T = OI$, $\tau_j = \dfrac{j}{\beta} + z_j$, con $\{z_j\}$ v. a. independientes e idénticamente distribuidas según la densidad común g de soporte en $\left[-\dfrac{1}{2\beta}, \dfrac{1}{2\beta}\right]$. Por tanto, para $i > j$, $\tau_i - \tau_j = \dfrac{i-j}{\beta} + \left(z_i - z_j\right) \Rightarrow$ la densidad asociada a $\tau_i - \tau_j$, para $i > j$ es $\gamma\left(t - \dfrac{i-j}{\beta}\right)$, siendo:

$$\gamma(t) = \left(g * g\right)(t) = \int_{-1/\beta}^{1/\beta} g(t-v) g(v)\,\mathrm{d}v$$

Nótese que las condiciones impuestas a g conducen a que γ sea simétrica y tenga soporte $\left[-1/\beta, 1/\beta\right]$.

Entonces aproximando $T_{n,j}(x,y)$, para $j \neq 0$, por su valor esperado respecto a la v. a. τ_j se obtiene que:

$$T_{n,j}(x,y) = \left(\sum_{k=1}^{n} h_k^\eta\right)^{-2} \sum_{i=1}^{n-j} h_{i+j}^\eta h_i^\eta \int_{-1/\beta}^{1/\beta} \mathrm{Cov}\left\{K_{i+j}\left(x - X\left(t + \frac{j}{\beta}\right)\right),\right.$$

$$\left. K_i\left(y - X(0)\right)\right\} \gamma(t)\, \mathrm{d}t \tag{16}$$

Sean $\xi = K_{i+j}\left(x - X\left(t + \frac{j}{\beta}\right)\right)$, $\eta = K_i\left(y - X(0)\right)$ y $\mu = \dfrac{2q}{1-q}$.

Se tiene que:

$$\mathrm{E}\,|\xi|^{2+\mu} = \frac{1}{h_{i+j}}\, q_{i+j}(x) \quad \text{y} \quad \mathrm{E}\,|\xi|^{2+\mu} = \frac{1}{h_i}\, q_i(y)$$

siendo

$$q_i(z) = \int \frac{1}{h_i}\left|K\left(\frac{z-u}{h_i}\right)\right|^{2+\mu} f(u)\,\mathrm{d}u \xrightarrow[i \to \infty]{} f(z) \int |K(u)|^{2+\mu}\,\mathrm{d}u \tag{17}$$

en virtud del lema 1 y de la hipótesis de continuidad establecida sobre f.

Por aplicación del Lema de Doe en (16):

$$|T_{n,j}(x,y)| \leq 10\left(\alpha\left(t + \frac{j}{\beta}\right)\right)^q \left(\sum_{j=1}^{n} h_j^\eta\right)^{-2} \sum_{i=1}^{n-j} h_{i+j}^\eta h_i^\eta \left(h_{i+j} h_i\right)^{-\frac{1+q}{2}} \cdot$$

$$\cdot \left(q_{i+j}(x) q_i(y)\right)^{\frac{1-q}{2}} \leq$$

$$\leq 10\left(\alpha\left(t + \frac{j}{\beta}\right)\right)^q \left(\sum_{j=1}^{n} h_j^\eta\right)^{-2} \sum_{i=1}^{n} h_i^{2\eta-1} h_i^{-q} \left(q_i(x) q_i(y)\right)^{\frac{1-q}{2}} \tag{18}$$

Supliendo (18) en (14) se tiene, en virtud del Lema de Toeplitz y de (17)

$$|\mathcal{R}_n(x,y)| \leq 20 V(\eta) \frac{1}{n h_n^{1+q}} \left(f(x) f(y)\right)^{\frac{1-q}{2}} K_q^{1-q} \sum_{j=1}^{\infty} \left(\alpha\left(\frac{j}{\beta}\right)\right)^q \tag{19}$$

Llevando ahora las cotas obtenidas en (15) y (19) a (13) obtendremos por H4 la cota de Cov $\left(\hat{f}_n(x), \hat{f}_n(y)\right)$ dada en (5).

Para obtener las expresiones asintóticas de la covarianza de $\hat{f}_n(x)$ dadas en (6) y (7), se considera una sucesión de enteros positivos c_n tal que $c_n \uparrow \infty$, $h_n c_n \downarrow 0$ y $h_n^{2q} c_n \uparrow \infty$ para algún $0 < q < 1/2$. Se considera entonces la descomposición del término $\mathcal{R}_n(x,y)$:

$$\mathcal{R}_n(x,y) = (T_1 + T_{-1}) + \sum_{j=2}^{c_n} (T_j + T_{-j}) + \sum_{j=c_n+1}^{n-1} (T_j + T_{-j}) = \mathcal{S}_1 + \mathcal{S}_2 + \mathcal{S}_3 \tag{20}$$

Se desarrolla cada uno de estos tres sumandos y se obtiene:

$$T_1(x,y) = \left(\sum_{i=1}^{n} h_i^\eta\right)^{-2} \sum_{j=1}^{n-1} h_{j+1}^\eta h_j^\eta \int_{-1/\beta}^{1/\beta} \mathrm{E}\left\{K_{j+1}\left(x - X\left(t + \frac{1}{\beta}\right)\right)\cdot\right.$$

$$\left.\cdot\ K_j\left(y - X(0)\right)\right\} \gamma(t)\,\mathrm{d}t - \left(\sum_{i=1}^{n} h_i^\eta\right)^{-2} \sum_{j=1}^{n-1} h_{j+1}^\eta h_j^\eta \mathrm{E}\left\{K_j\left(y - X(0)\right)\right\}$$

$$\int_{-1/\beta}^{1/\beta} \mathrm{E}\left\{K_{j+1}\left(x - X\left(t + \frac{1}{\beta}\right)\right)\right\} \gamma(t)\,\mathrm{d}t = A_n + B_n \tag{21}$$

Del Lema 1 y el Lema de Toeplitz se sigue que $B_n = O(1/n)$. Y de la desigualdad de Cauchy-Schwartz y la hipótesis H5 se sigue:

$$A_n \le \left(\sum_{i=1}^{n} h_i^\eta\right)^{-2} \sum_{j=1}^{n-1} h_{j+1}^\eta h_j^\eta \left\{\int_0^{1/\beta} \gamma\left(t - \frac{1}{\beta}\right)\,\mathrm{d}t\right\} \frac{1}{h_j^{1/2} h_{j+1}^{1/2}}$$

$$\left\{\int K_{j+1}^2(x-u) f(u) h_{j+1}\,\mathrm{d}u \int K_j^2(y-v) f(v) h_j\,\mathrm{d}v\right\}^{1/2} \quad \Rightarrow$$

$A_n \le \dfrac{1}{n h_n} V(\eta) \left(f(x) f(y)\right)^{1/2} \dfrac{1}{2} D_k$, y por simetría se obtiene:

$$\mathcal{S}_1 \le \frac{1}{n h_n} V(\eta) \left(f(x) f(y)\right)^{1/2} D_k \tag{22}$$

A continuación se prueba que los sumandos $\mathcal{S}_2$ y $\mathcal{S}_3$ tienen cotas de orden inferior a $(n h_n)^{-1}$.

$$\mathcal{S}_2 = 2 \sum_{j=2}^{c_n} T_j = 2 \left(\sum_{i=1}^{n} h_i^\eta\right)^{-2} \sum_{j=2}^{c_n} \sum_{i=1}^{n-j} h_{i+j}^\eta h_i^\eta$$

$$\int_{-1/\beta}^{1/\beta} \left[\int\int K_{i+j}(x-u)K_i(y-v)\left(f\left(u,v,t+\frac{j}{\beta}\right)-f(u)f(v)\right)\mathrm{d}u\,\mathrm{d}v\right]\gamma(t)\,\mathrm{d}t$$

Utilizando H5 y el hecho de ser f acotada se obtiene que:

$$\begin{aligned}
\mathcal{S}_2 &\leq \text{Cte.}\left(\sum_{i=1}^{n} h_i^{\eta}\right)^{-2}\sum_{j=1}^{c_n}\sum_{i=1}^{n} h_i^{2\eta}\int_{-1/\beta}^{1/\beta}\gamma(t)\,\mathrm{d}t = \\
&= \text{Cte.}\, c_n\left(\sum_{i=1}^{n} h_i^{\eta}\right)^{-2}\sum_{i=1}^{n} h_i^{2\eta} = \\
&= \text{Cte.}\, c_n\,\frac{\theta_{2\eta}}{\theta_\eta^2}\frac{c_n}{n} = O(c_n/n) = o(nh_n)^{-1} \qquad (23)
\end{aligned}$$

Para acotar el sumando $\mathcal{S}_3$ se utiliza el Lema 2 (Doe), con $\mu = \dfrac{4q}{1-2q}$

$$\begin{aligned}
\mathcal{S}_3 &\leq \left(\sum_{j=1}^{n} h_j^{\eta}\right)^{-2}\sum_{i=1}^{n} h_i^{2\eta-1}h_i^{-2q}\left(q_i(x)q_i(y)\right)^{\frac{1-2q}{2}}\,10\sum_{j=c_n+1}^{n-1}\left(\alpha\left(\frac{j-1}{\beta}\right)\right)^{2q} \leq \\
&\leq \frac{1}{nh_n^{2q+1}}V(\eta)\left(f(x)f(y)\right)^{\frac{1-2q}{2}}K_q^{1-2q}\sum_{j=c_n}^{\infty}\left(\alpha\left(\frac{j}{\beta}\right)\right)^{2q} \leq \\
&\leq \frac{1}{nh_n^{2q+1}}V(\eta)\left(f(x)f(y)\right)^{\frac{1-2q}{2}}K_q^{1-2q}\frac{1}{c_n}\beta\int_0^{\infty}\left(\alpha(t)\right)^{2q}\mathrm{d}t \leq
\end{aligned}$$

$$(24) \leq O\left(\frac{1}{nc_nh_n^{1+2q}}\right) = O\left(\frac{1}{nh_n}\frac{1}{h_n^{2q}c_n}\right) = o\frac{1}{nh_n}$$

La última desigualdad se obtiene utilizando H4. Ahora sustituyendo (22), (23) y (24) en (20) se sigue que: $\mathcal{R}_n(x,y) \leq \dfrac{1}{nh_n}V(\eta)\left(f(x)f(y)\right)^{1/2}D_k$, de esto y la expresión (15) se obtiene el resultado (6) del Teorema.

Si se utiliza la hipótesis H6 para acotar el término A_n se sigue de manera inmediata que $A_n = o\dfrac{1}{nh_n}$, de donde se obtiene la expresión (7).

■

La prueba del Teorema 3 sigue la misma línea que la del Teorema 2. La única modificación importante se debe a la nueva estructura de T (que es ahora la de un proceso de renovación) y afecta a la distribución de probabilidad de la variable $\tau_i - \tau_j$, $i > j$. En este caso y debido a que la sucesión de tiempos intermedios $\{t_i\}$ es independiente e idénticamente distribuida según una distribución $G(t)$ sobre $[0, \infty)$; la distribución de

$$\tau_{i+j} - \tau_i = \sum_{\ell=i+1}^{i+j} t_\ell = \sum_{\ell=1}^{j} t_\ell$$

es la j–ésima convolución de G consigo mismo. Si se denota a ésta por $G_j(t)$, al aproximar $T_{n,j}(x,y)$ por su valor esperado respecto a τ_j, con $j > 0$, tal y como se hizo en (16), se obtiene:

$$\mathrm{T}_{n,j}(x,y) = \left(\sum_{k=1}^{n} h_k^\eta\right)^{-2} \sum_{i=1}^{n-j} h_{i+j}^\eta h_i^\eta \int_0^\infty \mathrm{Cov}\,\{\mathrm{K}_{i+j}\,(x - \mathrm{X}(t)), \mathrm{K}_i\,(y - \mathrm{X}(O))\}\ \mathrm{d}G_j(t)$$

Demostración del TEOREMA 4

Se supone que el modelo estocástico (X, T) tiene una estructura OI, siendo análoga la demostración para el supuesto de que la estructura fuese PR.

Se desea probar que:

$$(25) \qquad \sqrt{nh_n}\,\frac{\hat{f}_n(x) - \mathrm{E}\,\hat{f}_n(x)}{\sigma_f} \xrightarrow{d} \mathrm{N}(0,1), \quad \text{con } \sigma_f = V(\eta) f(x) D_k$$

Sean

$$(26) \qquad Y_j = h_j^\eta \left(K_j\,(x - X(\tau_j)) - \mathrm{E}\,K_j\,(x - X(\tau_j))\right)$$

$$(27) \qquad \mathcal{S}_n = \sum_{j=1}^{n} Y_j = \sum_{j=1}^{n} h_j^\eta \left(K_j\,(x - X(\tau_j)) - \mathrm{E}\,K_j\,(x - X(\tau_j))\right)$$

Entonces, en virtud de la definición de $\hat{f}_n$ en (2) se tiene que:

$$(28) \qquad \mathcal{S}_n = \left(\sum_{j=1}^{n} h_j^\eta\right) \left(\hat{f}_n(x) - \mathrm{E}\,\hat{f}_n(x)\right) \quad \text{con} \quad \mathrm{E}[\mathcal{S}_n] = 0$$

y dado que el teorema actual verifica las hipótesis del teorema 2 se tiene por (10):

$$(29) \qquad \lim_{n\to\infty} nh_n \operatorname{Var}\left(\hat{f}_n(x)\right) = \sigma_f$$

De (27) y (28) se deduce que probar (25) es equivalente a probar:

$$(30) \qquad \frac{\mathcal{S}_n}{\sigma(\mathcal{S}_n)} \xrightarrow{d} \mathrm{N}(0,1),$$

Pero de (26), (29) y la hipótesis H2, ii) se tiene también que:

$$\frac{nh_n}{n^2 h_n^{2\eta}}\sigma^2(\mathcal{S}_n) = \left(\frac{1}{n}\sum_{k=1}^{n}\frac{h_k^\eta}{h_n^\eta}\right)^2 \left(nh_n \operatorname{Var}\left(\hat{f}(x)\right)\right) \longrightarrow \theta_\eta^2\sigma_f^2$$

Por tanto bastará demostrar que:

$$(31) \qquad \left(nh_n^{2\eta-1}\right)^{-(1/2)} \mathcal{S}_n \xrightarrow{d} \mathrm{N}\left(0, \theta_\eta\sigma_f\right)$$

Para ello se siguen los siguientes pasos:

Paso 1: Descomposición en bloques de $\mathcal{S}_n$

Por las hipótesis H9 y H10 existe una sucesión $\{r_n\} \subset \mathbb{N}$, $\{r_n\} \uparrow \infty$ tal que

$$(32) \qquad r_n q_n = o\left(nh_n^{3-2\gamma+4\eta}\right)^{1/2}$$

y

$$(33) \qquad r_n\left(nh_n^{-1}\right)^{1/2}\sum_{k=q_n}^{\infty}\alpha\left[k/\beta\right]^{1-\gamma} \longrightarrow 0$$

A partir de $\{r_n\}$ se define una nueva sucesión $\{p_n\} \subset \mathbb{N}$ en la forma:

$$(34) \qquad p_n = \left[\frac{(nh_n)^{1/2}}{r_n}\right]$$

donde $[a]$ denota la función parte entera de a.

De H9, H10, (32), (33) y (34) se deducen las siguientes relaciones

$$(35) \qquad \frac{q_n}{p_n h_n^{2\eta-\gamma+1}} \simeq \frac{q_n r_n}{\left(nh_n^{3-2\gamma+4\eta}\right)^{1/2}} \longrightarrow 0 \qquad \left(\Rightarrow \frac{q_n}{p_n} \longrightarrow 0 \ \text{ y } \ \frac{q_n}{p_n h_n^{2\eta}} \longrightarrow 0\right)$$

$$(36) \qquad \frac{p_n}{nh_n^{2\eta-\gamma+1}} \simeq \frac{(nh_n)^{1/2}}{r_n nh_n^{2\eta-\gamma+1}} = \frac{h_n}{r_n\left(nh_n^{3-2\gamma+4\eta}\right)^{1/2}} \longrightarrow 0 \qquad \left(\Rightarrow \frac{p_n}{n} \longrightarrow 0\right)$$

$$\frac{1}{h_n^{2\eta-\gamma+1}}\sum_{i=q_n}^{\infty}\alpha\left(\frac{i}{\beta}\right)^{1-\gamma} = \left(\left(nh_n^{-1}\right)^{1/2}\sum_{i=q_n}^{\infty}\alpha\left(\frac{i}{\beta}\right)^{1-\gamma}\right)\frac{h_n}{\left(nh_n^{3-2\gamma+4\eta}\right)^{1/2}} \longrightarrow 0$$

(37)

$$\frac{n}{p_n}\alpha\left(\frac{q_n}{\beta}\right) \leq \frac{n}{p_n}\sum_{i=q_n}^{\infty}\alpha\left(\frac{i}{\beta}\right)^{1-\gamma} \simeq r_n\left(nh_n^{-1}\right)^{1/2}\sum_{i=q_n}^{\infty}\alpha\left(\frac{i}{\beta}\right)^{1-\gamma} \longrightarrow 0$$

(38)

A continuación sea, para cada $n \in \mathbb{N}$, la partición de $\mathcal{S}_n$ en $2k_n + 1$ subconjuntos donde $k_n = \left[n/(p_n + q_n)\right]$ realizada en la siguiente forma:

$$\mathcal{S}_n = \mathcal{S}_n^1 + \mathcal{S}_n^2 + \mathcal{S}_n^3 = \sum_{j=0}^{k-1}\beta_j + \sum_{j=0}^{k-1}\pi_j + \theta_k \tag{39}$$

con

$$\beta_j = \sum_{i=1}^{p} Y_{m_j+i}, \quad \pi_j = \sum_{i=p+1}^{p+q} Y_{m_j+i}, \quad \theta_k = \sum_{i=m_k+1}^{n} Y_i \tag{40}$$

siendo $m_j = j(p+q)$, $j = 0, \ldots, k$; de tal forma que cada β_j constituye un bloque "grande" sumando p_n variables, cada π_j uno "pequeño" sumando q_n variables y θ_k un bloque residual.

PASO 2: CUANDO $n \longrightarrow \infty$ SE PRUEBA QUE:

$$\frac{1}{nh_n^{2\eta-1}}\,\mathrm{E}\left(\left|\mathcal{S}_n^2\right|\right)^2 \longrightarrow 0 \quad \text{y} \quad \frac{1}{nh_n^{2\eta-1}}\,\mathrm{E}\left(\left|\mathcal{S}_n^3\right|\right)^2 \longrightarrow 0$$

Primeramente nótese que por ser K acotado y verificar H1 también $\int |K(x)|^{2/\gamma}$ d$x < \infty$ para $0 < \gamma < 1$, y por el Lema 1, apdo. i:

$$\mathrm{E}\left\{\left|k_i(x - X(\tau_i))\right|^{2/\gamma}\right\} = h_i^{-(2/\gamma)+1}\int\left|h_i^{-1}K\left(\frac{x-u}{h_i}\right)\right|^{2/\gamma} f(u)\,\mathrm{d}u =$$
$$= h_i^{-(2/\gamma)+1}\Phi_i(x)$$

con

$$\Phi_i(x) \xrightarrow{i} f(x)\int |K(u)|^{2/\gamma}\,\mathrm{d}u \tag{41}$$

De forma análoga

$$\text{(42)} \qquad \operatorname{Var} Y_i \leq h_i^{2\eta} \operatorname{E} K_i^2\big(x - X(\tau_i)\big) = h_i^{2\eta-1}\Lambda_i(x)$$

siendo $\Lambda_i(x) := h_i \operatorname{E} K_i^2\big(x - X(\tau_i)\big)$, que en virtud del apartado ii del Lema 1 converge a $f(x)D_k$ cuando i tiende a ∞ dado que f es, por hipótesis, contínua.

Entonces:

$$\text{(43)} \qquad \operatorname{E}\left(\left|\mathcal{S}_n^2\right|\right)^2 \leq \sum_{j=0}^{k-1} \operatorname{Var} \pi_j + 2 \sum_{i=0}^{k-1} \sum_{\substack{j=0 \\ i\neq j}}^{k-1} \left|\operatorname{Cov}\{\pi_i, \pi_j\}\right| = A + B$$

donde

$$\text{(44)} \qquad \operatorname{Var} \pi_j = \sum_{i=1+p}^{p+q} \operatorname{Var} Y_{m_j+i} + \sum_{r=p+1}^{p+q} \sum_{\substack{\ell=p+1 \\ r<\ell}}^{p+q} \left|\operatorname{Cov}\left\{Y_{m_j+r}, Y_{m_j+\ell}\right\}\right|$$

Si $2\eta - 1 > 0$, entonces por (42) y el carácter decreciente de h_n se tiene:

$$\text{(45)} \quad \frac{1}{h_n^{2\eta-1}} \sum_{i=1+p}^{p+q} \operatorname{Var} Y_{m_j+i} \leq \sum_{i=1+p}^{p+q} \left(\frac{h_1}{h_n}\right)^{2\eta-1} \Lambda_{m_j+i}(x) \leq \frac{q_n}{h_n^{2\eta-1}} \operatorname{Cte.}_1(x)$$

Si $2\eta - 1 < 0$, entonces $\left(\dfrac{h_i}{h_n}\right)^{2\eta-1} < 1$, y la cota sería $q_n \operatorname{Cte.}_2(x)$.

En virtud del Lema de Doe, si γ es el del enunciado y denotando por $\|Y_i\|$ a $\|Y_i\| = \left(\operatorname{E}\left\{|Y_i|^{2/\gamma}\right\}\right)^{\gamma/2}$, se tiene:

$$\frac{1}{h_n^{2\eta-1}} \sum_{r=p+1}^{p+q} \sum_{\substack{\ell=p+1 \\ r<\ell}}^{p+q} \left|\operatorname{Cov}\left\{Y_{m_j+r}, Y_{m_j+\ell}\right\}\right| \leq$$

$$\leq \frac{10}{h_n^{2\eta-1}} \sum_{r=p+1}^{p+q} \sum_{\substack{\ell=p+1 \\ r<\ell}}^{p+q} \left\|Y_{m_j+r}\right\| \left\|Y_{m_j+\ell}\right\| \left[\alpha\left(\tau_s - \tau_\ell\right)\right]^{1-\gamma}$$

Pero por (41) y ser h_n decreciente:

$$\begin{aligned} \|Y_i\| \leq h_i^{\eta} h_i^{-1+\gamma/2} \Phi_i^{\gamma/2}(x) \leq h_1^{\eta} h_n^{-1+\gamma/2} \Phi_i^{\gamma/2}(x) \Rightarrow \\ \text{Máx.}\left\{\|Y_i\|;\quad 1 \leq i \leq n\right\} \leq h_n^{-1+\gamma/2} \operatorname{Cte}_3(x) \end{aligned}$$

Así pues:

$$\frac{1}{h_n^{2\eta-1}} \sum_{\substack{r=p+1 \\ r<\ell}}^{p+q} \sum_{\ell=p+1}^{p+q} \left|\text{Cov}\left\{Y_{m_j+r}, Y_{m_j+\ell}\right\}\right| \leq$$

$$\leq \text{Cte}_4(x) h_n^{\gamma-1-2\eta} \sum_{\substack{r=p+1 \\ r<\ell}}^{p+q} \sum_{\ell=p+1}^{p+q} \left[\alpha\left(\tau_s - \tau_\ell\right)\right]^{1-\gamma} \leq$$

$$\leq \text{Cte}_4(x) h_n^{\gamma-1-2\eta} \sum_{r=1}^{q-1} (q_n - r) \left[\alpha\left(\frac{r-1}{\beta}\right)\right]^{1-\gamma} \leq$$

$$\leq \text{Cte}_4(x) h_n^{\gamma-1-2\eta} q_n \sum_{r=1}^{\infty} \left[\alpha\left(\frac{r}{\beta}\right)\right]^{1-\gamma}$$

donde se ha de notar que el último factor es necesariamente finito por (38).

Tasladando esta cota y la obtenida en (45) a (44):

$$(46) \quad \frac{1}{nh_n^{2\eta-1}} B = \frac{1}{nh_n^{2\eta-1}} \sum_{j=0}^{k-1} \text{Var}\, \pi_j \leq \text{Cte}_1(x) \frac{k_n q_n}{nh_n^{2\eta-1}} + \text{Cte}_5(x) \frac{k_n q_n}{nh_n^{2\eta-\gamma-1}}$$

Y la convergencia de (46) a cero se sigue de (35) y (36).

De forma similar se procede a acotar $\frac{1}{nh_n^{2\eta-1}} B$. En efecto:

$$B \leq 2 \sum_{\substack{i=0 \\ i>j}}^{k-1} \sum_{j=0}^{k-1} \sum_{s=p+1}^{p+q} \sum_{t=p+1}^{p+q} \left|\text{Cov}\left\{Y_{m_i+s}, Y_{m_j+t}\right\}\right|$$

Como $i > j$, los índices $(m_i + s)$ y $(m_j + t)$ difieren cuando menos en p unidades, por tanto, volviendo a aplicar el Lema de Doe y dado que $\frac{q_n}{p_n} \longrightarrow 0$:

$$(47) \quad \frac{1}{nh_n^{2\eta-1}} B \leq 4 \sum_{s=1}^{n-p} \sum_{t=s+p}^{n} \left|\text{Cov}\left\{Y_s, Y_t\right\}\right| \leq \text{Cte}_6(x) h_n^{\gamma-1-2\eta} \sum_{i=q_n}^{\infty} \alpha\left(\frac{i}{\beta}\right)^{1-\gamma}$$

Y (47) converge a cero por (37), lo que unido a la convergencia a cero de (46) y sustituido en (43) prueba que: $\frac{1}{nh_n^{2\eta-1}} \text{E}\left(\left|S_n^2\right|\right)^2 \longrightarrow 0.$

Finalmente:

$$\frac{1}{nh_n^{2\eta-1}} \mathrm{E}\left(|\mathcal{S}_n^3|\right)^2 = \frac{1}{nh_n^{2\eta-1}} \left(\sum_{i=m_k+1}^{n} \mathrm{Var}\, Y_i + 2 \sum_{\substack{\ell=m+1 \\ \ell<s}}^{n} \sum_{s=m_k+1}^{n} \mathrm{Cov}\left\{Y_\ell, y_s\right\} \right)$$

Idénticos argumentos a los establecidos para acotar A y B permiten obtener:

$$\frac{1}{nh_n^{2\eta-1}} \mathrm{E}\left(|\mathcal{S}_n^3|\right)^2 \leq \frac{\mathrm{Cte}_7(x)}{n} \left\{ |\Delta_{k_n}^3| + \frac{|\Delta_{k_n}^3|}{h_n^{2\eta+1-\gamma}} \right\}$$

siendo

$$|\Delta_{k_n}^3| = n - k_n(p_n + q_n) = n - \left[\frac{n}{p_n + q_n}\right](p_n + q_n) \leq$$

$$\leq n - \left(\frac{n}{p_n + q_n} - 1\right)(p_n + q_n) = (p_n + q_n)$$

por tanto:

$$\frac{1}{nh_n^{2\eta-1}} \mathrm{E}\left(|\mathcal{S}_n^3|\right)^2 \leq \frac{\mathrm{Cte}_8(x)}{n} \left\{ 1 + \frac{1}{h_n^{2\eta+1-\gamma}} \right) (p_n + q_n) \longrightarrow 0 \quad \text{cuando } n \to \infty$$

en virtud de (36).

PASO 3: CUANDO $n \longrightarrow \infty$ SE TIENE QUE:

$$\left| \mathrm{E}\left(e^{iu\mathcal{S}_n^1}\right) - \prod_{j=0}^{k_n-1} \left(\mathrm{E}\, e^{iu\beta_j}\right) \right| \longrightarrow 0 \tag{48}$$

Para demostrar (48), y equivalentemente la independencia asintótica de los bloques grandes, se utiliza el siguiente resultado, que es una extensión del Teorema de Volkonskii-Rozanov (1959): "Sea $V_j = F_j(B_{j_G})$ $j = 1, \ldots, N$, y $|F_j(x)| \leq 1$. Y sea H la $\sigma-$álgebra generada por el proceso de los instantes de muestreo $\{\tau_k : k = 1, 2, \ldots\}$. Y sea X un proceso fuertemente mixing con coeficientes $\alpha(t)$, entonces se verifica:

$$\left| \mathrm{E}\left(\prod_{i=1}^{N} V_i\right) - \left(\prod_{i=1}^{N} \mathrm{E}\, V_i\right) \right| = \left| \mathrm{E}\left\{ \mathrm{E}\left(\prod_{i=1}^{N} V_i | H\right) - \left(\prod_{i=1}^{N} \mathrm{E}\,(V_i|H)\right) \right\} \right| \leq$$

(49) $\leq 4(N-1)\alpha(q_n/\beta)$"

Haciendo $F_j(x) = e^{-itx}$ en el Lema anterior se obtiene:

$$\left| \mathrm{E}\left(e^{iu\mathcal{S}_n^1}\right) - \prod_{j=0}^{k_n-1}\left(\mathrm{E}\, e^{iu\beta_j}\right)\right| \leq 4(k_n+1)\alpha\left(\frac{q_n}{\beta}\right) \simeq 4\frac{n}{p_n+q_n}\alpha\left(\frac{q_n}{\beta}\right) \simeq$$

$$\simeq 4\frac{n}{p_n}\alpha\left(\frac{q_n}{\beta}\right)$$

Y ahora (38) permite probar (48).

PASO 4: SE PRUEBA:

$$\frac{1}{nh_n^{2\eta-1}}\sum_{j=0}^{k-1}\mathrm{E}\,\beta_j^2 \longrightarrow \theta_\eta^2\sigma_f^2 \tag{50}$$

Procediendo como en (44) para la $\mathrm{Var}\,\pi_j$ se demuestra que el segundo sumando de:

$$\mathrm{E}\,\beta_j^2 = \sum_{r=1}^{p}\mathrm{Var}\,Y_{m_j+r} + 2\sum_{i=1}^{p}\sum_{\substack{\ell=1\\ i<\ell}}^{p}\left|\mathrm{Cov}\,\left\{Y_{m_j+i}, Y_{m_j+\ell}\right\}\right|$$

Por (26), (29) y el Lema 1 (ii):

$$\frac{1}{nh_n^{2\eta-1}}\,\mathrm{E}\,\left|\mathcal{S}_n^1\right|^2 \longrightarrow \theta_\eta^2\sigma_f^2$$

de ambos hechos y del paso 2 de la demostración se deduce (50).

PASO 5: $\mathcal{S}_n^1$ VERIFICA LA HIPÓTESIS CLÁSICA DE LINDEBERG-FELLER

Según (48) y (50), los sumandos β_j en $\mathcal{S}_n^1$ son asintóticamente independientes y tales que la suma normalizada de sus varianzas es $\theta_\eta^2\sigma_f^2$. Por ello la condición de Lindeberg-Feller para la normalidad asintótica de $\mathcal{S}_n^1$ toma la forma: para $\epsilon > 0$,

$$g_n(\epsilon) = \frac{1}{nh_n^{2\eta-1}\sigma_f^2\theta_\eta^2}\sum_{j=0}^{k-1}\mathrm{E}\,\left(\beta_j^2 I_{\left\{|\beta_j|\geq\epsilon\sigma_f\theta_\eta\left(nh_n^{2\eta-1}\right)^{1/2}\right\}}\right) \longrightarrow 0 \text{ si } n\to\infty \tag{51}$$

Por el Lema 1, $\mathrm{E}\,\left(K_i(x - X(\tau_i))\right)$ es convergente y dado que h_i^η decrece a cero y K es acotado por hipótesis:

$$|Y_i| \leq \text{Cte.}\left(h_i^{1-\eta} + h_i^{\eta}\right) \leq \text{Cte.}\, h_i^{1-\eta} \leq \text{Cte.}\, h_n^{1-\eta} \quad \forall i \leq n \Rightarrow$$

(52) $$\Rightarrow \text{Máx}\,\{|\beta_j|,\quad 0 \leq j \leq k_n - 1\} \leq \text{Cte.}\, p_n h_n^{1-\eta}$$

Por tanto en (51) se tiene:

$$g_n(\epsilon) = \frac{k_n p_n^2 h_n^{2\eta-2}}{n h_n^{2\eta-1} \sigma_f^2 \theta_\eta^2} \max_{0 \leq j \leq k_n - 1} P\left\{|\beta_j| \geq \epsilon \sigma_f \theta_\eta \left(n h_n^{2\eta-1}\right)^{1/2}\right\} =$$

$$g_n(\epsilon) = \text{Cte} \cdot k \left(\frac{p_n}{(n h_n)^{1/2}}\right) \max_{0 \leq j \leq k_n - 1} P\left\{|\beta_j| \geq \epsilon \sigma_f \theta_\eta \left(n h_n^{2\eta-1}\right)^{1/2}\right\}$$

Ahora bien, de (52)

$$\frac{\max\limits_{0 \leq j \leq k_n - 1} |\beta_j|}{\sigma_f \theta_\eta \left(n h_n^{2\eta-1}\right)^{1/2}} \leq \text{Cte.}\, \frac{p_n h_n^{\eta-1}}{\left(n h_n^{2\eta-1}\right)^{1/2}} = \text{Cte.} \frac{p_n}{(n h_n)^{1/2}} = \frac{1}{r_n} \longrightarrow 0$$

Y, en consecuencia, el $\max\limits_{0 \leq j \leq k_n - 1} P\left\{|\beta_j| \geq \epsilon \sigma_f \theta_\eta \left(n h_n^{2\eta-1}\right)^{1/2}\right\} = 0$ a partir de un n suficientemente grande, deduciéndose entonces (51).

Resumiendo:

$$\left.\begin{array}{l} \text{Por (50) y (51), } \left(n h_n^{2\eta-1}\right)^{-(1/2)} \mathcal{S}_n^1 \qquad \longrightarrow \text{N}\,(0, \theta_\eta \sigma_f) \\ \text{y por el paso 2, } \left(n h_n^{2\eta-1}\right)^{-(1/2)} \left(\mathcal{S}_n^2 + \mathcal{S}_n^3\right) \longrightarrow 0 \text{ en probabilidad} \end{array}\right\} \Rightarrow$$

$$\left(n h_n^{2\eta-1}\right)^{-(1/2)} \mathcal{S}_n \xrightarrow{d} \text{N}(0, \theta_\eta \sigma_f)$$

■

4. BIBLIOGRAFÍA

[1] **Blum-Boyles** (1981). "Random sampling from a continous parameter stochastic process". *Analytical Methods in Probability Theory, Lecture Notes in Mathematics,* **86**, Springer-Verlag.

[2] **Bradley, R.** (1981). "Central Limit Theorems under Weak Dependence". *Journal of Multivariate Analysis,* **11**, 1–16.

[3] **Deheuvels** (1974). "Conditions nécessaires et suffisantes de convergence ponctuelle presque sûre". *C.R. Acad. Sci. Paris Ser. A,* **278**, 1217–20.

[4] **Doe** (1973). "A note on empirical processes for strong mixing processes". *Ann. Prob.,* **1**, 870–875.

[5] **Gyorfi-Hardle-Sarda-Vieu** (1989). "Nonparametric curve estimation from time series". *Lecture Notes in Statistics,* **60**.

[6] **Marron, J. & Hardle, W.** (1986). "Random aproximations to some measures of accuracy in nonparametric curve estimation". *Journal of Multivariate Analysis,* **20**, 91–113.

[7] **Masry, E.** (1983). "Probability Density Estimation from Sampled Data". *IEEE, vol. IT-29,* **5**, 696–709.

[8] **Masry, E.** (1986). "Recursive Probability Density Estimation for Weakly Dependent Stationary Processes". *IEEE, vol. IT-32,* **2**, 254–267.

[9] **Peligrad, M.** (1985). "Recent advances in the central limit theorem and its weak invariance principle for mixing sequences of random variables (A Survey)". En *Dependence in Probability and Statistics.* E. Eberlein, M. Taqqu ed. Birkhauser.

[10] **Rosenblatt** (1970). "Density estimates and Markov Sequences". En *Nonparametric Techniques in Statistical Inference.* Cambridge University Press, 199–210.

[11] **Robinson** (1983). "Nonparametric estimators for time series". *J. Time Series Anal.,* **40**, 185–207.

[12] **Silverman, B.W.** (1986). *Density Estimation for Statistics and Data Analysis.* Chapman and Hall.

[13] **Stoyanov-Robinson** (1991). "Semiparametric and nonparametric inference from irregular observations on continuous time stochastic processes". *Nonparametric Functional Estimation and related Topics,* 553–558, Kluwer Academic Publishers.

[14] **Volkonskii-Rozanov** (1959). "Some limit theorems for random functions I". *Theory Prob. Appl.,* **4**, 178–197.

[15] **Wertz** (1985). "Sequential and Recursive Estimators of the Probability Density". *Statistics, 16,* **2**, 277–295.

[16] **Wolverton-Wagner** (1969). "Recursive estimates of probability densit". *IEEE Trans. Systems Sci. Cybernet,* **5**, 246–247.

ENGLISH SUMMARY:

ESTIMATION OF THE PROBABILITY DENSITY FROM RANDOM SAMPLING

José A. Vilar and Juan M. Vilar

1. INTRODUCTION. DEFINITIONS

Data are not always collected at equally spaced time intervals and there are various causes by which data may be irregularly recorded over time: small random deviations, the own random experimental design...

In this paper we study the asymptotic behaviour of a recursive nonparametric estimate of the probability density function associated with a stationary continuous-time process X, where the sampling instants are assumed to be random.

1.1. The model

Let the continuous-discrete stochastic model be defined as the pair (X, T) where:

1. The first component $X = \{X(t);\, t \in \mathbb{R}\}$ is a continuous time process, which is assumed to be strictly stationary asn strongly mixing ($\alpha-$mixing), with marginal density function $f(x)$ continuous and bounded.
2. The second component, $T = \{\tau_k;\, k \in \mathbb{N}\}$, consists of a strictly increasing sequence of random times that we confine to two specific structures:

- **"IO" STRUCTURE (Irregularly Observations)** where:

$$\tau_k = \frac{k}{\beta} + z_k \qquad k = 0, \pm 1, \dots \quad \beta > 0$$

being z_k's i.i.d. random variables with a symmetric density function $g(t)$ supported on $[\frac{-1}{2\beta}, \frac{1}{2\beta}]$.

- "**RP STRUCTURE (Renewal Process)** where:

$$\tau_k = \sum_{i=1}^{k} \quad k = 1, 2, \ldots \quad \text{and} \quad \tau_0 \equiv 0$$

where $\{t_i\}_{i=1}^{+\infty}$ is a sequence of i.i.d. random variables with a common absolutely continuous distribution $G(t)$ on $[0, \infty)$, satisfying: $\mathrm{E}[t_i] = \beta^{-1}$.

Both structures turn to a progressive randomness of the sampling data: the IO structure is the mixture of a periodic sampling and a random deviation, and the RP structure is completely random.

Throughout this paper we also suppose that X and T are independent.

1.2. Definition of the estimate

The nonparametric estimation of an unknown probability density function $f(x)$ has received much attention in recent years (see Silverman (1986) for independent observations and Gyorfy e.a. (1989) for dependent observations).

Masry (1983) has studied this same stochastic model for the nonrecursive type kernel estimator of the density function defined by (1):

$$f_n(x) = \frac{1}{n}\sum_{j=1}^{n} K_n(x - X_j), \quad K_n(u) = \frac{1}{h_n}K\left(\frac{u}{h_n}\right)$$

being $K(u)$ the Kernel function and h_n a bandwidth sequence.

But the recursive estimator has a clear advantage over nonrecursive estimator as they can be updated with each additional observation. Thus we will consider the recursive type Kernel estimate defined by (2):

$$\hat{f}_n(x) = H_n^{-1}\sum_{j=1}^{n} K_j(x - X(\tau_j)), \quad \text{with } \eta \in \mathbb{R} \quad \text{and } H_n = \left(\sum_{j=1}^{n} h_j^{\eta}\right)$$

This estimator, which one was introduced by Deheuvels (1974), can be computed recursively by (3). From equally spaced data, Wolverton-Wagner (1969) have used (2) with $\eta = 0$ for independent observations, and this same estimator has been considered by Masry (1986) under the assumption of mixing conditions.

2. BIAS AND VARIANCE

Under weak regularity conditions on K and $\{h_n\}$, the theorem 1 related to bias of $\hat{f}_n$ was established.

We remark that the asymptotic bias expression (4) obtained in theorem 1 depends on Kernel function K, on the underlying probability density f and their derivatives, and on parameter η. Specifically, large values of η result in estimations with a large bias. But the bias of the estimation does not depend neither on the particular sampling scheme used nor on the dependence structure type.

Finally, comparing this expression to that in nonrecursive case it is seen that the recursive estimate has a larger bias than a nonrecursive one.

In theorems 2 and 3 we establish bounds for the asymptotic covariance of $\hat{f}_n(x)$ when $T \equiv$ IO and $T \equiv$ RP respectively.

If $T \equiv$ IO, expression (5) is obtained under a weak constraint on the mixing coefficients $\{\alpha_t\}$ (see (H4), and it is interesting to remark some observations about this result:

- The sampling scheme (according to IO structure) contributes to an increasing of the covariance, which is proportional to β.
- The convergence rate of the bound is $O\big(nh_n^{1+q}\big)^{-1}$ compared to $O\big(nh_n\big)^{-1}$ for regular observations.
- But, under the considerably more restrictive assumptions H4 with $0 < q < \frac{1}{2}$, and H5, the new bound for the asymptotic covariance (see(6)) is independent of β and the corresponding rate of convergence is now $O\big(nh_n\big)^{-1}$.
- Finally, under condition H6, the contribution of the sampling scheme is asymptotically negligible.
- These results are extensions to the recursive case of results by Masry (1983) for nonrecursive case. In this sense, it is interesting to note that nonrecursive estimate has a larger variance than a recursive one (see remark 6).

The performance of the estimate $\hat{f}_n(x)$ when $T \equiv$ RP is very similar such as it is shown in teorem 3. In this case the influence of random sampling is present through the renewal density function $h(t)$, which is supposed to be bounded (see (9)). Analogous conclusions to those in theorem 2 are now made for theorem 3.

Combining the results of theorem 1 on the bias of the estimator $\hat{f}_n(x)$ and of theorem 2 (for IO structure) or theorem 3 (for RP structure) on the variance, the quadratic-mean consistency and the corresponding rate of $\hat{f}_n(x)$ is immediately obtained in corollaries 1 and 2 respectively.

3. ASYMPTOTIC NORMALITY

In this section, asymptotic normality of the recursive estimate defined by (2) is established.

In first place we write:

$$(nh_n)^{\frac{1}{2}}\left(\hat{f}_n(x)-f(x)\right)=(nh_n)^{\frac{1}{2}}\left(\hat{f}_n(x)-\mathrm{E}\,\hat{f}_n(x)\right)+(nh_n)^{\frac{1}{2}}\left(\mathrm{E}\,\hat{f}_n(x)-f(x)\right)$$

Following Pelligrad (1985) *(Bernstein method)* the asymptotic normality of $\hat{f}_n(x)$ centered at $\mathrm{E}\,\hat{f}_n(x)$ is stated in theorem 4 under H9 and H10.

Finally the sketch of the proofs is given in the last section of the paper.