

ESTIMACIÓN DE LA FUNCIÓN DE DISTRIBUCIÓN CONDICIONAL CON CENSURA Y TRUNCAMIENTO

M^a del Carmen Iglesias Pérez

16 de noviembre de 2001

Índice General

| | |
|---|----|
| Introducción | 5 |
| 1 Introducción a los modelos de censura, truncamiento y censura con truncamiento. Estimación de la función de distribución. | 11 |
| 1.1 El modelo de censura aleatoria por la derecha. | 11 |
| 1.2 Estimación de la función de distribución con censura. | 15 |
| 1.2.1 El estimador de Kaplan y Meier: definición. | 17 |
| 1.2.2 El estimador de Kaplan y Meier: propiedades. | 21 |
| 1.3 El modelo de truncamiento aleatorio por la izquierda. | 28 |
| 1.4 Estimación de la función de distribución con truncamiento. | 31 |
| 1.4.1 El estimador de Linden-Bell: definición. | 32 |
| 1.4.2 El estimador de Linden-Bell: propiedades. | 34 |
| 1.5 El modelo de truncamiento aleatorio por la izquierda y censura aleatoria por la derecha. | 39 |
| 1.6 Estimación de la función de distribución con censura y truncamiento. | 42 |
| 1.6.1 El estimador de Turnbull: definición. | 42 |
| 1.6.2 El estimador de Turnbull: propiedades. | 44 |
| 2 Estimación no paramétrica de la función de distribución condicional en presencia de censura, truncamiento y censura con truncamiento. | 51 |
| 2.1 Estimación no paramétrica de la función de distribución condicional. | 52 |
| 2.2 Estimación no paramétrica de la función de distribución condicional con censura. | 60 |
| 2.3 Estimación no paramétrica de la función de distribución condicional con censura y truncamiento. | 69 |
| 2.3.1 Definición del estimador. | 71 |
| 2.3.2 Propiedades del estimador: características generales. | 74 |
| 2.3.3 Propiedades del estimador: aspectos asintóticos. | 86 |

| | | |
|-------|---|-----|
| 2.3.4 | Lemas y demostraciones de los principales resultados. | 90 |
| 3 | Bootstrap para la función de distribución condicional en presencia de censura y truncamiento. | 135 |
| 3.1 | Antecedentes. | 135 |
| 3.1.1 | El método de remuestreo Bootstrap. | 135 |
| 3.1.2 | El método Bootstrap para datos con censura. | 136 |
| 3.1.3 | El método Bootstrap para datos con censura y truncamiento. | 138 |
| 3.2 | Bootstrap condicional con censura y truncamiento. | 140 |
| 3.2.1 | Definición de una aproximación bootstrap. | 140 |
| 3.2.2 | Propiedades de la aproximación bootstrap. | 142 |
| 3.2.3 | Lemas y demostraciones. | 144 |
| 4 | Aplicación práctica. | 189 |
| 4.1 | Selección de la ventana. | 192 |
| 4.2 | Análisis condicionales. | 195 |
| | Bibliografía | 207 |

Introducción

Una de las partes más importantes de la Estadística Matemática es la Inferencia Estadística cuyo objetivo es el estudio de características de una población partiendo de la información contenida en una muestra de la misma. La forma de abordar los problemas de inferencia da lugar a dos tipos de contexto: el paramétrico y el no paramétrico. Los métodos paramétricos suponen que los datos provienen de una distribución conocida (normal, exponencial,...) y tratan de estimar los parámetros que definen esa distribución. Los métodos no paramétricos tratan de establecer el modelo probabilístico que ha generado los datos introduciendo hipótesis muy generales sobre el modelo. Así, pues, la inferencia paramétrica resolverá problemas de estimación y contraste de parámetros (media, varianza, parámetros asociados a una distribución específica,...) mientras que en inferencia no paramétrica aparecerá el problema de la estimación de curvas (función de distribución, función de densidad,...).

En múltiples ocasiones interesa explicar la relación existente entre una variable respuesta Y y una variable explicativa o covariable X . En este contexto, uno puede plantearse estimar la distribución conjunta, la curva de regresión o la función de distribución de Y condicionada a X , entre otras características relevantes. Al igual que la estimación de la distribución de una v.a., la estimación de las funciones anteriores puede llevarse a cabo desde dos ópticas distintas: la paramétrica y la no paramétrica.

Esta memoria tiene que ver con la estimación no paramétrica de curvas. En particular se abordarán cuestiones relacionadas con estimadores de la función de distribución sin y con covariables en los contextos específicos de censura y truncamiento.

El campo de la estimación no paramétrica de curvas ha sufrido un importante desarrollo, sobre todo en los últimos decenios (ver por ejemplo, Eubank (1988) y Härdle (1990)) y hoy en día existe toda una gama de técnicas disponibles: el método núcleo, el método de los k puntos próximos, la estimación spline y la estimación por polinomios locales pueden considerarse como las más notables. Entre estos

procedimientos hemos elegido el tipo núcleo, por ser uno de los más utilizados y estudiados.

El método núcleo presenta la ventaja de la flexibilidad: se adapta a la estimación de funciones de distribución, densidad o regresión, al diseño ...jo y al aleatorio, y además, presenta una formulación algebraica relativamente fácil de tratar teóricamente. Entre los tipos de pesos núcleo que aparecen en la literatura, utilizaremos los propuestos por Nadaraya (1964) y Watson (1964) por ser los más apropiados en el contexto de diseño aleatorio (truncamiento implica diseño aleatorio) y por disponer de resultados teóricos de interesante aplicación.

Concretamente, dada una m.a.s. $f(X_i; Y_i)_{i=1}^n$; el estimador tipo núcleo con pesos de Nadaraya-Watson de la función de regresión $m(x) = E(Y | X = x)$ viene dado por:

$$\hat{m}_n(x) = \frac{\sum_{i=1}^n K\left(\frac{x_i - x}{h}\right) Y_i}{\sum_{j=1}^n K\left(\frac{x_i - x_j}{h}\right)} = \sum_{i=1}^n B_{ni}(x) Y_i$$

donde K es una función núcleo (función real, continua, acotada, simétrica y que integra 1) y h el conocido parámetro de suavización de la estimación no paramétrica.

Intuitivamente, $\hat{m}_n(x)$ no es más que un promedio local, es decir, $m(x)$ se estima por una media ponderada de valores Y_i ; donde las ponderaciones, $B_{ni}(x)$; tienen en cuenta la distancia entre los valores de X_i y x , y el parámetro de suavizado, h ; es el encargado de regular el tamaño de la vecindad de x donde se pondera.

Obviamente, el estimador de Nadaraya-Watson de la función de distribución condicional $F(y | x) = P(Y \leq y | X = x)$ vendrá dado por:

$$\hat{F}_n(x) = \frac{\sum_{i=1}^n K\left(\frac{x_i - x}{h}\right) 1_{F_{Y_i} \leq y}}{\sum_{j=1}^n K\left(\frac{x_i - x_j}{h}\right)} = \sum_{i=1}^n B_{ni}(x) 1_{F_{X_i} \leq y}$$

sin más que tener en cuenta que $F(y | x) = E(1_{F_{Y_i} \leq y} | X = x)$:

La construcción de modelos que resuelvan los problemas de estimación de curvas anteriormente citados, necesita conocer una muestra aleatoria simple de la población de interés: $f(X_i)_{i=1}^n$ para la estimación de la distribución; $f(X_i; Y_i)_{i=1}^n$ para estimar la función de distribución condicional o la curva de regresión. Sin embargo, esta condición, que en principio parecería la mínima necesaria para construir un buen modelo, no siempre se puede conseguir. La censura y el truncamiento denotan dos situaciones distintas donde no se tiene una m.a.s. de las variables a estudiar. En ambos contextos, que pueden aparecer a la vez o por separado, hay una pérdida

de información de las variables de interés y es necesario estudiar nuevas técnicas encaminadas a extraer lo mejor posible la información, que sobre la población inicial, contienen los datos observados.

La censura modeliza situaciones en la cuales o se observa la variable de interés o se observa otra variable, que se denomina censura, y que de algún modo impide ver la realización de la primera. Más concretamente, si nos centramos en la censura aleatoria por la derecha, observaremos la variable $(Z; \pm)$ donde $Z = \min\{Y; C\}$; $\pm = 1$ si $Y < C$ y $\pm = 0$ si $Y \geq C$ siendo Y la variable de interés y C la variable censura. La variable indicadora \pm señala que en cada momento conocemos el tipo de variable que observamos y precisamente Z coincidirá con la de menor valor. Supongamos, como ejemplo, que Y representa el tiempo de vida de un enfermo de cancer y que en la muestra de estudio hay varios casos en que éste no ha podido observarse totalmente: debido a un cambio de domicilio, o porque ha ocurrido un fallecimiento debido a causas ajenas a la enfermedad, por abandono del individuo en estudio, etc. Esta situación se adapta perfectamente a la descrita de censura puesto que en ocasiones no se dispone de la variable de interés debido a que hay otra variable, la censura, que la precede impidiendo su observación. Otros ejemplos de esta situación aparecen en el campo de la Ingeniería (duración de una pieza electrónica, tiempo de reparación de un sistema,...), de la Economía (permanencia de un individuo en paro, duración de huelgas), etc.

Otra de las situaciones donde hay una pérdida de información en los datos y que, al igual que la censura, es objeto de estudio en este trabajo, es el truncamiento. El problema de los datos truncados aparece cuando la variable de interés no se observa siempre (caso de una m.a.s.) sino sólo cuando se cumple determinada condición, y en este caso, además de observar la variable en estudio, obtenemos el valor de otra variable involucrada en la condición de observación y que se denomina truncamiento. Para explicar conceptos, centrémonos en el truncamiento aleatorio por la izquierda. Si Y denota la variable de interés y T la variable de truncamiento observaremos pares $(Y_i; T_i)$ sólo cuando $T_i \leq Y_i$, en otro caso no habrá observación. Supongamos, como ejemplo, que Y es el tiempo de duración de una determinada enfermedad y T mide el tiempo desde que la enfermedad comienza en un individuo hasta que éste entra en estudio. En cierto tipo de enfermedades el especialista tendrá datos correspondientes a enfermedades de una cierta duración mientras que perderá los datos relativos a periodos cortos o asintomáticos, ya que en estos casos los pacientes no acudirán a consulta. Cualquier estudio que no tenga en cuenta esta situación proporcionará resultados sesgados.

Esta memoria pretende dar una visión de la situación actual respecto a la estimación no paramétrica de la función de distribución para datos censurados y/o truncados, tanto en ausencia como en presencia de covariables. Pretende también,

aportar nuevos resultados al problema anteriormente mencionado. A este ...n, dedica sus cuatro capítulos, que se organizan de la forma siguiente:

- ² En el primer capítulo se presentan los modelos de censura, truncamiento y de censura con truncamiento con los que trabajaremos, definiendo las funciones más importantes que aparecen ligadas a cada uno de ellos, así como las condiciones para su identificabilidad. Para cada modelo, se hace una revisión del llamado estimador límite producto, que resulta ser el estimador no paramétrico de máxima verosimilitud de la función de distribución, exponiendo sus propiedades más relevantes: consistencia, normalidad asintótica, representaciones casi seguras, etc.
- ² El segundo capítulo plantea un análisis similar al realizado en el capítulo anterior, pero en presencia de covariables. De este modo, se aborda el problema de la estimación de la función de distribución condicional. Una vez expuestos los resultados ya existentes, se presenta un nuevo estimador para dicha función de distribución condicional en el contexto de censura y truncamiento. Para este nuevo estimador se obtienen propiedades de interés: una representación casi segura, la consistencia fuerte uniforme, su distribución asintótica y la convergencia débil del proceso.
- ² En el tercer capítulo se propone un método de remuestreo bootstrap para datos censurados y truncados con covariables, con el fin de obtener una versión bootstrap del estimador definido en el capítulo anterior. Además, se deriva una representación casi segura bootstrap, que es la llave para demostrar la consistencia del método de remuestreo propuesto. La aproximación bootstrap obtenida, representa una alternativa a la distribución normal asintótica (capítulo dos) y evita la estimación de los complicados parámetros que definen dicha distribución.
- ² El cuarto y último capítulo de este trabajo ilustra el comportamiento del estimador anteriormente estudiado para la función de distribución condicional con censura y truncamiento, en un problema con datos reales. Es un capítulo interesante en cuanto que muestra como el estimador condicional proporciona una información importante sobre el problema en cuestión (relativo al análisis sobre la mortalidad en pacientes diabéticos), información que en un análisis incondicional permanece oculta. Sin embargo, debemos poner de manifiesto algunos problemas prácticos que han quedado abiertos y que tenemos intención de abordar en breve. El primero y más importante de ellos es la estimación de un parámetro ventana óptimo en este contexto. Hemos adaptado un criterio de tipo validación cruzada introducido por Bowman, Hall

y Prvan (1998) para la estimación tipo núcleo de funciones de distribución, pero es necesario estudiar a fondo su comportamiento para el nuevo estimador introducido en este trabajo. También somos conscientes de la necesidad de un estudio de simulación que muestre el comportamiento del estimador de la función de distribución condicional con censura y truncamiento en distintas situaciones teóricas y prácticas.

Además de los problemas abiertos en el campo práctico, y sobre todo de la vía que supone el estudio de un parámetro de suavizado óptimo en este nuevo contexto, nos gustaría destacar la importancia del camino teórico recorrido, sobre todo de las representaciones casi seguras, como una herramienta importante para aplicar en nuevos campos de estudio: en el análisis de propiedades de estimadores de interés relacionados con la distribución condicional, en el planteamiento y análisis de contrastes de hipótesis para este contexto, como paso que facilita el estudio de cuestiones semejantes para otras características relevantes como pueden ser la función cuantil o la mediana, etc.

Otro tema atractivo es el de la aplicación práctica del método bootstrap condicional con censura y truncamiento, propuesto y validado teóricamente en el capítulo tres, que permite la aproximación de la distribución y el cálculo de bandas de confianza para características condicionales de interés, y que implica el estudio de métodos para el cálculo de la ventana piloto y necesaria en el remuestreo.

Capítulo 1

Introducción a los modelos de censura, truncamiento y censura con truncamiento. Estimación de la función de distribución.

1.1 El modelo de censura aleatoria por la derecha.

Como se ha puesto de mani...esto en la introducción de esta memoria, el término censura hace referencia a situaciones en las que hay una pérdida de información en la variable de interés, en el sentido de que en ocasiones no se observa esta variable sino otra variable, que se denomina censura, y que de algún modo impide ver la realización de la primera. Este problema aparece de forma habitual cuando se estudian variables de tiempos de vida (o de fallo) y por ello no es extraño que los primeros trabajos sobre censura aparezcan ligados al Análisis de Supervivencia.

El Análisis de Supervivencia se ocupa del estudio de los tiempos de fallo de un conjunto de individuos. Por tiempo de fallo se entiende el período transcurrido desde un instante inicial hasta un instante ...nal (fallo) que han de ser de...nidos previamente y de forma precisa por el investigador. Esta formulación general se adapta a multitud de problemas. Así, el tiempo de fallo puede medir el período transcurrido entre la entrada en estudio de un paciente con una enfermedad terminal y su muerte, o el tiempo de duración de una máquina eléctrica desde su instalación; también puede responder al tiempo transcurrido desde el tratamiento de un enfermo

con una determinada medicación hasta su alta, o la permanencia de un individuo en paro hasta la obtención de su primer empleo. Todos estos problemas tienen en común que el suceso fallo está perfectamente determinado de antemano y puede ocurrir, a lo sumo, una vez en cada individuo. A su vez, verifican que la variable de interés, el tiempo de fallo, es una variable no negativa.

Es fácil darse cuenta que en este contexto aparecen con frecuencia problemas de seguimiento que proporcionan muestras censuradas. Si el estudio que hemos diseñado termina antes del fallo de algunos individuos tendremos una información incompleta de su tiempo de fallo, pues sólo dispondremos del tiempo transcurrido hasta un instante inicial conocido (previo al fallo) y que se suele denominar censura. Lo mismo ocurrirá si un paciente terminal abandona el estudio o muere por causas ajenas a su enfermedad, o si una máquina sometida a un estudio de seguimiento es robada cuando todavía funciona.

La modelización de una situación de censura tiene en cuenta las siguientes variables:

- 1) La variable de interés (tiempo de fallo) que denotaremos por Y .
- 2) La variable censura, C , que mide el tiempo transcurrido desde el instante inicial hasta el instante de censura.
- 3) La variable observada, Z , que será $Z = \min\{Y, C\}$; puesto que se observa el tiempo de fallo siempre que sea inferior al tiempo de censura, en caso contrario se observa el tiempo de censura.
- 4) Una variable que indica cuando se observa un fallo y cuando una censura:

$$\delta = \begin{cases} 1 & \text{si } Z = Y \quad (\text{fallo}) \\ 0 & \text{si } Z = C \quad (\text{censura}) \end{cases}$$

Todas estas variables aparecen, implícitamente, en la muestra observada bajo censura, dada por

$$f(Z_1; \delta_1); (Z_2; \delta_2); \dots; (Z_n; \delta_n)g$$

y que es el punto de partida de cualquier estudio estadístico en este contexto.

5) En ocasiones interesa estudiar la relación de ciertas variables, X (covariables), con el tiempo de fallo. Generalmente estas variables se observan al comienzo del estudio de seguimiento y por ello los modelos más habituales suponen que la censura no afecta a las covariables. Es por esto, que en el caso de datos censurados con covariables observaremos una muestra del tipo

$$f(X_1; Z_1; \delta_1); (X_2; Z_2; \delta_2); \dots; (X_n; Z_n; \delta_n)g$$

Una cuestión importante a tener en cuenta es que la censura C puede considerarse como fija (conocida a priori por el investigador) o como una variable aleatoria

(su valor no está predeterminado de antemano), dando lugar a modelos de censura...ja y modelos de censura aleatoria, respectivamente.

También es interesante observar que hasta ahora hemos considerado un mecanismo de censura por la derecha. Análogamente se puede de...nir un modelo de censura por la izquierda: se observará una m.a.s. de la variable $(Z; \pm)$ donde $Z = \min fY; Cg$ y $\pm = 1_{fC} \cdot Yg$:

Uno de los modelos de censura más estudiados es el modelo de censura aleatoria por la derecha, y es el que utilizaremos a lo largo de toda esta memoria. Este modelo parte de una m.a.s. de $(Z; \pm)$ donde:

$$Z = \min fY; Cg; \pm = 1_{fY} \cdot Cg$$

y supone además que

$$Y \text{ y } C \text{ son v.a. independientes}$$

Por comodidad y ya que ello no supone pérdida alguna de generalidad, trabajaremos con variables no negativas.

Asociadas al mecanismo probabilístico de este modelo de censura se de...nen varias funciones importantes:

1) La función de distribución de la variable tiempo de fallo, $F(y) = P(Y \cdot y)$ y su función de supervivencia o ...abilidad, $S(y) = 1 - F(y) = P(Y > y)$:

2) La función de distribución de la variable de censura, $G(y) = P(C \cdot y)$ y su función de supervivencia o ...abilidad, $1 - G(y) = P(C > y)$:

3) La función de distribución de la variable observada Z , $H(y) = P(Z \cdot y)$ y su función de supervivencia o ...abilidad, $1 - H(y) = P(Z > y) = P(\min fY; Cg > y) = P(Y > y; C > y)$:

Bajo la hipótesis de que Y y C sean independientes se tiene

$$1 - H(y) = (1 - F(y))(1 - G(y)) \tag{1.1}$$

4) La subdistribución de la variable observada, Z , cuando es una observación no censurada: $H_1(y) = P(Z \cdot y; \pm = 1)$ y que se puede escribir de la siguiente forma

$$H_1(y) = P(Y \cdot y; Y \cdot C) = \int_0^y \int_t^{\infty} dG(u) dF(t) = \int_0^y \frac{dH_1(t)}{1 - G(t)} dF(t) \tag{1.2}$$

5) La subdistribución de la variable observada, Z , cuando es una observación censurada

$$H_0(y) = P(Z \cdot y; \pm = 0) = \int_0^y \int_t^{\infty} 1 - F(t) dG(t)$$

6) La función de azar o razón de fallo definida por

$$r(y) = \lim_{h \downarrow 0} \frac{P(y < Y \leq y + h | Y \geq y)}{h} = \frac{\frac{dF(y)}{dy}}{1 - F(y)}$$

y que representa el riesgo instantáneo de fallo de un individuo en el instante y ; sabiendo que ha vivido hasta ese instante.

7) La función razón de fallo acumulada, que se define a partir de la función anterior mediante la expresión:

$$\Lambda(y) = \int_0^y r(t) dt = \int_0^y \frac{dF(t)}{1 - F(t)}$$

La función Λ determina de forma única la función de distribución F mediante la relación

$$1 - F(y) = e^{-\Lambda_c(y)} \prod_{a_i \leq y} (1 - \Delta a_i) \quad (1.3)$$

$\Delta a_i = a_i - y_i$

donde Λ_c denota la parte continua de Λ , A el conjunto de puntos donde Λ tiene discontinuidades de salto y $\Delta a_i = \Lambda(a_i) - \Lambda(a_i^-)$ la magnitud del salto de Λ en a_i :

Esta propiedad, cuya demostración se puede ver en la página 301 de Shorack y Wellner (1986), jugará un papel fundamental en la construcción de estimadores de la función de distribución de la variable Y :

En el presente modelo, también es importante definir:

8) Los extremos del soporte para una función de distribución arbitraria W , dados por

$$a_W = \inf\{t : W(t) > 0\} \quad \text{y} \quad b_W = \sup\{t : W(t) < 1\} \quad (1.4)$$

Tiene interés considerar dichos extremos para las funciones F , G y H : a_F ; b_F ; a_G ; b_G ; a_H y b_H , respectivamente, y estudiar como se relacionan.

La expresión (1.1) implica que

$$H(t) > 0, \quad 1 - H(t) < 1, \\ 1 - F(t) < 1 \text{ o } 1 - G(t) < 1, \quad F(t) > 0 \text{ o } G(t) > 0$$

y

$$H(t) < 1, \quad 1 - H(t) > 0, \\ 1 - F(t) > 0 \text{ y } 1 - G(t) > 0, \quad F(t) < 1 \text{ y } G(t) < 1$$

y por tanto

$$\begin{aligned} a_H &= \min f a_F ; a_G g \\ b_H &= \min f b_F ; b_G g : \end{aligned} \quad (1.5)$$

1.2 Estimación de la función de distribución con censura.

Cuando se dispone de una muestra completa de datos $f Y_i g_{i=1}^n$; la función de distribución empírica es un estimador no paramétrico de la función de distribución poblacional, F ; con el criterio de máxima verosimilitud:

$$\max L_n = \max_{f p_1 ; p_2 ; \dots ; p_n 2 R = p_i > 0 \text{ y } p_i = 1 g} P_{p_1 \in p_2 \in \dots \in p_n}$$

Una forma habitual de presentar este estimador es

$$F_n(y) = \frac{\# f Y_i \cdot y g}{n} = \sum_{i=1}^n 1_{f Y_i \cdot y g} \frac{1}{n}$$

En términos de Análisis de Supervivencia diríamos que la probabilidad de fallo en el intervalo $I = (0; y]$ es decir, $F(y)$, es estimada por la proporción muestral de fallos en I , esto es, por el cociente entre el número de fallos ocurridos en I y el número de individuos a riesgo en el instante $t = 0$: Explícitamente, si para el intervalo $I = (0; y]$ de...nimos:

$n = n^0$ de individuos a riesgo en el instante $t = 0$

$m = n^0$ de fallos en I

entonces $F_n(y) = m/n$ y, de forma complementaria, la probabilidad de supervivencia en el intervalo $I = (0; y]$ se estima mediante $1 - F_n(y) = 1 - \frac{m}{n}$

A diferencia del contexto de datos completos, en presencia de datos censurados no observamos una muestra de tamaño n de la variable Y , sino la muestra $f(Z_i; \pm_i) g_{i=1}^n$; donde Z_i coincide con Y_i sólo en los casos $\pm_i = 1$: Si nos planteamos estimar $F(y)$ mediante la muestra $f(Z_i; \pm_i) = \pm_i = 1 g = f Y_k g_{k=1}^m$ (evidentemente $m \cdot n$), estaremos reduciendo la estimación a un contexto de datos completos con menor tamaño muestral, y, naturalmente, despreciando la información proporcionada por los datos censurados de que el fallo para un individuo censurado siempre ocurre después de su tiempo observado. Precisamente, el objetivo de los modelos para datos con censura está en aprovechar dicha información.

Nosotros nos centraremos en el problema de estimar la función de distribución en presencia de censura, es decir de buscar un estimador alternativo a la función de distribución empírica (de tamaño n), que cuando hay censura no se puede calcular.

El primer punto de referencia en este tema es, indiscutiblemente, el trabajo de Kaplan y Meier de 1958. En este artículo se estudian y comparan varios estimadores de la función de distribución para datos con censura: los estimadores de tipo actuarial, el estimador de muestra reducida y el estimador límite producto. Éste último, también llamado estimador de Kaplan-Meier, presenta bastantes propiedades deseables: es el estimador no paramétrico máximo-verosímil para datos censurados; si lo comparamos con el estimador de muestra reducida, sale ganando en cuanto a que se calcula sin necesidad de conocer los tiempos de censura para las observaciones no censuradas, es siempre monótono y tiene menor varianza que el primero; frente a los estimadores actuariales presenta la ventaja de que evita la selección de una partición del tiempo en subintervalos, partición que determina la obtención de cada estimador. Si bien, es justo mencionar que el estimador límite producto presenta problemas cuando la hipótesis de independencia entre tiempos de fallo y tiempos de censura no se verifica y, en tales casos, será conveniente plantearse la utilización de estimadores alternativos.

Las ventajas anteriormente mencionadas han convertido al estimador límite producto en el estimador más utilizado y estudiado en el contexto de datos censurados, y ésta es la principal razón de que hayamos elegido este estimador como protagonista de la presente memoria.

Después del ya citado artículo de Kaplan y Meier, donde también se calculan la media y varianza del estimador, hay que esperar hasta los años setenta y ochenta para encontrar propiedades adicionales, paralelas a las ya conocidas para la función de distribución empírica, como son: la normalidad asintótica puntual y sobre intervalos compactos (Breslow y Crowley, 1974), el comportamiento asintótico sobre toda la recta real (Gill, 1983), la convergencia uniforme (Földes y Rejto, 1981) o cotas del tipo Berry-Esseen (Chang y Rao, 1989). A partir de los años ochenta, empiezan a estudiarse representaciones del estimador de Kaplan-Meier como una suma de variables independientes más un término de error cuyo orden se va mejorando progresivamente: Lo y Singh (1985), Burke, Csörgö y Horváth (1981, 1988) y Lo, Mack y Wang (1989) son las principales referencias, y recientemente se han multiplicado los resultados sobre el estudio de las integrales respecto al estimador de Kaplan-Meier ($\int_0^{\infty} x dF_n^c$, donde F_n^c denota el estimador de Kaplan-Meier y \cdot una función arbitraria) para las cuales se ha derivado la consistencia fuerte, la convergencia en distribución y el estudio del sesgo y de la varianza: Stute y Wang (1993) y Stute (1994, 1995, 1996).

A continuación desarrollamos con detalle las cuestiones citadas que nos parecen

de mayor interés.

1.2.1 El estimador de Kaplan y Meier: definición.

El estimador de Kaplan y Meier o estimador límite producto, se puede derivar siguiendo distintos caminos. Aquí exponemos tres de ellos.

1^{er} método. Consiste en dividir el intervalo $I = (0; y]$ en una serie de subintervalos de la forma:

$$I_1 = (0; y_1]; \quad I_2 = (y_1; y_2]; \quad I_3 = (y_2; y_3]; \quad \dots; \quad I_k = (y_{k-1}; y_k = y]:$$

Si p_i denota la probabilidad de supervivencia en I_i (esto es, la probabilidad de que un individuo a riesgo en el instante y_{i-1} sobreviva a y_i); entonces, la probabilidad de supervivencia en I se puede escribir como

$$P_i F(y) = p_1 p_2 \dots p_k$$

y para estimar F bastará con estimar las probabilidades de dicho producto.

Precisamente, el calificativo de "producto" que aparece en el nombre del estimador de Kaplan-Meier hace referencia a la idea de estimar la probabilidad de supervivencia para el intervalo I mediante un producto: el de las estimaciones de las probabilidades de supervivencia para cada uno de los subintervalos I_i . En cuanto al calificativo de "límite" lo da el hecho de que este estimador elige los subintervalos I_i de forma que sólomente contengan una observación muestral (se podría decir que acerca los intervalos al límite). Con esta forma de elegir los subintervalos, resulta sencillo estimar las probabilidades de fallo en cada I_i que serán: 0 si Z_i es un tiempo censurado y $1/n_i$ si Z_i es un tiempo de fallo ($n_i = n^\circ$ de individuos a riesgo al comienzo de I_i). Así, las estimaciones de las probabilidades de supervivencia en I_i (complementarias de las probabilidades de fallo), serán:

$$\hat{p}_i = \begin{cases} 1 - \frac{1}{n_i} & \text{si } Z_i = Y_i \\ 1 & \text{si } Z_i = C_i \end{cases}$$

y el estimador de Kaplan y Meier para la función de distribución F , vendrá dado por:

$$\hat{F}_n^c(y) = \prod_{y_i \leq y} \left(1 - \frac{1}{n_i} \right)$$

Si denotamos por $Z_{1:n} \cdot Z_{2:n} \cdot \dots \cdot Z_{n:n}$ los estadísticos ordenados de la variable Z y por $n_{[1:n]}, n_{[2:n]}, \dots, n_{[n:n]}$ los respectivos concomitantes para la variable

\pm ; podemos escribir el estimador de Kaplan-Meier de la siguiente forma:

$$1 - F_n^c(y) = \prod_{i=Z_{1:n}, y}^{\mu} \frac{1 - \frac{\pm_{[i:n]}}{n_i}}{i_i - 1} = \prod_{i=Z_{1:n}, y}^{\mu} \frac{1}{i_i + 1} \quad (1.6)$$

Si se contempla la posibilidad de empates, la expresión anterior se convierte en:

$$1 - F_n^c(y) = \prod_{i=Z_{1:n}, y}^{\mu} \frac{e_{[i]}}{n_{[i]}} \quad \prod_{\pm_{[i:n]}}$$

siendo $e_{[i]}$ y $n_{[i]}$ el número de empates en $Z_{i:n}$ y el número de Z_j estrictamente mayores que $Z_{i:n}$, respectivamente. Es importante tener en cuenta que en caso de empate un dato no censurado precede, por convención, a uno censurado.

2º método. Esta forma de derivar el estimador de Kaplan-Meier se debe a Efron y la exponemos por su carácter intuitivo (hace lo que a todos se nos ocurriría hacer para aprovechar la información de los tiempos censurados). Consiste en ordenar la muestra de observaciones $f(Z_i; \pm_i)g_{i=1}^n$ de la forma:

$$i \quad \pm_{[1:n]} \quad \pm_{[1:n]} \quad ; \quad i \quad \pm_{[2:n]} \quad \pm_{[2:n]} \quad ; \quad \dots ; \quad i \quad \pm_{[n:n]} \quad \pm_{[n:n]} \quad \pm_{[n:n]}$$

y proceder de forma iterativa para la definición de los saltos de F_n^c , que denotaremos por dF_n^c ; de acuerdo con el esquema siguiente:

$$k = 0: (dF_n^c)^0(Z_{i:n}) = \frac{1}{n}; \quad i = 1; \dots; n:$$

$$k = 1: \text{Si } \pm_{[1:n]} = 1 \text{ entonces } (dF_n^c)^1(Z_{i:n}) = \frac{1}{n}; \quad i = 1; 2; \dots; n:$$

Si $\pm_{[1:n]} = 0$ el tiempo de fallo de ese individuo es estrictamente mayor que $Z_{1:n}$ y por tanto el peso que tiene asignado ($\frac{1}{n}$) se reparte de forma equitativa ($\frac{1}{n(n_i - 1)}$) entre los $Z_{2:n}; \dots; Z_{n:n}$ restantes, de forma que

$$(dF_n^c)^1(Z_{i:n}) = \frac{1}{n} + \frac{1}{n(n_i - 1)} = \frac{1}{n} \left(1 + \frac{1}{n_i - 1} \right) \quad \begin{matrix} (dF_n^c)^1(Z_{1:n}) = 0; \\ i = 2; 3; \dots; n; \end{matrix}$$

$$k = 2: \text{Si } \pm_{[2:n]} = 1 \text{ entonces } (dF_n^c)^2(Z_{i:n}) = (dF_n^c)^1(Z_{i:n}); \quad i = 1; \dots; n:$$

Si $\pm_{[2:n]} = 0$ el tiempo de fallo de ese individuo es estrictamente mayor que $Z_{2:n}$ y por tanto el peso que tiene asignado, $(dF_n^c)^1(Z_{2:n})$; se reparte de forma

equitativa entre los $Z_{3:n}; \dots; Z_{n:n}$ restantes, de forma que

$$\begin{aligned} (dF_n^c)^2(Z_{1:n}) &= (dF_n^c)^1(Z_{1:n}); & (dF_n^c)^2(Z_{2:n}) &= 0; \\ (dF_n^c)^2(Z_{i:n}) &= (dF_n^c)^1(Z_{i:n}) + \frac{(dF_n^c)^1(Z_{2:n})}{n_i - 2} = \\ &= (dF_n^c)^1(Z_{i:n}) \left(1 + \frac{1}{n_i - 2} \right) & i &= 3; \dots; n: \end{aligned}$$

(en la última igualdad se ha tenido en cuenta que para $i \geq 2$ los valores $(dF_n^c)^1(Z_{i:n})$ son iguales).

$k \geq 3$: Si $\pm_{[k:n]} = 1$ entonces $(dF_n^c)^k(Z_{i:n}) = (dF_n^c)^{k-1}(Z_{i:n}); \quad i = 1; \dots; n$:

Si $\pm_{[k:n]} = 0$ el tiempo de fallo de ese individuo es estrictamente mayor que $Z_{k:n}$ y por tanto el peso que tiene asignado, $(dF_n^c)^{k-1}(Z_{k:n})$; se reparte de forma equitativa entre los $Z_{k:n}; \dots; Z_{n:n}$ restantes, de forma que

$$\begin{aligned} (dF_n^c)^k(Z_{i:n}) &= \begin{cases} (dF_n^c)^{k-1}(Z_{i:n}) & \text{si } i < k \\ 0 & \text{si } i = k \\ (dF_n^c)^{k-1}(Z_{i:n}) + \frac{(dF_n^c)^{k-1}(Z_{k:n})}{n_i - k} & \text{si } i > k \end{cases} \\ &= (dF_n^c)^{k-1}(Z_{i:n}) \left(1 + \frac{1}{n_i - k} \right) \end{aligned}$$

Al analizar el paso n se tendrán los saltos de...nitivos.

Este procedimiento muestra claramente que el salto de la función estimadora en un punto Z_i depende únicamente de lo sucedido en los Z_i anteriores. Así por ejemplo, para una muestra ordenada de tamaño $n > 5$ con censuras en el 2º y 5º valores se tendrán los siguientes saltos:

| i | $dF_n^c(Z_{i:n})$ |
|-----|---|
| 1 | $\frac{1}{n}$ |
| 2 | 0 |
| 3 | $\frac{1}{n-3} \left(1 + \frac{1}{n_i-2} \right)$ |
| 4 | $\frac{1}{n-3} \left(1 + \frac{1}{n_i-2} \right)$ |
| 5 | 0 |
| 6 | $\frac{1}{n-3} \left(1 + \frac{1}{n_i-2} \right) \left(1 + \frac{1}{n_i-5} \right)$ |

que de forma general se pueden escribir de la siguiente forma:

$$dF_n^c(Z_{i:n}) = \frac{\pm_{[i:n]}}{n} \prod_{k < i} \left(1 + \frac{1}{n_i - \pm_{[k:n]}} \right)$$

Es fácil comprobar que esta última expresión da los saltos del estimador límite producto derivado por el primer método (ver la Propiedad 1.2 del apartado siguiente). Por tanto, el estimador de la función de distribución F obtenido con ambos procedimientos es el mismo.

3^{er} método. Obtiene el estimador de Kaplan-Meier utilizando la teoría de los procesos empíricos. La idea que subyace en este procedimiento de estimación es la de buscar una expresión equivalente a la función teórica F que dependa de cantidades empíricamente estimables, es decir, debe involucrar cantidades teóricas del vector $(Z; \pm)$; puesto que la muestra observada bajo censura es $f(Z_i; \pm_i)g_{i=1}^n$:

En este sentido, la función razón de fallo acumulada, Λ ; juega un papel crucial, ya que, además de caracterizar la función F mediante la relación (1.3), se expresa de forma sencilla en función de cantidades empíricamente estimables: para $y < b_H$; se verifica que

$$\Lambda(y) = \int_0^y \frac{dH_1(t)}{1 - G(t)}; \quad (1.7)$$

La expresión anterior se obtiene mediante sencillas operaciones y utilizando las relaciones dadas en (1.1), (1.2) y (1.5):

$$\Lambda(y) = \int_0^y \frac{dF(t)}{1 - G(t)} = \int_0^y \frac{(1 - G(t)) dF(t)}{(1 - G(t))(1 - G(t))} = \int_0^y \frac{dH_1(t)}{1 - G(t)};$$

Ahora, si estimamos H y H_1 por sus correspondientes empíricas H_n y H_{1n} , dadas por

$$H_n(z) = \sum_{i=1}^n 1_{fZ_i \cdot z} \frac{1}{n}$$

$$H_{1n}(z) = \sum_{i=1}^n 1_{fZ_i \cdot z; \pm_i = 1g} \frac{1}{n}$$

y sustituimos H y H_1 por H_n y H_{1n} en la expresión (1.7), obtenemos un estimador de Λ ; conocido como el estimador de Nelson-Aalen, y cuya expresión es

$$\Lambda_n^c(z) = \int_0^z \frac{dH_{1n}(t)}{1 - G_n(t)} = \sum_{i=1}^n \frac{1_{fZ_i \cdot y; \pm_i = 1g}}{n \cdot 1 - G_n(Z_i)} = \sum_{i=1}^n \frac{1_{fZ_i \cdot y; \pm_i = 1g}}{n \cdot (\text{rango}(Z_i) - 1)};$$

El estimador Λ_n^c es una función en escalera con saltos en los puntos $fZ_i = Y_i g_{i=1}^n$ y por tanto la relación entre F y Λ dada por la ecuación (1.3) sugiere estimar F de

forma que:

$$\begin{aligned}
 1_i F_n^c(y) &= \prod_{i=1}^{\infty} \frac{1_{f_{Z_i \cdot y; \pm_i=1g}}}{n_i \text{ rango}(Z_i) + 1} = \\
 &= \prod_{i=Z_i \cdot y}^{\infty} \frac{1_{\pm_i}}{n_i \text{ rango}(Z_i) + 1} \\
 &= \prod_{i=Z_i \cdot y}^{\infty} \frac{1}{n_i \text{ rango}(Z_i) + 1}
 \end{aligned} \tag{1.8}$$

que resulta ser, una vez más, el estimador de Kaplan-Meier.

1.2.2 El estimador de Kaplan y Meier: propiedades.

Hemos englobado bajo la denominación común de "propiedades" algunas observaciones sencillas, pero bastante útiles, que se derivan de la propia definición del estimador junto con los resultados relevantes citados al principio de esta sección.

Propiedad 1.1 El estimador de Kaplan-Meier coincide con la función de distribución empírica en el caso de ausencia de censura.

A partir de la expresión (1.6) del estimador y utilizando que cuando no hay censuras todos los \pm toman el valor 1 y las Z coinciden con los valores de Y se llega a

$$\begin{aligned}
 1_i F_n^c(y) &= \prod_{i=Y_{i:n}, y}^{\infty} \frac{1}{n_i i + 1} = \prod_{i=Y_{i:n}, y}^{\infty} \frac{n_i i}{n_i i + 1} = \\
 &= \frac{n_i}{n} \frac{1}{n_i} \frac{2}{n_i} \frac{3}{n_i} \dots \frac{n_i k}{n_i k + 1} = \frac{n_i k}{n} = 1_i \prod_{i=1}^k \frac{1_{f_{Y_{i:n}, yg}}}{n}
 \end{aligned}$$

donde $k = \max_{i=Y_{i:n}, y} (evidentemente k = \sum_{i=1}^n 1_{f_{Y_{i:n}, yg}})$:

Propiedad 1.2 El estimador de Kaplan-Meier tiene forma escalonada con saltos en las observaciones no censuradas. La magnitud del salto en $Z_{i:n}$; $dF_n^c(Z_{i:n})$, puede expresarse de las siguientes formas

$$\frac{1_{\pm_{[i:n]}}}{n} \prod_{k=1}^{\infty} \frac{1}{n_i k} = \frac{1_{\pm_{[i:n]}}}{n_i i + 1} \prod_{k=1}^{\infty} \frac{1_{f_{Z_{i+1:n}}}}{n_i k + 1}$$

La primera expresión se deriva del 2º método de obtención del estimador; la segunda expresión se obtiene al calcular $F_n^c(Z_{i:n})$ y $F_n^c(Z_{i-1:n})$ utilizando (1.6). La igualdad entre ambas se prueba fácilmente, pues

$$\begin{aligned} & \frac{\prod_{k=1}^{i-1} \left(1 + \frac{1}{n_{i-k}}\right)}{n} \prod_{k=1}^{i-1} \frac{1}{n_{i-k}} \\ &= \frac{\prod_{k=1}^{i-1} \frac{n_{i-k+1}}{n_{i-k}}}{n} \prod_{k=1}^{i-1} \frac{1}{n_{i-k}} = \\ &= \frac{\prod_{k=1}^{i-1} \frac{n_{i-k+1}}{n_{i-k}}}{n} \prod_{k=1}^{i-1} \frac{1}{n_{i-k}} \end{aligned}$$

Propiedad 1.3 El estimador de Kaplan-Meier puede interpretarse como el estimador no paramétrico máximo-verosímil de F .

La demostración se puede ver en la sección 5 del artículo de Kaplan y Meier (1958) o también en Wang (1987).

Propiedad 1.4 El estimador de Kaplan-Meier para la función de distribución G de la variable censura C viene dado por

$$1 - G_n^c(y) = \prod_{i=Z_{i:n}, y} \frac{1 - \frac{1}{n_{i+1}}}{n_{i+1}} = \prod_{i=Z_{i:n}, y} \frac{1}{n_{i+1}}$$

Es inmediato sin más que tener en cuenta que una observación censurada de Y_i corresponde a una observación no censurada de C_i y viceversa.

Propiedad 1.5 El producto de los estimadores de Kaplan-Meier para las funciones de supervivencia de las variables Y y C coincide con la distribución empírica de la supervivencia de la variable Z , es decir

$$(1 - F_n^c(y))(1 - G_n^c(y)) = 1 - H_n(y):$$

Explicitando las fórmulas del primer miembro de la igualdad anterior se tiene

$$\begin{aligned} & \prod_{i=Z_{i:n}, y} \frac{1}{n_{i+1}} \prod_{i=Z_{i:n}, y} \frac{1}{n_{i+1}} \\ &= \prod_{i=Z_{i:n}, y} \frac{1}{n_{i+1}} \end{aligned}$$

y

$$\sum_{i=Z_{i:n}, y} \frac{1}{n_i} = \sum_{i=1}^n \frac{1_{fZ_{i:n}, yg}}{n}$$

como ya se vió en la Propiedad 1.1.

Propiedad 1.6 Bajo la suposición de que las distribuciones F y G sean continuas, se veri...ca:

a) Puntualmente en cada $y \cdot T < b_H$; que

$$P_{\bar{n}}(F_n^c(y) - F(y)) \xrightarrow{d} N(0; \mathcal{K}(y))$$

donde $\mathcal{K}(y) = (1 - F(y))^2 \int_0^y (1 - H(t))^2 dH_1(t)$:

b) Globalmente en $D[0; T]$; se tiene

$$P_{\bar{n}}(F_n^c - F) \xrightarrow{d} Z(t)$$

donde Z es un proceso Gaussiano de media 0 y función de covarianzas

$$\text{Cov}(Z(y); Z(t)) = (1 - F(y))(1 - F(t)) \int_0^{y \wedge t} (1 - H(s))^2 dH_1(s)$$

siendo $y \wedge t = \min\{y, t\}$, $F : ([0; T]; \mathbb{R})$ el conjunto de las funciones que van del intervalo $[0; T]$ en \mathbb{R} y $D[0; T] = \{f \in C([0; T]; \mathbb{R}) : f \text{ continua por la derecha y con discontinuidades, a lo sumo, de salto}\}$; con la topología de Skorohod (Billingsley, pag.111).

Es la normalidad asintótica puntual y sobre intervalos compactos y fue probada por Breslow y Crowley en 1974.

Propiedad 1.7 Suponiendo que la distribución F es continua y que $\int_0^{b_H} \frac{dF(s)}{1 - G(s)} < \infty$, se veri...ca:

$$P_{\bar{n}}(F_n^c - F) \xrightarrow{d} Z \text{ en } D[0; b_H]$$

donde Z es el proceso Gaussiano definido en el apartado b) de la propiedad anterior.

Esta propiedad muestra el comportamiento asintótico del estimador de Kaplan-Meier sobre toda la recta real y se debe a Gill (1983).

Propiedad 1.8 Bajo la suposición de que las distribuciones F y G sean continuas y para $T < b_H$, se verifica:

$$\sup_{0 \leq y \leq T} |F_n^c(y) - F(y)| = O\left(\frac{\sqrt{\ln n}}{n}\right) \quad \text{c.s.}$$

Este orden de convergencia uniforme fuerte se debe a Földes y Rejto (1981).

Propiedad 1.9 Bajo la suposición de que las distribuciones F y G sean continuas, se verifica puntualmente en cada $y < b_H$; que

$$\sup_{x \in \mathbb{R}} \left| \frac{F_n^c(y) - F(y)}{\sqrt{V(y)}} - \Phi_N(x) \right| \leq K \frac{1 + V(y)}{(1 - H(y))^2} n^{-1/2}$$

donde Φ_N denota la función de distribución de una $N(0; 1)$ y $V(y)$ viene dada por $V(y) = \frac{1}{4}(y) = (1 - F(y))^2$ ($\frac{1}{4}(y)$ ha sido definida en el apartado a) de la propiedad 1.6).

Esta cota del tipo Berry-Esseen fue calculada por Chang y Rao en 1989.

Propiedad 1.10 Bajo la suposición de que las distribuciones F y G sean continuas, se verifica en cada $y < T < b_H$; que

$$F_n^c(y) = F(y) + \frac{1}{n} \sum_{i=1}^n \varepsilon_i(y) + r_n(y)$$

donde

$$\varepsilon_i(y) = (1 - F(y)) \frac{1_{F(Z_i) \leq y; \pm 1 = 1g}}{1 - H(Z_i)} - \int_0^y \frac{1_{f(t) \leq Z_i g}}{(1 - H(t))^2} dH_1(t)$$

y

$$\sup_{0 \leq y \leq T} |r_n(y)| = O\left(\frac{\sqrt{\ln n}}{n}\right) \quad \text{c.s.}$$

Además, para ε_i ; se tiene que

$$\sup_{0 \leq y \leq T} E |r_n(y)|^2 = O\left(\frac{\ln n}{n}\right) :$$

Este resultado, debido a Lo y Singh (1985), permite escribir la diferencia entre el estimador de Kaplan-Meier y la función de distribución teórica como una suma de variables independientes de media cero mas un término de error cuyo orden es conocido. La importancia de esta representación casi segura es que permite trabajar con la parte principal del estimador (mucho más manejable al tratarse de una suma de v.a. en lugar de un producto) y derivar propiedades de interés como, por ejemplo, la normalidad asintótica y la convergencia débil del proceso.

Propiedad 1.11 En las mismas condiciones y con la misma notación de la propiedad anterior se verifica que

$$F_n^c(y) = F(y) + \frac{1}{n} \sum_{i=1}^n \epsilon_i(y) + r_n(y)$$

donde

$$\sup_{0 \leq y \leq T} |r_n(y)| = O\left(\frac{\mu \ln n}{n}\right) \quad \text{c.s.}$$

y

$$\sup_{0 \leq y \leq T} E |r_n(y)|^2 = O\left(\frac{\mu^2 \ln n}{n}\right) \quad \text{8.1.1}$$

Esta mejora en el orden del error para la representación casi segura se puede ver en Lo, Mack y Wang (1989), cuya demostración se basa en un resultado relevante de Burke, Csörgö y Horváth (1988). También se puede derivar como caso particular de las últimas representaciones casi seguras obtenidas para el estimador límite producto de la función de distribución bajo censura y truncamiento (que detallaremos más adelante), suponiendo ausencia de truncamiento.

Propiedad 1.12 Bajo la suposición de que las distribuciones F y G no tengan saltos en común y denotando por A al conjunto de los valores discretos de H, se verifica, con probabilidad uno y en media, que

$$\begin{aligned} \lim_{n \rightarrow \infty} \int_Z \phi(y) dF_n^c(y) &= \int_Z \phi(y) dF(y) + \sum_{a_i \in A} \phi(a_i) [F(a_i) - F(a_i^-)] \\ &= \int_{\phi(y) < b_H} \phi(y) dF(y) + 1_{\phi(b_H) > 2A} \phi(b_H) [F(b_H) - F(b_H^-)] \end{aligned}$$

donde ϕ es cualquier función F-integrable.

En particular si F es continua y $b_H = b_F$ la expresión anterior se reduce a

$$\lim_{n \rightarrow \infty} \int_Z \phi(y) dF_n^c(y) = \int_Z \phi(y) dF(y) \quad (1.9)$$

Este potente resultado de consistencia puede verse en Stute y Wang (1993).

Además, tomando distintas funciones particulares ψ ; se obtienen propiedades importantes, entre las que cabe destacar:

² La consistencia fuerte puntual de F_n^c para $y \in [0, b_H]$:

$$\lim_{n \rightarrow \infty} F_n^c(y) = F(y) \quad \text{c.s.}$$

Basta tomar $\psi(y) = 1_{(0, y]}$:

² La consistencia fuerte de los estimadores de momentos determinados a partir de F_n^c . Particularmente para la media y varianza teóricas se tiene:

$$\lim_{n \rightarrow \infty} \int_{-\infty}^{\infty} y dF_n^c(y) = E(Y) \quad \text{c.s.}$$

$$\lim_{n \rightarrow \infty} \int_{-\infty}^{\infty} y^2 dF_n^c(y) - \left(\int_{-\infty}^{\infty} y dF_n^c(y) \right)^2 = \text{Var}(Y) \quad \text{c.s.}$$

Basta tomar $\psi_1(y) = y$ en el primer caso y $\psi_2(y) = y^2$ en el segundo.

² La consistencia uniforme fuerte para F_n^c en el intervalo $[0, b_H]$:

$$\sup_{y \in [0, b_H]} |F_n^c(y) - F(y)| \rightarrow 0 \quad \text{c.s.}$$

En Stute (1976) se considera una aproximación a la convergencia uniforme a partir de la convergencia fuerte puntual, como alternativa a la demostración clásica que utiliza acotaciones exponenciales. Esta aproximación junto con (1.9) demuestra el resultado de convergencia uniforme.

Propiedad 1.13 Con la notación de la propiedad anterior y bajo ciertas condiciones de integrabilidad (bastante generales) sobre la función ψ ; se verifica que

$$\sqrt{n} \int_{-\infty}^{\infty} \psi(y) dF_n^c(y) - \int_{-\infty}^{\infty} \psi(y) d\tilde{F}(y) \xrightarrow{d} N(0, \frac{1}{4})$$

donde

$$\tilde{F}(y) = \begin{cases} F(y) & \text{si } y < b_H \\ F(b_H) + \frac{1}{2} \frac{F(b_H) - F(b_H)}{b_H - b_H} & \text{si } y = b_H \\ F(b_H) & \text{si } y > b_H \end{cases}$$

y $\frac{1}{4} = \text{Var} \left(\int_{-\infty}^{\infty} \psi(Z) dZ + \int_{-\infty}^{\infty} \psi(Z) dZ \right)$:

Este resultado se puede ver en Stute (1995) y claramente generaliza los resultados de normalidad asintótica ya citados, obtenidos por Breslow y Crowley o Lo y Singh, para el estimador de Kaplan-Meier. Las funciones ϕ_0 , ϕ_1 y ϕ_2 tienen una expresión complicada, que en el caso particular de que F y G sean funciones continuas se simplifica dando lugar a

$$\begin{aligned} \phi_0(y) &= \frac{1}{1 - G(y)}; & \phi_1(y) &= \frac{1}{1 - H(y)} \int_y^{Z_{b_H}} \lambda(w) dF(w) \\ \phi_2(y) &= \int_y^{Z_{b_H}} \lambda(w) \int_y^{Z_{b_H}} 1_{FV < y; v < w} (1 - H(v))^2 (1 - F(v)) dG(v) dF(w); \end{aligned}$$

Propiedad 1.14 Utilizando la notación de la Propiedad 1.12 y suponiendo que las distribuciones F y G no tienen saltos en común, que H es continua y $\lambda \geq 0$; se tiene que

$$\int_0^Z \lambda(y) G(y) H^{n_i-1}(y) dF(y) \cdot \text{sesgo} \int_0^Z \lambda(y) dF_n^c(y)$$

y

$$\text{sesgo} \int_0^Z \lambda(y) dF_n^c(y) \cdot \int_0^Z \lambda(y) (1 - G(y)) \int_0^y \frac{H^{n_i-1}(s) dG(s)}{(1 - G(y))^2} dF(y)$$

La demostración de esta acotación para el sesgo de una integral de Kaplan-Meier se debe a Stute (1994). En dicha referencia se presenta, además, una expresión exacta de dicho sesgo.

Propiedad 1.15 Con la notación e hipótesis de la Propiedad 1.13, se verifica, con probabilidad 1, que

$$\lim_{n \rightarrow \infty} n \widehat{\text{Var}}^{\text{Jack}} = \frac{3}{4}$$

donde $\widehat{\text{Var}}^{\text{Jack}}$ denota el estimador jackknife de la varianza de $\int_0^R \lambda(y) dF_n^c(y)$, tomando $\lambda_{[n:n]}^{\pm} = 0$ en el caso de que ocurra la situación $\lambda_{[n_i-1:n]} = 0$ y $\lambda_{[n:n]} = 1$.

De este modo se tiene que, bajo ciertas condiciones de integrabilidad de la función λ ; el estimador jackknife de la varianza de una integral de Kaplan-Meier estima consistentemente la varianza límite. La demostración es de Stute (1996).

1.3 El modelo de truncamiento aleatorio por la izquierda.

Otra de las situaciones donde hay una pérdida de información en los datos y que, al igual que la censura, es objeto de estudio en este trabajo es la presencia de datos truncados. El problema del truncamiento aparece cuando la variable de interés no se observa siempre (caso de una m.a.s. para datos completos) sino solamente cuando se cumple una determinada condición, y en este caso, además de observar la variable en estudio obtenemos el valor de otra variable involucrada en la condición de observación y que se denomina truncamiento. Fijémonos, para empezar, en los siguientes ejemplos:

- Supongamos que la variable de interés, Y_i ; es el tiempo de duración de una determinada enfermedad y que otra variable, T_i ; mide el tiempo desde que la enfermedad comienza en un individuo hasta que éste entra en estudio. Evidentemente, el especialista tendrá datos correspondientes a los pacientes que entran en el estudio, esto es, individuos con $T_i \cdot Y_i$:

- En Astronomía tiene interés estudiar la luminosidad de objetos astronómicos. Si Y mide el brillo de un objeto observado desde la tierra, uno sólo puede detectar objetos que sean suficientemente brillantes, es decir, observaremos objetos con $T_i \cdot Y_i$, donde T representa el brillo mínimo detectable a una cierta distancia.

- En la exploración de reservas de petróleo, la probabilidad de encontrar un yacimiento es proporcional al tamaño del yacimiento. Si Y mide el tiempo hasta encontrar un yacimiento y X el tamaño del yacimiento, la muestra de tiempos observados estará formada por Y_i que superen una cierta barrera de tiempo que dependerá de X_i .

Todos estos ejemplos tienen en común que proporcionan una muestra de observaciones ($Y_i; T_i$) solamente cuando $T_i \cdot Y_i$ y, en otro caso, no hay observación. Estamos ante un mecanismo denominado truncamiento por la izquierda que consiste en que una variable llamada de truncamiento, T ; impide observar la variable de interés, Y ; cuando ésta última se sitúa a su izquierda ($Y < T$).

Como primera observación conviene notar que el problema del truncamiento difiere del problema de la censura en un aspecto fundamental: cuando hay censura la muestra del estudio es una muestra aleatoria de todos los individuos cuyo tiempo de fallo Y interesa estudiar, aunque posteriormente dicho tiempo de fallo pueda no ser completamente observado; por el contrario el truncamiento proporciona una muestra aleatoria de una parte de los individuos a estudiar: aquellos que verifican una cierta condición sobre la variable Y ($Y \geq T$); condición que nosotros no controlamos y que conlleva la aleatoriedad en el tamaño de la muestra de trabajo.

A continuación definiremos el modelo de truncamiento aleatorio por la iz-

quierda, modelo que recoge estas peculiaridades.

Sea $(T_i; Y_i) \quad i = 1; 2; \dots; N$ una muestra aleatoria simple de un cierto vector aleatorio formado por dos variables positivas $(T; Y)$, donde Y es la variable de interés y T es otra variable denominada variable de truncamiento. Supongamos que uno sólo observa aquellos pares $(T_i; Y_i)$ para los cuales $T_i \cdot Y_i$ (aunque el índice i no tiene porque corresponder con su valor inicial). Entonces n es aleatorio y N desconocido. La muestra observada es de la forma:

$$\begin{array}{|c|} \hline (T_i; Y_i) \quad \text{con } T_i \cdot Y_i \\ \hline i = 1; 2; \dots; n \quad (n \cdot N) \\ \hline \end{array}$$

Además, una condición habitual en este modelo es

$$\boxed{Y \text{ y } T \text{ son v.a. independientes.}}$$

En este modelo de truncamiento aleatorio por la izquierda, se definen varias funciones importantes relacionadas con las variables aleatorias Y y T ; que nosotros supondremos no negativas, y que son:

1) La función de distribución de la variable de interés, $F(y) = P(Y \leq y)$ y su función de supervivencia, $S(y) = 1 - F(y) = P(Y > y)$:

2) La función de distribución de la variable de truncamiento, $L(y) = P(T \leq y)$ y su función de supervivencia, $1 - L(y) = P(T > y)$:

3) La función de distribución de la variable Y condicionada a que $T \cdot Y$; $F^{\circ}(y) = P(Y \leq y | T \cdot Y)$:

Llamando $\circ = P(T \cdot Y)$ a la probabilidad de ausencia de truncamiento, que necesariamente será mayor que cero (en caso contrario no tendríamos observaciones), se tiene que

$$F^{\circ}(y) = \frac{P(Y \leq y; T \cdot Y)}{P(T \cdot Y)} = \int_0^y \int_0^t dL(s) dF(t) = \int_0^y \int_0^t L(s) dF(t) : \quad (1.10)$$

4) La función de distribución de la variable T condicionada a que $T \cdot Y$, $L^{\circ}(y) = P(T \leq y | T \cdot Y)$ y que se puede escribir de la forma

$$\begin{aligned} L^{\circ}(y) &= \frac{P(T \leq y; T \cdot Y)}{P(T \cdot Y)} = \int_0^y \int_0^t dF(s) dL(t) = \\ &= \int_0^y \int_0^t L_i F(t) dL(t) : \quad (1.11) \end{aligned}$$

5) La función auxiliar $C(y) = P(T \cdot y \cdot Y \leq T \cdot Y)$ que se puede expresar como

$$C(y) = \int_0^y \int_0^y dL(s)dF(t) = \int_0^y L(y) dF(t) : \quad (1.12)$$

6) La función de azar de Y definida por

$$r(y) = \lim_{h \downarrow 0} \frac{P(y < Y \leq y + h | Y = y)}{h} = \frac{dF(y)}{dF(y)}$$

7) La función de azar acumulada, que se define a partir de la función anterior mediante la expresión:

$$\Lambda(y) = \int_0^y r(t)dt = \int_0^y dF(t)$$

y que al igual que en la situación de datos censurados jugará un papel fundamental para construir estimadores de la función de distribución de la variable Y ; al caracterizarla de forma única mediante la relación (análoga a (1.3)):

$$1 - F(y) = e^{-\Lambda_c(y)} \prod_{a_i \leq y} (1 - \Delta a_i) \quad (1.13)$$

donde Λ_c denota la parte continua de Λ , A el conjunto de puntos donde Λ tiene discontinuidades de salto y $\Delta a_i = \Lambda(a_i) - \Lambda(a_i^-)$ la magnitud del salto de Λ en a_i :

En el presente modelo, también es importante definir y analizar los siguientes conceptos:

8) Sean $a_F; b_F; a_L; b_L; a_{F^*}; b_{F^*}; a_{L^*}$ y b_{L^*} los extremos de los soportes de las funciones F, L, F^* y L^* respectivamente, según la definición dada por (1.4). Es sencillo comprobar (mediante un mero análisis gráfico) que:

$$\begin{aligned} a_{F^*} &= \max\{a_F; a_L\} \\ b_{F^*} &= b_F \\ a_{L^*} &= a_L \\ b_{L^*} &= \min\{b_L; b_F\} \end{aligned} \quad (1.14)$$

9) $\otimes = P(T \cdot Y)$; ya hemos dicho que representa la probabilidad de ausencia de truncamiento, es decir, de la región del plano donde las variables $(Y; T)$ son

observables, por tanto su probabilidad debe ser estrictamente positiva: $\theta > 0$: Además, θ se puede expresar como

$$\theta = P(T < Y) = \int_0^{Z_1} \int_0^t dL(s) dF(t) = \int_0^{Z_1} L(t) dF(t)$$

o también

$$\theta = P(T < Y) = \int_0^{Z_1} \int_t^{Z_1} dF(s) dL(t) = \int_0^{Z_1} (1 - F(t)) dL(t):$$

Es interesante notar que de forma análoga al truncamiento por la izquierda, se puede plantear un mecanismo de truncamiento por la derecha, que proporcionará observaciones sólo cuando la variable de interés Y no supere a la variable de truncamiento T . Entonces, la muestra de datos observados será $(Y_i; T_i)$ con $T_i \geq Y_i$.

Un ejemplo relevante de esta situación aparece al estudiar la distribución del periodo de latencia del virus del SIDA, definido éste como el tiempo que transcurre entre la infección por el virus hasta el desarrollo de la enfermedad. Los datos correspondientes a individuos contaminados por transfusiones de sangre proporcionan una fuente de información importante al permitir conocer la fecha de infección, pero presentan el problema de que corresponden a individuos que han desarrollado la enfermedad dentro de cierto intervalo de tiempo, siendo desconocido el número total de personas infectadas por dicha vía. Si llamamos Y al tiempo de latencia y T al tiempo que transcurre desde la infección hasta que el individuo entra en estudio sólo observaremos individuos cuando $Y_i < T_i$: Este problema concreto ha sido estudiado por Lagakos, Barraj y De Gruttola (1988) considerando el tiempo cronológico de forma invertida, de manera que el problema se transforma en otro susceptible de análisis en presencia de truncamiento por la izquierda.

A partir de este momento estudiaremos problemas de estimación para datos truncados por la izquierda.

1.4 Estimación de la función de distribución con truncamiento.

Ya hemos visto que en presencia de datos truncados por la izquierda observamos una muestra $(T_i; Y_i)$ con $T_i < Y_i$; para $i = 1; 2; \dots; n$ (n aleatorio). Nuestro objetivo en esta sección es estimar la función de distribución de la variable Y teniendo en cuenta la peculiaridad de estos datos, ya que trabajar con la función de distribución empírica construída con las observaciones Y_i proporcionaría resultados sesgados (es intuitivo pensar que una muestra así observada tiende a perder valores pequeños de la variable Y).

Este problema se aborda por primera vez en el campo de la Astronomía; es Linden-Bell quien publica en 1971 en la revista *Monthly Notices of the Royal Astronomical Society* un estimador de la función de distribución para datos truncados que resulta ser el estimador máximo verosímil de la función de distribución con truncamiento. Las propiedades del estimador de Linden-Bell, similar en muchos aspectos al estimador de Kaplan y Meier (es un estimador límite producto), comenzaron a estudiarse con más retraso que las de éste último: la consistencia es debida a Woodroffe (1985), distintas propiedades sobre la convergencia asintótica se estudian en Woodroffe (1985), Wang, Jewell y Tsai (1986) y Keiding y Gill (1990). Esto dio lugar a que se dispusiese de algunas representaciones del estimador de Kaplan-Meier como suma de variables i.i.d. más un término de error, representaciones que se fueron trasladando progresivamente al campo del truncamiento: Chao y Lo (1988), Stute (1993a) y Arcones y Giné (1995) son las principales referencias.

En lo que resta de esta sección expondremos con detalle las cuestiones anteriormente citadas.

1.4.1 El estimador de Linden-Bell: definición.

Definiremos el estimador de Linden-Bell siguiendo a Woodroffe (1985).

En primer lugar, Woodroffe demuestra que la función de distribución F de la variable Y sólo puede ser estimada a partir de las observaciones $(T_i; Y_i)$ con $T_i \leq Y_i$; si se verifica la condición (denominada condición de identificabilidad de F) de

$$a_L \leq a_F: \quad (1.15)$$

Bajo dicha condición es sencillo demostrar que la función de azar acumulada Λ se expresa en función de cantidades empíricamente estimables:

$$\Lambda(y) = \int_0^y \frac{dF^*(t)}{C(t)} \quad 0 \leq y < 1: \quad (1.16)$$

Para ello basta tener en cuenta que $dF^*(y) = \int_0^y L(t) dF(t)$ (ver expresión (1.10)) y $C(y) = \int_0^y L(t) (1 - F(t)) dt$ (ver expresión (1.12)). Además $L(t) \geq 0$; $\int_0^y L(t) dt > a_F$ (ya que $a_L \leq a_F$) y entonces, mediante sencillas operaciones, se llega al resultado deseado:

$$\Lambda(y) = \int_0^y \frac{dF(t)}{\int_0^t L(u) (1 - F(u)) du} = \int_0^y \frac{\int_0^t L(u) dF(u)}{\int_0^t L(u) (1 - F(u)) du} = \int_0^y \frac{dF^*(t)}{C(t)}:$$

Ahora, si estimamos F^* y C por sus correspondientes empíricas F_n^* y C_n , dadas

por

$$F_n^a(y) = \sum_{i=1}^n 1_{F_{Y_i} \cdot y} \frac{1}{n}$$

$$C_n(y) = \sum_{i=1}^n 1_{F_{T_i} \cdot y \cdot Y_i} \frac{1}{n}$$

y sustituimos F^a y C por F_n^a y C_n en la expresión (1.16), obtenemos un estimador de Λ ; cuya expresión es

$$\Lambda_n^t(y) = \int_0^y \frac{dF_n^a(t)}{C_n(t)} = \sum_{i=1}^n \frac{1_{F_{Y_i} \cdot y}}{nC_n(Y_i)} = \sum_{i=1}^n \frac{1_{F_{Y_i} \cdot y}}{\prod_{j=1}^n 1_{F_{T_j} \cdot Y_i \cdot Y_j}}$$

El estimador Λ_n^t es una función en escalera con saltos en los puntos $Y_1; Y_2; \dots; Y_n$ y por tanto, la relación entre F y Λ dada por la ecuación (1.13) sugiere estimar F de forma que:

$$\int_0^y F_n^t(y) = \int_0^y \frac{1_{F_{Y_i} \cdot y}}{nC_n(Y_i)} = \int_0^y \frac{1_{F_{Y_i} \cdot y}}{\prod_{j=1}^n 1_{F_{T_j} \cdot Y_i \cdot Y_j}} \quad (1.17)$$

estimador conocido con el nombre de estimador de Linden-Bell.

El estimador de Linden-Bell es un estimador límite producto puesto que (utilizando términos del Análisis de Supervivencia) estima la probabilidad de supervivencia en un intervalo $(0; y]$ mediante el producto de las supervivencias estimadas en cada uno de los subintervalos de una partición de $(0; y]$ construida de forma que en cada uno de dichos subintervalos sólo haya una observación (intervalos límite).

Denotando la partición por

$$I_1 = (0; y_1]; \quad I_2 = (y_1; y_2]; \quad I_3 = (y_2; y_3]; \quad \dots; \quad I_k = (y_{k-1}; y_k = y]$$

lo anterior significa que

$$\int_0^y F_n^t(y) = \hat{p}_1 \hat{p}_2 \dots \hat{p}_k$$

donde

$$\hat{p}_i = \begin{cases} \frac{1}{2} \int_0^y \frac{1}{n_i} & \text{si } y_i = Y_i \\ 1 & \text{si } y_i = T_i \end{cases} \quad i = 1; 2; \dots; k$$

y $n_i = n^0$ de individuos a riesgo al comienzo de I_i .

Por la naturaleza del truncamiento y el diseño de los subintervalos, n_i es el número de individuos observados hasta y_{i-1} (incluido) menos los que han muerto hasta dicho instante, es decir, para los $y_i = Y_i$

$$n_i = \sum_{j=1}^X 1_{f_{T_j} \cdot Y_i > y} - \sum_{j=1}^X 1_{f_{Y_j} < Y_i} = \sum_{j=1}^X 1_{f_{T_j} \cdot Y_i \cdot Y_j > y}$$

y por tanto

$$1_i F_n^t(y) = \sum_{i=1}^Y \frac{1_i \prod_{j=1}^{\tilde{A}} 1_{f_{Y_j} \cdot y}}{\prod_{j=1}^n 1_{f_{T_j} \cdot Y_i \cdot Y_j}}$$

En el caso de que se consideren empates, la expresión anterior se convierte en:

$$1_i F_n^t(y) = \sum_{i=Y_i \cdot y}^Y \frac{1_i \prod_{j=1}^{\tilde{A}} e_i}{\prod_{j=1}^n 1_{f_{T_j} \cdot Y_i \cdot Y_j}}$$

siendo e_i el número de empates en Y_i :

Finalmente, hemos obtenido el estimador de Linden-Bell mediante un procedimiento alternativo al utilizado por Woodroffe.

1.4.2 El estimador de Linden-Bell: propiedades.

Todas las propiedades que se exponen a continuación parten de las hipótesis ya mencionadas de independencia entre Y y T ; así como $a_L \cdot a_F$:

Propiedad 1.16 El estimador de Linden-Bell coincide con la función de distribución empírica en el caso de ausencia de truncamiento.

A partir de la expresión (1.17) del estimador y utilizando que cuando no hay truncamiento todos las $f_{T_j}^n$ toman el valor 0 y las $f_{Y_j}^n$ constituyen una m.a.s. de tamaño n de toda la variable de interés se llega a

$$\begin{aligned} 1_i F_n^t(y) &= \sum_{i=1}^Y \frac{1_i \prod_{j=1}^{\tilde{A}} 1_{f_{Y_j} \cdot y}}{\prod_{j=1}^n 1_{f_{T_j} \cdot Y_i \cdot Y_j}} = \sum_{i=Y_{i:n} \cdot y}^Y \frac{1_i}{n_i (i-1)} = \\ &= \sum_{i=Y_{i:n} \cdot y}^Y \frac{n_i}{n_i (i+1)} = \frac{n_i}{n} \frac{1}{n_i} \frac{2}{n_i} \frac{3}{n_i} \dots \frac{k}{n_i} = \\ &= \frac{n_i}{n} = \sum_{i=1}^X \frac{1}{n} 1_{f_{Y_{i:n}} \cdot y} \end{aligned}$$

donde $k = \max_{i=Y_{i:n} \cdot y} \prod_{j=1}^{\tilde{A}} 1_{f_{Y_j} \cdot y}$: (evidentemente $k = \prod_{j=1}^{\tilde{A}} 1_{f_{Y_j} \cdot y}$):

Propiedad 1.17 El estimador de Linden-Bell tiene forma escalonada con saltos en las observaciones Y_i . La magnitud del salto en Y_i , $dF_n^t(Y_i)$; es igual a

$$dF_n^t(Y_i) = \frac{1 - F_n^t(Y_{i-1})}{nC_n(Y_i)};$$

Esta expresión se obtiene al calcular $F_n^t(Y_i) - F_n^t(Y_{i-1})$ utilizando (1.17):

$$\begin{aligned} & F_n^t(Y_i) - F_n^t(Y_{i-1}) \\ &= \prod_{j=1}^i \frac{1 - F_n^t(Y_{j-1})}{nC_n(Y_j)} - \prod_{j=1}^{i-1} \frac{1 - F_n^t(Y_{j-1})}{nC_n(Y_j)} \\ &= \prod_{j=1}^{i-1} \frac{1 - F_n^t(Y_{j-1})}{nC_n(Y_j)} \left(\frac{1 - F_n^t(Y_{i-1})}{nC_n(Y_i)} - 1 \right) \\ &= \prod_{j=1}^{i-1} \frac{1 - F_n^t(Y_{j-1})}{nC_n(Y_j)} \left(\frac{1 - F_n^t(Y_{i-1})}{nC_n(Y_i)} - \frac{1 - F_n^t(Y_{i-1})}{nC_n(Y_i)} \right) \\ &= \prod_{j=1}^{i-1} \frac{1 - F_n^t(Y_{j-1})}{nC_n(Y_j)} \frac{1 - F_n^t(Y_{i-1})}{nC_n(Y_i)} \left(\frac{1}{nC_n(Y_i)} - 1 \right) = \frac{1 - F_n^t(Y_{i-1})}{nC_n(Y_i)} \end{aligned}$$

Propiedad 1.18 El estimador de Linden-Bell es el estimador no paramétrico de máxima verosimilitud de F , siempre que $\sum_{j=1}^n 1_{F_{T_j} \cdot Y_i \cdot Y_j} > 0$; para todos los valores observados Y_i excepto el mayor.

La demostración se puede ver en Wang (1987).

Propiedad 1.19 Si $b_L \cdot b_F$; el estimador de Linden-Bell para la función de distribución L de la variable de truncamiento T viene dado por

$$L_n^t(y) = \prod_{i=1}^n \frac{1_{F_{T_i} > y}}{nC_n(T_i)} = \prod_{i=1}^n \frac{1_{F_{T_i} > y}}{\prod_{j=1}^n 1_{F_{T_j} \cdot T_i \cdot Y_j}}; \quad (1.18)$$

Este resultado está en Woodroffe (1985) y se deriva bajo la condición de identidad...abilidad de L dada por

$$b_L \cdot b_F; \quad (1.19)$$

El proceso es análogo al seguido para obtener el estimador F_n^t ; consistente en encontrar una expresión que caracterice L y que dependa de cantidades empíricamente estimables. Dicha expresión viene dada por $O_L(y) = \frac{dL(y)}{L(y)}$ que determina la distribución L de forma única mediante la relación:

$$L(y) = e^{\int_{-\infty}^y O_L(t) dt} \prod_{a_i > y} (1 - O_L(a_i)) \quad (1.20)$$

donde O_L^c denota la parte continua de O_L , A el conjunto de puntos donde O_L tiene discontinuidades de salto y $O_L \text{ faig} = O_L(a_i) - O_L^-(a_i)$ la magnitud del salto de O_L en a_i :

Además, $O_L(y) = \int_y^{\infty} \frac{dL(t)}{L(t)}$ se puede escribir como

$$\int_y^{\infty} \frac{dL(t)}{L(t)} = \int_y^{\infty} \frac{\sum_{i=1}^n (1 - F(t_i)) dL(t)}{\sum_{i=1}^n (1 - F(y_i)) L(t)} = \int_y^{\infty} \frac{dL^a(t)}{C(t)}$$

sin más que utilizar (1.11), (1.12) y (1.19).

Sustituyendo en el último término de la igualdad de arriba L^a y C por sus estimadores empíricos y utilizando el resultado (1.20) se obtiene el estimador L_n^t :

Propiedad 1.20 Bajo la suposición de que las distribuciones F y L sean continuas e idéntificables ($a_L = a_F$ y $b_L = b_F$), se verifica que:

$$\sup_{y>0} \int_y^{\infty} F_n^t(y) - F(y) \rightarrow 0:$$

Esta propiedad de consistencia uniforme débil se debe a Woodroffe (1985).

Propiedad 1.21 Si las distribuciones F y L son continuas e idéntificables y se verifica alguna de las dos siguientes condiciones:

- i) $a_L < a_F$
- ii) $a_L = a_F$ y $\int_{a_F}^{\infty} \frac{1}{L} dF < 1$

entonces, globalmente en el intervalo $[a_F; b]$; donde $a_F < b < b_F$; se tiene que

$$\int_{a_F}^b F_n^t - F \rightarrow Z(t);$$

donde Z es el proceso

$$Z(t) = (1 - F(t)) \int_0^t C^{-2} (Y dL^a - T dF^a) + \frac{Y(t)}{C(t)}$$

siendo $(Y; T)$ el límite de $(Y_n; T_n)$ cuando $n \rightarrow \infty$; procesos definidos por

$$\begin{aligned} Y_n(t) &= \int_{a_F}^t [F_n^a(t) - F^a(t)] \\ T_n(t) &= \int_{a_F}^t [L_n^a(t) - L^a(t)]: \end{aligned}$$

Es la convergencia débil sobre intervalos compactos y fue probada por Woodroffe (1985).

Propiedad 1.22 Bajo la suposición de que las distribuciones F y L sean continuas y $a_L < a < a_F$ · $b < b_L$, se veri...ca globalmente en el intervalo $[a; b]$ que

$$P_{\frac{1}{n}} F_n^t \rightarrow F \text{ en } Z(t)$$

donde Z es un proceso Gaussiano de media 0 y función de covarianzas dada por

$$\text{Cov}(Z(y); Z(t)) = (1 - F(y))(1 - F(t)) \int_a^{y \wedge t} \frac{dF^\alpha(s)}{(C(s))^2}$$

con $y \wedge t = \min\{y, t\}$:

Esta propiedad de Wang, Jewell y Tsai (1986) proporciona una forma explícita de la estructura de covarianzas del proceso límite obtenido por Woodroffe (1985) y de...nido en la propiedad anterior.

Los resultados de consistencia y normalidad asintótica expuestos en las tres últimas propiedades pueden verse también (derivados desde un enfoque diferente) en el artículo de Keiding y Gill (1990).

Propiedad 1.23 Bajo la suposición de que las distribuciones F y L sean continuas e identi...cables se tiene:

a) si $a_L < a_F$, para cada $y = 0 \cdot y \cdot b < b_F$; se veri...ca

$$F_n^t(y) = F(y) + \frac{1}{n} \sum_{i=1}^n \omega_i(y) + r_n(y)$$

donde

$$\omega_i(y) = (1 - F(y)) \frac{1 - F(y_i)}{C(y_i)} \int_0^y \frac{1 - F(t)}{(C(t))^2} dF^\alpha(t) \tag{1.21}$$

y

$$\sup_{0 \cdot y \cdot b} |r_n(y)| = O\left(\frac{\ln n}{n}\right) \quad \text{c:s:}$$

b) si $a_L = a_F$; $\lim_{x \rightarrow 0^+} \frac{F(x)}{L(x)} = 0$ y $\int_0^1 \frac{1}{L} dF < 1$ entonces

$$F_n^t(y) = F(y) + \frac{1}{n} \sum_{i=1}^n \omega_i(y) + r_n(y)$$

donde $\omega_i(y)$ está de...nido según (1.21) y

$$\sup_{0 \cdot y \cdot b} |r_n(y)| = o\left(\frac{1}{n}\right) \quad \text{c:s:}$$

Este resultado, debido a Chao y Lo (1988) constituye una representación casi segura del estimador de Linden-Bell, paralela a la dada por Lo y Sing (1985) para el estimador de Kaplan-Meier. Como aplicación inmediata los autores derivan la convergencia débil del proceso $\sqrt{n}(F_n^t - F)$, propiedad que enunciamos a continuación:

Propiedad 1.24 Bajo las hipótesis de la propiedad anterior, en el intervalo $[0; b]$ con $b < b_F$ el proceso $\sqrt{n}(F_n^t - F)$ converge débilmente a un proceso Gaussiano de media cero y función de covarianzas

$$C(y, t) = \int_0^{\min(y, t)} (C(s))^{-2} dF^{\alpha}(s)$$

donde $y \wedge t = \min(y, t)$:

Propiedad 1.25 Bajo la suposición de que las distribuciones F y L sean continuas e idéntificables y que $\int_{a_F}^b \frac{1}{L^2} dF < 1$; se verifica, uniformemente en $a_F \leq y \leq b < b_F$; que

$$F_n^t(y) = F(y) + \frac{1}{n} \sum_{i=1}^{\infty} \eta_i(y) + r_n(y)$$

donde $\eta_i(y)$ está definido según (1.21) y

$$\sup_{a_F \leq y \leq b} |r_n(y)| = O\left(\frac{\mu \ln^3 n}{n}\right) \quad c.s.$$

Esta mejora en el orden del error para la representación casi segura se puede ver en Stute (1993a). La demostración tiene interés en si misma puesto que aplica resultados sobre U procesos obtenidos también por Stute (1993b).

Propiedad 1.26 Bajo la suposición de que las distribuciones F y L sean continuas e idéntificables y que $\int_{a_F}^b \frac{1}{L^3} dF < 1$; se verifica, uniformemente en $a_F \leq y \leq b < b_F$; que

$$F_n^t(y) = F(y) + \frac{1}{n} \sum_{i=1}^{\infty} \eta_i(y) + r_n(y)$$

donde $\eta_i(y)$ está definido según (1.21) y

$$\sup_{a_F \leq y \leq b} |r_n(y)| = O\left(\frac{\mu \ln \ln n}{n}\right) \quad c.s.$$

Arcones y Giné (1995) han obtenido esta representación con error de tasa óptima, a costa de imponer una condición un poco más restrictiva que la de Stute: $\int_{a_F}^b \frac{1}{L^3} dF < 1$:

1.5 El modelo de truncamiento aleatorio por la izquierda y censura aleatoria por la derecha.

Hasta ahora hemos definido dos modelos distintos: el modelo de truncamiento aleatorio por la izquierda y el modelo de censura aleatoria por la derecha, que se correspondían con problemas donde se observaban variables truncadas, por un lado, y censuradas, por otro. Sin embargo, no es difícil darse cuenta que en muchos problemas, sobre todo de Análisis de Supervivencia, hay una presencia de datos, que son, a la vez, truncados y censurados y que es necesario un nuevo modelo para estudiar esta situación. Un ejemplo ilustrativo de un problema de este tipo puede ser el siguiente:

- Se quiere estudiar el tiempo de duración, Y ; de una enfermedad determinada. Si T mide el tiempo desde que la enfermedad comienza en un individuo hasta que éste entra en estudio, el especialista sólo tendrá datos correspondientes a aquellos individuos que entren en el estudio, esto es, pacientes con $T_i \cdot Y_i$ (truncamiento por la izquierda). Una vez que un sujeto entra en estudio puede no observarse el tiempo completo de duración de su enfermedad, Y . Debido a distintas causas: fallecimiento por accidente, cambio de residencia, finalización del estudio, etc. se observa, tan sólo, un tiempo C menor que Y (censura por la derecha).

El modelo teórico que recoge esta situación, denominado modelo de truncamiento aleatorio por la izquierda y censura aleatoria por la derecha, es el siguiente:

Sea $(T_i; Y_i; C_i)$ $i = 1; 2; \dots; N$ una muestra aleatoria simple de un cierto vector aleatorio formado por tres variables $(T; Y; C)$, que supondremos no negativas, donde Y es la variable de interés, T es otra variable denominada variable de truncamiento y C es una tercera variable denominada de censura. Supongamos que uno sólo observa aquellos pares $(T_i; Z_i; \pm_i)$ para los cuales $T_i \cdot Z_i$; donde $Z_i = \min(Y_i; C_i)$ y $\pm_i = 1_{T_i \cdot C_i}$. Entonces n es aleatorio y N desconocido.

La muestra observada es de la forma:

$$\begin{array}{l} (T_i; Z_i; \pm_i) \quad \text{con } T_i \cdot Z_i \\ i = 1; 2; \dots; n \quad (n \cdot N) \end{array}$$

Además, condiciones habituales en este modelo son:

$$Y \text{ es independiente de } (T; C)$$

o

$$Y, T \text{ y } C \text{ son mutuamente independientes.}$$

Sobra decir que la segunda condición implica la primera.

8) La distribución de la variable de truncamiento, T , condicionada a que $T \leq Z$:
 $L^a(y) = P(T \leq y | T \leq Z)$:

Si $T; Y$ y C se suponen mutuamente independientes, L^a se puede escribir como:

$$L^a(y) = \frac{P(T \leq y; T \leq Z)}{P(T \leq Z)} = \int_0^Z y \frac{dL^a(t)}{1 - F(t)} dL(t) : \quad (1.24)$$

9) La función auxiliar $C(y) = P(T \leq y | T \leq Z)$ que, debido a la independencia entre las variables Y y C , se puede expresar mediante

$$C(y) = \int_0^y L(y) \frac{dF(t)}{1 - F(t)} : \quad (1.25)$$

Si también se supone independencia entre T y C , entonces

$$C(y) = \int_0^y L(y) \frac{dG(t)}{1 - G(t)} = \int_0^y L(y) \frac{dF(t)}{1 - F(t)} : \quad (1.26)$$

10) La función de azar de Y definida por

$$r(y) = \lim_{h \downarrow 0} \frac{P(y < Y \leq y + h | Y \leq y)}{h} = \frac{\frac{dF(y)}{dy}}{1 - F(y)} :$$

11) La función de azar acumulada, que se define a partir de la función anterior mediante la expresión:

$$\Lambda(y) = \int_0^y r(t) dt = \int_0^y \frac{dF(t)}{1 - F(t)}$$

y que, al igual que en modelos anteriores, jugará un papel fundamental para construir estimadores de la función de distribución de la variable Y ; al caracterizarla de forma única mediante la relación (análoga a (1.3) y (1.13)):

$$1 - F(y) = e^{-\Lambda_c(y)} \prod_{a_i \leq y} (1 - \Delta \Lambda_{fa_i g}) \quad (1.27)$$

$fa_i 2A = a_i \cdot yg$

donde Λ_c denota la parte continua de Λ , A el conjunto de puntos donde Λ tiene discontinuidades de salto y $\Delta \Lambda_{fa_i g} = \Lambda(a_i) - \Lambda(a_i^-)$ la magnitud del salto de Λ en a_i :

También es importante definir y analizar los siguientes conceptos:

12) Anteriormente hemos definido los extremos del soporte de una función de distribución (ver (1.4)). En el presente modelo tiene interés considerar las siguientes relaciones:

$$\begin{aligned} a_H &= \min a_F; a_{Gg} \\ b_H &= \min b_F; b_{Gg} \end{aligned} \quad (1.28)$$

y también:

$$\begin{aligned} a_H^* &= \max f_{a_H}; a_L^* \\ b_L^* &= \min f_{b_L}; b_H^* \end{aligned} \quad (1.29)$$

13) $\mathbb{P}(T \cdot Z)$ representa la probabilidad de ausencia de truncamiento, es decir, de la región del plano donde las variables $(T; Z)$ son observables, por tanto su probabilidad debe ser estrictamente positiva: $\mathbb{P} > 0$:

Si Y , T y C son mutuamente independientes, \mathbb{P} se puede expresar de las siguientes formas:

$$\begin{aligned} \mathbb{P} &= P(T \cdot Z) = \int_0^{Z_1} \int_0^{Z_t} dL(s)dH(t) = \int_0^{Z_1} L(t)dH(t) \\ \mathbb{P} &= P(T \cdot Z) = \int_0^{Z_1} \int_t^{Z_1} dH(s)dL(t) = \int_0^{Z_1} (1 - H(t)) dL(t) \end{aligned}$$

1.6 Estimación de la función de distribución con censura y truncamiento.

La primera referencia sobre un estimador no paramétrico de la función de distribución F para datos censurados y truncados se encuentra en la revista *Journal of the Royal Statistical Society* y está ...mada por Turnbull (1976), quien en su trabajo utiliza una idea sobre autoconsistencia para construir un algoritmo que obtiene el estimador de máxima verosimilitud de F :

Las primeras propiedades del estimador de Turnbull, que resulta ser un estimador de tipo límite producto, al igual que los de Kaplan-Meier y Linden-Bell, fueron estudiadas por Tsai, Jewell y Wang (1987), quienes derivaron su comportamiento asintótico, y por Lai y Ying (1991), que estudiaron la consistencia fuerte uniforme y la convergencia débil del proceso para una versión modificada del citado estimador bajo hipótesis más generales.

En cuanto a las representaciones casi seguras que descomponen al estimador en una suma de variables i.i.d. más un término de error, se deben a Gijbels y Wang (1993), Zhou (1996) y Zhou y Yip (1999).

A continuación, expondremos con detalle todas estas cuestiones.

1.6.1 El estimador de Turnbull: definición.

Llegaremos al estimador de Turnbull utilizando un procedimiento similar al usado por Woodroffe para estimar la función de distribución en el contexto de datos truncados. Por tanto, seguimos los siguientes pasos:

1. Expresar la función de azar acumulada Λ ; en función de cantidades empíricamente estimables:

$$\begin{aligned}\Lambda(y) &= \int_0^y \frac{dF(t)}{1 - F(t)} = \int_0^y \frac{P(T \cdot t \cdot C) dF(t)}{P(T \cdot t \cdot C)(1 - F(t))} \\ &= \int_0^y \frac{dH_1^a(t)}{C(t)}.\end{aligned}\quad (1.30)$$

Se han utilizado las expresiones (1.22) y (1.25).

Una consideración importante es que las igualdades anteriores son ciertas si $P(T \cdot t \cdot C) > 0$; $8t \in [a_F; y)$: Si se trabaja bajo la hipótesis de que Y , T y C sean mutuamente independientes, esto se consigue imponiendo que

$$a_L \cdot a_H \leq y < b_H = \min\{b_F; b_G\}$$

que constituyen las condiciones de identificabilidad de F .

2. Definir un estimador de la función Λ :

$$\Lambda_n^{tc}(y) = \int_0^y \frac{dH_{1n}^a(t)}{C_n(t)} = \sum_{i=1}^n \frac{1_{fZ_i \cdot y; \pm_i = 1g}}{nC_n(Z_i)} = \sum_{i=1}^n \frac{1_{fZ_i \cdot y; \pm_i = 1g}}{\prod_{j=1}^n 1_{fT_j \cdot Z_i \cdot Z_j g}};\quad (1.31)$$

Para ello hemos sustituido en la expresión (1.30) H_1^a y C por sus correspondientes estimadores empíricos H_{1n}^a y C_n , dados por

$$\begin{aligned}H_{1n}^a(y) &= \sum_{i=1}^n 1_{fZ_i \cdot y; \pm_i = 1g} \frac{1}{n} \\ C_n(y) &= \sum_{i=1}^n 1_{fT_i \cdot y \cdot Z_i g} \frac{1}{n}.\end{aligned}$$

3. Definir un estimador de la función F , a partir de la relación dada por la ecuación (1.27). Como el estimador Λ_n^{tc} es una función en escalera con saltos en los puntos $fZ_i \cdot \pm_i = 1g$, dicha relación sugiere estimar F de forma que:

$$1 - F_n^{tc}(y) = \prod_{i=1}^n \frac{1_{fZ_i \cdot y; \pm_i = 1g}}{nC_n(Z_i)} = \prod_{i=1}^n \frac{1_{fZ_i \cdot y; \pm_i = 1g}}{\prod_{j=1}^n 1_{fT_j \cdot Z_i \cdot Z_j g}};\quad (1.32)$$

estimador conocido con el nombre de estimador de Turnbull.

Al igual que el estimador de Linden-Bell, en el caso de que se consideren empates, la expresión anterior se convierte en:

$$\hat{F}_n^{tc}(y) = \prod_{i=1}^n \frac{e_i}{\sum_{j=1}^n 1_{f_{T_j} \cdot Z_i \cdot Z_j} g}$$

siendo $e_i = \prod_{j=1}^n 1_{f_{(Z_j; \pm_j)} = (Z_i; 1)} g$.

1.6.2 El estimador de Turnbull: propiedades.

Todas las propiedades que se exponen a continuación parten de las hipótesis ya mencionadas de independencia entre Y y $(T; C)$ así como $a_L \cdot a_H$:

Propiedad 1.27 El estimador de Turnbull coincide:

- con el estimador de Kaplan-Meier en el caso de ausencia de truncamiento.
- con el estimador de Linden-Bell en el caso de ausencia de censura.
- con la función de distribución empírica en el caso de ausencia de censura y truncamiento.

A partir de la expresión (1.32) del estimador y utilizando que cuando no hay truncamiento todos los $f_{T_i} g_{i=1}^n$ toman el valor 0 y las $f_{(Z_i; \pm_i)} g_{i=1}^n$ constituyen una m.a.s. de tamaño n de toda la variable $(Z; \pm)$ se obtiene el apartado a):

$$\hat{F}_n^{tc}(y) = \prod_{i=1}^n \frac{1_{f_{Z_i \cdot y; \pm_i = 1} g}}{\sum_{j=1}^n 1_{f_{T_j} \cdot Z_i \cdot Z_j} g} = \prod_{i=Z_i: n \cdot y} \frac{1_{\pm[i:n]}}{n_i (i-1)}$$

(estimador de Kaplan-Meier dado por la ecuación (1.6)).

El apartado b) se prueba partiendo de la expresión del estimador de Turnbull (ecuación (1.32)) y utilizando que cuando no hay censura la muestra $(T_i; Z_i; \pm_i) = (T_i; Y_i; 1)$; $\delta_i = 1; \dots; n$: Así, se llega al estimador de Linden-Bell (expresión (1.17)):

$$\hat{F}_n^{tc}(y) = \prod_{i=1}^n \frac{1_{f_{Z_i \cdot y; \pm_i = 1} g}}{nC_n(Z_i)} = \prod_{i=1}^n \frac{1_{f_{Y_i \cdot y} g}}{nC_n(Y_i)}$$

Para probar c) se procede de forma análoga. Utilizando que $T_i = 0$, $(Z_i; \pm_i) = (Y_i; 1)$; $\delta_i = 1; \dots; n$ y n no es aleatorio, sino el tamaño del total de la muestra, se

Llega a:

$$\begin{aligned} 1_i F_n^{tc}(y) &= \prod_{i=1}^n \frac{\tilde{A}}{1_i \prod_{j=1}^n \frac{1_{fY_i \cdot yg}}{1_{fY_i \cdot Y_j g}}} = \prod_{i=Y_{i:n} \cdot y}^n \frac{1}{1_i (i-1)} = \\ &= \prod_{i=Y_{i:n} \cdot y}^n \frac{\mu}{n_i i} = \frac{n_i}{n} \frac{1}{n_i} \frac{2}{n_i} \frac{3}{n_i} \dots \frac{k}{n_i} = \\ &= \frac{n_i}{n} = 1_i \prod_{i=1}^n \frac{1}{n} 1_{fY_{i:n} \cdot yg} \end{aligned}$$

donde $k = \max_{i=Y_{i:n} \cdot yg}$ (evidentemente $k = \prod_{i=1}^n 1_{fY_{i:n} \cdot yg}$):

Propiedad 1.28 El estimador de Turnbull tiene forma escalonada con saltos en las observaciones no censuradas. La magnitud del salto en Z_i , $dF_n^{tc}(Z_i)$; es igual a

$$dF_n^{tc}(Z_i) = \frac{\pm_i}{nC_n(Z_i)} 1_i F_n^{tc}(Z_i)^\pm :$$

Esta expresión se obtiene al calcular $F_n^{tc}(Z_i) - F_n^{tc}(Z_i^-)$ utilizando (1.32):

$$\begin{aligned} F_n^{tc}(Z_i) - F_n^{tc}(Z_i^-) &= 1_i F_n^{tc}(Z_i^-) - 1_i F_n^{tc}(Z_i)^\pm = \\ &= \prod_{j=1}^n \frac{\tilde{A}}{1_i \frac{1_{fZ_j < Z_i; \pm_j = 1g}}{nC_n(Z_i^-)}} - \prod_{j=1}^n \frac{\mu}{1_i \frac{1_{fZ_j \cdot Z_i; \pm_j = 1g}}{nC_n(Z_i)}} = \\ &= \prod_{j=1}^n \frac{\tilde{A}}{1_i \frac{1_{fZ_j < Z_i; \pm_j = 1g}}{nC_n(Z_i^-)}} \left(1_i \left(1 + \frac{\pm_i}{nC_n(Z_i)} \right) - 1_i \right) = \frac{1_i F_n^{tc}(Z_i^-) \pm_i}{nC_n(Z_i)} \end{aligned}$$

Propiedad 1.29 El estimador de Turnbull es el estimador no paramétrico de máxima verosimilitud de F ; siempre que $\prod_{j=1}^n 1_{fT_j \cdot Z_i \cdot Z_j g} - \prod_{j=1}^n 1_{f(Z_j; \pm_j) = (Z_i; 1)g} > 0$; para todo los valores Z_i con $\pm_i = 1$; excepto el mayor.

Este resultado se puede ver en Wang (1987).

Propiedad 1.30 Si las variables T y C son independientes, el estimador de tipo límite producto para la función de distribución H de la variable $Z = \min\{Y; C\}$ viene dado por

$$1_i H_n^{tc}(y) = \prod_{i=1}^n \frac{\mu}{1_i \frac{1_{fZ_i \cdot yg}}{nC_n(Z_i)}} : \tag{1.33}$$

Estamos trabajando bajo la hipótesis de que las variables T , C e Y son mutuamente independientes. Por tanto $Z = \min fY; Cg$ es independiente de T y $a_L \cdot a_H$. Estas hipótesis constituyen las condiciones de identi...cabilidad de H bajo truncamiento (ver (1.15); el papel de F ahora lo juega H) y por tanto el estimador de H se obtiene sustituyendo en la expresión del estimador de Linden-Bell (ecuación (1.17)) las variables Y_i por las Z_i :

Propiedad 1.31 Si $b_L \cdot b_H$ y las variables T y C son independientes, el estimador de tipo límite producto para la función de distribución L de la variable de truncamiento T viene dado por

$$L_n^{tc}(y) = \prod_{i=1}^n \frac{1_{fT_i > yg}}{nC_n(T_i)} = \prod_{i=1}^n \frac{1_{fT_i > yg}}{\prod_{j=1}^n 1_{fT_j \cdot T_i \cdot Z_j}} \quad (1.34)$$

Las hipótesis de trabajo implican que $Z = \min fY; Cg$ es independiente de T y $b_L \cdot b_H$: Estas hipótesis constituyen las condiciones de identi...cabilidad de L bajo truncamiento (ver (1.19); el papel de Y ; con función de distribución F ; ahora lo juega Z ; con función de distribución H) y por tanto el estimador de L se obtiene sustituyendo en la expresión del estimador de Linden-Bell de L (ecuación (1.18)) las variables Y_i por las Z_i :

Propiedad 1.32 Si las variables T y C son independientes, el estimador de tipo límite producto para la función de distribución G de la variable C viene dado por

$$\prod_{i=1}^n G_n^{tc}(y) = \prod_{i=1}^n \frac{1_{fZ_i \cdot y; \pm_i = 0g}}{nC_n(Z_i)} \quad \text{para } y < b_H: \quad (1.35)$$

El proceso es análogo al seguido para obtener el estimador F_n^{tc} ; consistente en encontrar una expresión que caracterice G y que dependa de cantidades empíricamente estimables. Dicha expresión viene dada por $\Lambda_G(y) = \int_0^y \frac{dG(t)}{1 - G(t)}$ que determina la distribución G de forma única mediante la relación

$$\prod_{i=1}^n G(y) = e^{\int_0^y \Lambda_G^c(t) dt} \prod_{a_i \in A} \Lambda_G^d(a_i) \quad (1.36)$$

$f_{a_i} 2A = a_i \cdot yg$

donde Λ_G^c denota la parte continua de Λ_G , A el conjunto de puntos donde Λ_G tiene discontinuidades de salto y $\Lambda_G^d(a_i) = \Lambda_G(a_i) - \Lambda_G^c(a_i)$ la magnitud del salto de Λ_G en a_i :

Además, $\int_0^y \frac{dG(t)}{1 - G(t)}$ se puede escribir como

$$\int_0^y \frac{dG(t)}{1 - G(t)} = \int_0^y \frac{\int_0^t 1_{L(t)}(1 - F(t^-)) dG(t)}{\int_0^t 1_{L(t)}(1 - F(y^-)) (1 - G(t^-))} = \int_0^y \frac{dH_0^a(t)}{C(t)}$$

sin más que utilizar las expresiones (1.23) y (1.26).

Sustituyendo, en el último término de la igualdad anterior, H_0^a y C por sus estimadores empíricos y utilizando el resultado (1.36) se obtiene el estimador $G_n^{tc}(y)$; para $y < b_H = \min\{b_F, b_G\}$:

Propiedad 1.33 El producto de los estimadores de las funciones $1_i F$ y $1_i G$, $1_i F_n^{tc}$ y $1_i G_n^{tc}$ respectivamente, es igual al estimador de la función $1_i H$, $1_i H_n^{tc}$:

$$1_i F_n^{tc} \cdot 1_i G_n^{tc} = 1_i H_n^{tc}.$$

Basta multiplicar las expresiones (1.32) y (1.35) y tener en cuenta que, para cada i desde 1 hasta n ; se verifica que $\pm_i = 0$ o $\pm_i = 1$. Entonces:

$$\begin{aligned} 1_i F_n^{tc}(y) \cdot 1_i G_n^{tc}(y) &= \prod_{i=1}^n \frac{1_{f_{Z_i \cdot y; \pm_i=1g}}}{n C_n(Z_i)} \prod_{i=1}^n \frac{1_{f_{Z_i \cdot y; \pm_i=0g}}}{n C_n(Z_i)} \\ &= \prod_{i=\pm_i=1}^n \frac{1_{f_{Z_i \cdot yg}}}{n C_n(Z_i)} \prod_{i=\pm_i=0}^n \frac{1_{f_{Z_i \cdot yg}}}{n C_n(Z_i)} \\ &= \prod_{i=1}^n \frac{1_{f_{Z_i \cdot yg}}}{n C_n(Z_i)} = 1_i H_n^{tc}(y): \end{aligned}$$

Propiedad 1.34 Supongamos que $S = 1_i F$ es continua y la función de distribución conjunta de $(T; C)$ es continua. Además $P(T > a) > 0$ y $P(C > b) > 0$: Denotando por $S^a(y) = P(Y > y | Y > a)$ y

$$S_n^{atc}(y) = \prod_{i=Z_i > a}^n \frac{1_{f_{Z_i < y; \pm_i=1g}}}{\prod_{j=1}^n 1_{f_{T_j \cdot Z_j \cdot g}}}$$

se verifica, globalmente en el intervalo $[a; b]$; que:

$$P_n \{ S_n^{atc} - S^a \} \xrightarrow{d} Z(t)$$

donde Z es un proceso Gaussiano de media cero y función de covarianzas dada por

$$\text{Cov}(Z(y); Z(t)) = S^a(y)S^a(t) \int_a^{y \wedge t} \frac{dH_1^a(s)}{(C(s))^2}$$

con $y \wedge t = \min\{y, t\}$:

Esta propiedad de convergencia débil sobre intervalos compactos fue probada por Tsai, Jewell y Wang (1987).

Propiedad 1.35 Se consideran las funciones $C_{0N}(s) = \frac{1}{N} \prod_{j=1}^N 1_{F_{T_j} \cdot s \cdot C_j} g$ y $C_{0N}^a(s) = \frac{1}{N} \prod_{j=1}^N 1_{F_{T_j} \cdot C_j < sg}$; donde $(T_i; Y_i; C_i) \ i = 1; 2; \dots; N$ es el vector aleatorio (de datos completos (no observables)) que genera la muestra observada (de tamaño n) mediante truncamiento. Supongamos que $\liminf_{N \rightarrow \infty} \frac{1}{N} \int_0^1 (C_{0N} + C_{0N}^a) dF$ es estrictamente positivo.

Si denotamos por $F(y|u) = P(Y \cdot y|Y > u)$; por

$$F_{n;u}^{atc}(y) = \prod_{i=Z_i > u} \frac{1_{F_{Z_i} \cdot y; \pm i = 1} g}{\prod_{j=1}^n 1_{F_{T_j} \cdot Z_i \cdot Z_j} g}$$

y $Z_{1:n}; Z_{2:n}; \dots; Z_{n:n}$ es la muestra ordenada de los valores observados $fZ_i g_{i=1}^n$; entonces, para cada $c > 0$; $0 < \epsilon < 1$ y $0 < \delta < 1$; se tiene

$$\sup_{0 < y < 1} |F_{n;u}^{atc}(y) - F(y|u)| = O(n^{-c}) \cdot \min_{u < s < y} C_n(s) \cdot n^{-\delta} \rightarrow 0$$

donde $[\cdot]$ es la parte entera de \cdot :

Lai y Ying (1991) son los autores de esta propiedad de consistencia fuerte uniforme.

Propiedad 1.36 Bajo la suposición de que la distribución F sea continua y si $a_L < a_H$, se verifica, para cada $y = a_H \cdot y \cdot b < b_H$; que

$$F_n^{tc}(y) = F(y) + \frac{1}{n} \sum_{i=1}^n \kappa_i(y) + r_n(y)$$

donde

$$\kappa_i(y) = (1 - F(y)) \cdot \frac{1_{fZ_i \cdot y; \pm i = 1} g}{C(Z_i)} \int_0^y \frac{1_{fT_i \cdot t \cdot Z_i} g}{(C(t))^2} dH_1^a(t) \tag{1.37}$$

y

$$\begin{aligned} \sup_{0 < y < b} |r_n(y)| &= O\left(\frac{\mu_{\ln n}}{n}\right) \quad c:s: \\ E \sup_{0 < y < b} |r_n(y)| &= O(n^{-\epsilon}) \quad ; \text{ para } \epsilon > 0: \end{aligned}$$

Este resultado, obtenido por Gijbels y Wang (1993), constituye una representación casi segura del estimador de Turnbull, bajo la hipótesis de $a_L < a_H$; que logra un orden $\ln n = n$ para el término de error. Esto les permite aplicar el resultado al estudio de estimadores no paramétricos de las funciones de densidad y de razón de fallo.

Una aplicación más inmediata de la representación casi segura es la demostración de la convergencia débil del proceso $n^{1/2} \int F_n^{tc}(y) - F(y)$, propiedad que enunciamos a continuación:

Propiedad 1.37 En las hipótesis de la propiedad anterior, el proceso $n^{1/2} \int F_n^{tc} - F$ converge débilmente a un proceso Gaussiano de media cero y función de covarianzas

$$(1 - F(y))(1 - F(t)) \int_0^{y \wedge t} (C(s))^{-2} dH_1^{\alpha}(s)$$

en el intervalo $[0; b]$ con $b < b_H$:

Propiedad 1.38 Si las variables T y C son independientes, las distribuciones F , G y H son continuas, y se verifica alguna de las dos condiciones siguientes:

- i) $a_L < a_H$ y $\int_{a_H}^{\infty} \frac{1}{C^3} dH_1 < 1$
 - ii) $a_L = a_H$ y $\int_{a_H}^{\infty} \frac{1}{C^3} dH_1 < 1$
- entonces, se tiene

$$F_n^{tc}(y) = F(y) + \frac{1}{n} \sum_{i=1}^{\infty} \eta_i(y) + r_n(y)$$

uniformemente en $a_H \cdot y \cdot b < b_H$; donde $\eta_i(y)$ está definido según (1.37) y

$$\sup_{a_H \cdot y \cdot b} |r_n(y)| = O(n^{-1} \ln^{1+\epsilon} n) \quad \text{c.s. para } \epsilon > \frac{1}{2}$$

Esta representación casi segura, demostrada por Zhou (1996), extiende la obtenida por Gijbels y Wang (1993) al caso $a_L = a_H$ y representa una generalización de la representación dada por Stute para el estimador límite producto de la función de distribución para datos truncados al contexto más amplio de censura y truncamiento, que además, mejora el orden del error.

Recientemente, Zhou y Yip (1999) han mejorado este último resultado obteniendo un orden óptimo para el error, dado por $O(n^{-1} \ln \ln n)$ c.s. Además, en dicho resultado no imponen la hipótesis de independencia entre las variables T y C .

Capítulo 2

Estimación no paramétrica de la función de distribución condicional en presencia de censura, truncamiento y censura con truncamiento.

En múltiples ocasiones interesa explicar la relación existente entre una variable respuesta Y y una variable explicativa o covariable X . En este contexto, uno puede plantearse estimar la distribución conjunta, la curva de regresión o la función de distribución de Y condicionada a X , entre otras características relevantes. Además, la estimación de las funciones anteriores puede llevarse a cabo desde dos ópticas distintas: la paramétrica y la no paramétrica.

En este capítulo estudiaremos la estimación no paramétrica de la función de distribución de Y condicionada a X en presencia de censura, truncamiento y censura con truncamiento. Si bien previamente expondremos el contexto de trabajo y algunas propiedades relevantes de la estimación no paramétrica de la función de distribución condicional para datos completos, que más tarde serán necesarias en la demostración de nuevos resultados.

2.1 Estimación no paramétrica de la función de distribución condicional.

Supondremos un contexto de trabajo en el cual X e Y son variables aleatorias (diseño aleatorio), continuas y unidimensionales, y se dispone de una muestra observada $f(X_i; Y_i)_{i=1}^n$ i.i.d. según $(X; Y)$: Se quiere estimar la función de distribución de Y condicionada a X ; denotada por $F(y | x) = P(Y \leq y | X = x)$:

Para abordar esta estimación desde el punto de vista no paramétrico, se suele utilizar la propiedad de que la función de distribución condicional es la esperanza condicionada a X de la variable indicadora 1_{FY} . yg , esto es

$$F(y | x) = E[1_{FY} | X = x];$$

lo cuál permite estimar $F(y | x)$ mediante las técnicas usuales de estimación no paramétrica de la curva de regresión, sin más que tomar como variable dependiente la variable indicadora 1_{FY} .

Ya hemos comentado en la introducción de esta memoria que entre los procedimientos de estimación no paramétricos disponibles utilizaremos el tipo núcleo con pesos de Nadaraya (1964) y Watson (1964) por ser uno de los más apropiados en el contexto de diseño aleatorio y por disponer de resultados teóricos de interesante aplicación.

Concretamente, dada una m.a.s. $f(X_i; Y_i)_{i=1}^n$; el estimador de Nadaraya-Watson de la función de regresión $g(x) = E(Y | X = x)$ viene dado por:

$$\hat{g}_n(x) = \frac{\sum_{i=1}^n K\left(\frac{x_i - x}{h}\right) Y_i}{\sum_{i=1}^n K\left(\frac{x_i - x}{h}\right)} = \sum_{i=1}^n B_{ni}(x) Y_i$$

donde K es una función núcleo (función real, continua, acotada, simétrica y que integra 1) y h el conocido parámetro de suavización o ventana de la estimación no paramétrica.

Intuitivamente, $\hat{g}_n(x)$ no es más que un promedio local, es decir, $g(x)$ se estima por una media ponderada de valores Y_i ; donde las ponderaciones, $B_{ni}(x)$; tienen en cuenta la distancia entre los valores de X_i y x , y el parámetro de suavizado, h ; es el encargado de regular el tamaño del entorno de x donde se pondera. Un entorno demasiado amplio producirá una curva "sobresuavizada", dando lugar a una estimación de g sesgada. Por el contrario, si el entorno es demasiado pequeño se producirá una "infrasuavización" o estimación de excesiva variabilidad que se traduce en una curva \hat{g}_n con numerosos picos. Aquí se pone de mani...esto el importante problema de la selección del parámetro de suavizado, para tratar de encontrar un equilibrio entre sobre e infrasuavización.

Obviamente, el estimador tipo núcleo de la función de distribución condicional $F(y | x) = P(Y \leq y | X = x)$ vendrá dado por:

$$\hat{F}_n(y | x) = \frac{\sum_{i=1}^n K\left(\frac{x_i - x}{h}\right) 1_{F_{Y_i} \leq y}}{\sum_{j=1}^n K\left(\frac{x_i - x_j}{h}\right)} = \sum_{i=1}^n B_{ni}(x) 1_{F_{Y_i} \leq y} \quad (2.1)$$

sin más que tener en cuenta que $F(y | x) = E[1_{F_Y \leq y} | X = x]$:

Algunas propiedades importantes del estimador de Nadaraya-Watson para la función de distribución condicional se detallan a continuación. En todas ellas m denota la densidad marginal de la variable X ; que se supone existente.

Propiedad 2.39 Bajo las hipótesis:

- i) $m(x)$ y $F(y | x)$ son funciones continuas en x ;
 - ii) $\int K(t) dt < 1$ y $\lim_{t \rightarrow \pm 1} tK(t) = 0$;
 - iii) $h \rightarrow 0$ y $nh \rightarrow 1$ cuando $n \rightarrow \infty$;
- se verifica que:

$$\hat{F}_n(y | x) \xrightarrow{P} F(y | x) \quad \text{para cada } y \in \mathbb{R}$$

Esta propiedad de consistencia se puede ver en Härdle (1990) y es un caso particular de la consistencia del estimador de la función de regresión g ; tomando $g(x) = E[1_{F_Y \leq y} | X = x]$.

Propiedad 2.40 Suponiendo que:

- i) $m(x)$ y $F(\cdot | x)$ tienen derivadas de segundo orden continuas en un entorno de x y uniformemente acotadas.
- ii) K es una densidad simétrica acotada de media cero con soporte compacto.
- iii) $h \rightarrow 0$ y $nh \rightarrow 1$ cuando $n \rightarrow \infty$;

Entonces, en todo punto x tal que $m(x) > 0$; se verifica

$$\sup_{y \in \mathbb{R}} |\hat{F}_n(y | x) - F(y | x)| = O_p((nh)^{-1/2}) + O(h^2)$$

Es el Teorema 3.2.2 de la Tesis de Presedo Quindimil (1991).

Propiedad 2.41 Bajo las hipótesis:

- i) m es acotada y existe un intervalo $I = [x_1; x_2]$ contenido en el soporte de m tal que

$$\inf_{x \in I} m(x) > 0;$$

Además, m es uniformemente continua sobre un intervalo abierto que contenga a I :

ii) $F(y|x)$ es Lipschitz continua en x , uniformemente en y , esto es: existe una constante $C > 0$ tal que

$$\sup_{y \in \mathbb{R}} |F(y|x) - F(y|x^0)| \leq C |x - x^0|$$

iii) La función núcleo, K , es una función de densidad Lipschitz continua de orden 1; y existen constantes positivas: L, e_1, e_2 de forma que $e_1 \cdot \sup_{|t| \leq L} |K(t)| \leq e_2$:

iv) $h \rightarrow 0$ y $\ln n = (nh) \rightarrow 0$ cuando $n \rightarrow \infty$:

Se verifica que

$$\sup_{y \in \mathbb{R}; x \in I} |\hat{F}_n(y|x) - F(y|x)| = O\left(\frac{1}{(\ln n)(nh)}\right)^{1/2} + O(h) \quad \text{c.s.}$$

Este resultado de convergencia fuerte uniforme se deriva de Cheng y Cheng (1987).

Propiedad 2.42 Si se verifican las condiciones:

h1) Sea $I = [x_1; x_2]$ un intervalo contenido en el soporte de m verificando que

$$0 < \gamma = \inf_{x \in I_\pm} [m(x)] < \sup_{x \in I_\pm} [m(x)] = \Gamma < 1$$

para algún entorno $I_\pm = [x_1 \pm \epsilon; x_2 \pm \epsilon]$ con $\epsilon > 0$ y $0 < \pm \Gamma < 1$:

h2) Las funciones $m(x)$ y $F(y|x)$ tienen primera derivada respecto a x continua y acotada en $(y; x) \in \mathbb{R} \times I_\pm$:

h3) Las funciones $m(x)$ y $F(y|x)$ tienen segunda derivada respecto a x continua y acotada en $(y; x) \in \mathbb{R} \times I_\pm$:

h4) La función núcleo, K , es una función de densidad simétrica, que vale 0 fuera del intervalo $(-1; 1)$ y su variación total está acotada por $\int_{-1}^1 K(t) dt < +\infty$.

h5) El parámetro ventana $h = h_n$ verifica: $h \rightarrow 0$ y $\ln n = (nh) \rightarrow 0$:

Entonces:

$$\sup_{y \in \mathbb{R}; x \in I} |\hat{F}_n(y|x) - F(y|x)| = O\left(\frac{1}{(\ln n)(nh)}\right)^{1/2} + O(h^2) \quad \text{c.s.}$$

Si además imponemos $nh^5 = \ln n = O(1)$ se tiene:

$$\sup_{y \in \mathbb{R}; x \in I} |\hat{F}_n(y|x) - F(y|x)| = O\left(\frac{1}{(\ln n)(nh)}\right)^{1/2} \quad \text{c.s.}$$

Resultado probado por Dabrowska (1989) para funciones de supervivencia condicionales de variables no negativas: $S(y|x) = 1 - F(y|x)$; en un contexto de censura, pero que se verifica también para cualquier función de distribución bajo las hipótesis enunciadas.

Este resultado de convergencia fuerte uniforme será clave en la demostración de posteriores resultados novedosos y por ello detallamos su demostración (para funciones de supervivencia de v.a. no negativas):

Sea $C(u; v)$, $u = (u_1; u_2; \dots; u_{q_1})$; $v = (v_1; v_2; \dots; v_{q_2})$; $q_1 \geq 1$; $q_2 \geq 0$; una función de distribución conjunta sobre $[0; 1]^{q_1+q_2}$ con marginales uniformes y sea C_n la correspondiente función de distribución empírica para una muestra de tamaño n . Para cualquier rectángulo $R \subseteq [0; 1]^{q_1+q_2}$ tenemos $C(R) = \int_R dC(u; v)$ y $C_n(R) = \int_R dC_n(u; v)$ y podemos considerar la diferencia de ambas medidas que denotaremos $w_n(R) = C_n(R) - C(R)$:

Consideremos ahora rectángulos de la forma $R_{u;v} = \prod_{i=1}^{q_1} [u_{i1}; u_{i2}] \times \prod_{j=1}^{q_2} [v_{j1}; 1]$ y sean

$$w_n(a) = \sup_{\substack{0 \leq v_{j1} \leq 1; \\ u_{i1} \leq u_{i2} \leq a_i}} |w_n(R_{u;v})|$$

$$\mu_n(a) = \sup_{\substack{0 \leq v_{j1} \leq 1; \\ u_{i1} \leq u_{i2} \leq a_i}} C(R_{u;v})$$

donde $a = (a_1; a_2; \dots; a_{q_1}) \in [0; 1]^{q_1}$:

Dabrowska prueba (generalizando un resultado de Stute) el siguiente resultado:

Sea $s > 0$ tal que $2^{-\mu_n(a)} \leq e_0^{-\mu_n(a) s}$ donde $e_0 < +1$, entonces existen constantes $c_1; c_2 > 0$ verificando:

$$P(w_n(a) > s) \leq \frac{c_1}{(\min a_i)^{q_1}} \exp \left\{ -\frac{c_2 s^2}{\mu_n(a)^{3/4}} \right\} \quad (2.2)$$

Volvamos ahora a nuestro problema:

Tenemos Y una v.a. no negativa, con función de supervivencia condicionada a $X = x$, $S(t|x) = 1 - F(t|x) = P(Y > t | X = x)$ y supongamos Y y X continuas con distribución conjunta $M(t; x)$.

Utilizando la transformación cuantil obtenemos $M(t; x) = C(M_1(t); M_2(x))$, siendo $M_1; M_2$ las funciones de distribución marginales de M . Las distribuciones marginales de C son uniformes en $[0; 1]$: Para las distribuciones empíricas obtenemos, con probabilidad 1, $M_n(t; x) = C_n(M_1(t); M_2(x))$ y por tanto $M_n(t; x) - M(t; x) = C_n(M_1(t); M_2(x)) - C(M_1(t); M_2(x))$ con probabilidad 1. Si ahora consideramos los intervalos

$$R = [t_1; t_2] \times [x_1; x_2] \quad \text{y} \quad T = [M_1(t_1); M_1(t_2)] \times [M_2(x_1); M_2(x_2)]$$

y sus respectivas medidas teóricas y empíricas obtendremos, con probabilidad 1, que

$$\hat{F}_n(R) = M_n(R) \quad ; \quad M(R) = C_n(T) \quad ; \quad C(T) = \hat{F}_n(T):$$

Por otro lado la expresión del estimador núcleo con pesos Nadaraya-Watson de $S(t|x)$ se puede escribir como $\hat{S}_n(t|x) = \hat{m}_n^{-1}(x)\hat{A}_n(t;x)$; donde

$$\hat{A}_n(t;x) = \frac{1}{nh} \sum_{i=1}^n 1_{f_{Y_i} > t} K \left(\frac{X_i - X}{h} \right) = \frac{1}{h} \int_{f_{Y_i} > t} K \left(\frac{X_i - u}{h} \right) dM_n(y; u)$$

y

$$\hat{m}_n(x) = \frac{1}{nh} \sum_{j=1}^n K \left(\frac{X_j - X}{h} \right) = \hat{A}_n(0;x)$$

$\hat{m}_n(x)$ es el conocido estimador de Parzen-Rosenblatt de la densidad de X (Rosenblatt (1956), Parzen (1962)).

También ocurre que

$$\begin{aligned} S(t|x) &= \int_{f_{Y_i} > t} m(y|x) dy = \int_{f_{Y_i} > t} \frac{m(y;x)}{m(x)} dy = \\ &= \frac{\int_{f_{Y_i} > t} m(y;x) dy}{m(x)} = \frac{A(t;x)}{m(x)} \end{aligned}$$

donde

$$A(t;x) = \int_{f_{Y_i} > t} m(y;x) dy = S(t|x) m(x)$$

y

$$m(x) = A(0;x):$$

Observemos que

$$\begin{aligned} E\hat{A}_n(t;x) &= E \left[\frac{1}{nh} \sum_{i=1}^n 1_{f_{Y_i} > t} K \left(\frac{X_i - X}{h} \right) \right] = \frac{1}{h} E \left[\int_{f_{Y_i} > t} K \left(\frac{X_i - X}{h} \right) \right] = \\ &= \frac{1}{h} \int_{f_{Y_i} > t} K \left(\frac{X_i - u}{h} \right) dM(y; u): \end{aligned}$$

Entonces

$$\begin{aligned} & \hat{A}_n(t; x) - E\hat{A}_n(t; x) \\ &= \frac{1}{h} \int_{x_i}^{x_i+h} \int_{t}^{t+h} \frac{\mu_{x_i, u}}{h} dM_n(y; u) - \frac{1}{h} \int_{x_i}^{x_i+h} \int_{t}^{t+h} \frac{\mu_{x_i, u}}{h} dM(y; u) = \\ &= \frac{1}{h} \int_{x_i}^{x_i+h} \int_{t}^{t+h} \frac{\mu_{x_i, u}}{h} [dM_n(y; u) - dM(y; u)] \end{aligned} \quad (2.3)$$

y utilizando que las variables son continuas y haciendo integración por partes, obtenemos que

$$\begin{aligned} (2.3) &= \frac{1}{h} \int_{x_i}^{x_i+h} \int_{t}^{t+h} \frac{\mu_{x_i, u}}{h} 1_{fy>tg} [m_n(y; u) - m(y; u)] dy du = \\ &= \frac{1}{h} \int_{x_i}^{x_i+h} \int_{t}^{t+h} \frac{\mu_{x_i, u}}{h} (m_n(y; u) - m(y; u)) dy du \end{aligned}$$

donde B representa el intervalo comprendido entre x y u (puesto que $\int_x^u f(t)dt$ es una primitiva de f(u), \int_x^u se restringe al intervalo $(x - h; x + h)$).

Llamando $R(t; x; u) = [t; t+h] \cap [\min(x; u); \max(x; u)]$ obtenemos:

$$\begin{aligned} & \hat{A}_n(t; x) - E\hat{A}_n(t; x) \\ &= \frac{1}{h} \int_{x_i}^{x_i+h} \int_{t}^{t+h} \frac{\mu_{x_i, u}}{h} R_n(t; x; u) dK \frac{\mu_{x_i, u}}{h} = \\ &= \frac{1}{h} \int_w^w \int_{x_i}^{x_i+h} \frac{\mu_{x_i, M_2^{-1}(M_2(x) + w)}}{h} dK \end{aligned} \quad (2.4)$$

donde $w \in [M_2(x - h); M_2(x); M_2(x + h); M_2(x)]$.

Además, si $x \in I$;

$$(2.4) \cdot \sup_{T \in J} |j_n^{\otimes}(T)| \frac{1}{h}$$

donde

$$J = fT = [s; 1] \cap [v_1; v_2] = [0; s-1] \cup [v_1; v_2] \cup [2\Gamma h; \dots]$$

puesto que

$$\begin{aligned} & |j_{v_1; v_2}| \cdot |j_{M_2(x) + w(\max); M_2(x) + w(\min)}| = \\ &= |j_{M_2(x+h); M_2(x)}; M_2(x-h); M_2(x)}| \\ &= |j_{M_2(x+h); M_2(x-h)}| = |j_{M_2(x-h); M_2(x+h)}| = 2\Gamma h \end{aligned}$$

Llamando $k_k = \sup_{j \in I} |j|$, obtenemos

$$\hat{A}_n \stackrel{\circ}{=} E \hat{A}_n \stackrel{\circ}{=} \sup_{T \cup J} j^{\otimes n}(S) j \frac{\circ}{h}$$

y utilizando el resultado dado por (2.2) se tiene que

$$\begin{aligned} P \left\{ \hat{A}_n \stackrel{\circ}{=} E \hat{A}_n \stackrel{\circ}{=} > \epsilon \right\} &= P \left\{ \sup_{T \cup J} j^{\otimes n}(S) j \frac{\circ}{h} > \epsilon \right\} = \\ &= P \left\{ \sup_{T \cup J} j^{\otimes n}(S) j > \epsilon \frac{\circ}{h} \right\} = \\ &= \frac{c_1}{2\Gamma h} \exp \left\{ -i \frac{c_2 n h^{2a-2} \frac{3}{4}}{2^{1/2} \epsilon} \right\} = \\ &= \frac{c_1}{2\Gamma h} \exp \left\{ -i \frac{c_2 n h^{2a-2} \frac{3}{4}}{2^{1/2} \epsilon} \right\} = \frac{c_1^0}{h} \exp \left\{ -i c_2^0 n h^{2a-2} \right\} \end{aligned}$$

siempre que $2 \cdot \frac{c_1^0}{h} \exp \left\{ -i c_2^0 n h^{2a-2} \right\} < \epsilon$ y $\epsilon > e_0 \frac{c_1^0}{2\Gamma h}$ para alguna constante positiva e_0 :

Tomando $\epsilon = c \frac{\ln n}{nh} \mathbb{1}_{1=2}$, $c > 0$, las condiciones anteriores significan:

$$2 \cdot \frac{c_1^0}{h} \exp \left\{ -i c_2^0 n h^{2a-2} \right\} < c \frac{\ln n}{nh} \mathbb{1}_{1=2}$$

y

$$\frac{c_1^0}{h} \exp \left\{ -i c_2^0 n h^{2a-2} \right\} < e_0 \frac{c_1^0}{2\Gamma h} \mathbb{1}_{1=2} < c \frac{\ln n}{nh} \mathbb{1}_{1=2} \cdot e_0 2\Gamma$$

que se verificarán si $h \rightarrow 0$; $nh \rightarrow 1$ y $\ln n = (nh) \rightarrow 0$; y bajo estas hipótesis obtenemos:

$$\begin{aligned} \frac{c_1^0}{h} \exp \left\{ -i c_2^0 n h^{2a-2} \right\} &= \frac{c_1^0}{h} \exp \left\{ -i \frac{c_2^0 n h c^2 \ln n \frac{3}{4}}{nh} \right\} = \\ &= \frac{c_1^0}{h} \exp \left\{ -i c_2^0 c^2 \ln n \right\} = c_1^0 \frac{1}{h n c_2^0 c^2} \end{aligned} \quad (2.5)$$

Eligiendo c de forma apropiada conseguimos que (2.5) sea sumable, y aplicando el lema de Borel-Cantelli se llega a

$$\hat{A}_n \stackrel{\circ}{=} E \hat{A}_n \stackrel{\circ}{=} O \left(\frac{\ln n}{nh} \mathbb{1}_{1=2} \right) \quad c.s. \quad (2.6)$$

Estudiamos a continuación el orden de $\mathbb{E} \hat{A}_n(t; x) - A(t; x)$. Conviene trabajar con la expresión de $\mathbb{E} \hat{A}_n(t; x)$ del modo siguiente:

$$\begin{aligned} \mathbb{E} \hat{A}_n(t; x) &= \frac{1}{h} \mathbb{E} \int_{t-h}^{t+h} K\left(\frac{x_j - x}{h}\right) \frac{\mu_{x_j}}{h} dx_j = \\ &= \frac{1}{h} \mathbb{E} \int_{t-h}^{t+h} K\left(\frac{x_j - x}{h}\right) \frac{\mu_{x_j}}{h} dx_j = \\ &= \frac{1}{h} \mathbb{E} \int_{t-h}^{t+h} S(t; x) K\left(\frac{x_j - x}{h}\right) \frac{\mu_{x_j}}{h} dx_j = \\ &= \frac{1}{h} \int_{t-h}^{t+h} S(t; u) K\left(\frac{x_j - u}{h}\right) m(u) du \end{aligned} \tag{2.7}$$

Haciendo un cambio de variable $\frac{x_j - u}{h} = z$ y utilizando un desarrollo de Taylor de orden 2 para $S(t; x)$ y m en el punto x ($x \in I$), obtenemos:

$$\begin{aligned} (2.7) &= \int_{t-h}^{t+h} K(z) S(t; x + hz) m(x + hz) dz = \\ &= \int_{t-h}^{t+h} K(z) \left[S(t; x) + S'(t; x)hz + \frac{S''(t; x)}{2}h^2z^2 + O(h^3) \right] \frac{\mu_{x+hz}}{h} dz \\ &= \int_{t-h}^{t+h} K(z) \left[m(x) + m'(x)hz + \frac{m''(x)}{2}h^2z^2 + O(h^3) \right] dz \\ &= \int_{t-h}^{t+h} S(t; x)m(x) K(z) dz + h \int_{t-h}^{t+h} S'(t; x)m'(x) + S''(t; x)m(x) K(z) dz \\ &= \int_{t-h}^{t+h} zK(z) dz + \\ &+ h^2 \int_{t-h}^{t+h} \frac{1}{2} S''(t; x)m''(x) + \frac{1}{2} S''(t; x)m'(x) + S''(t; x)m(x) K(z) dz \\ &= \int_{t-h}^{t+h} z^2 K(z) dz + O(h^3) \end{aligned}$$

Utilizando las hipótesis sobre la función núcleo y sobre la acotación de S'' , m'' , S'' y m'' respecto a x se tiene de modo uniforme en $[0; +1) \in I$

$$\mathbb{E} \hat{A}_n(t; x) = S(t; x)m(x) + O(h^2)$$

Por tanto

$$\mathbb{E} \hat{A}_n(t; x) - A(t; x) = O(h^2)$$

que junto con (2.6) prueba el resultado deseado de

$$\hat{A}_n(t; x) - A(t; x) = O\left(\frac{\mu_{\ln n}^{11=2}}{nh}\right) + O(h^2) \quad \text{c.s.}$$

Si además $nh^5 = \ln n = O(1)$ se tiene

$$h^2 = \frac{\mu_{\ln n}^{11=2}}{nh} = \frac{\mu_{nh^5}^{11=2}}{\ln n} = O(1)$$

es decir

$$\hat{A}_n(t; x) - A(t; x) = O\left(\frac{\mu_{\ln n}^{11=2}}{nh}\right) \quad \text{c.s.}$$

En realidad estamos interesados en el orden de $\hat{S}_n(t; x) - S(t; x)$: Teniendo en cuenta la siguiente descomposición:

$$\begin{aligned} \hat{S}_n(t; x) - S(t; x) &= \frac{\hat{A}_n(t; x)}{\hat{m}_n(x)} - \frac{A(t; x)}{m(x)} = \\ &= \frac{\hat{A}_n(t; x) - A(t; x)}{\hat{m}_n(x)} + \frac{(m(x) - \hat{m}_n(x))A(t; x)}{\hat{m}_n(x)m(x)} \end{aligned}$$

y que $m(x) - \hat{m}_n(x)$ es un caso particular de $\hat{A}_n(t; x) - A(t; x)$ para $t = 0$, el orden de $\hat{S}_n(t; x) - S(t; x)$ coincide con el de $\hat{A}_n(t; x) - A(t; x)$:

2.2 Estimación no paramétrica de la función de distribución condicional con censura.

La estimación no paramétrica de la función de distribución condicional para datos censurados ha sido estudiada por varios autores. La definición de un estimador límite producto condicional, también llamado generalizado, para dicha función se puede ver, por primera vez, en Beran (1981). Dependiendo de que el diseño sea ...jo o aleatorio y del tipo de pesos de la estimación no paramétrica se estudian distintas variantes y propiedades del estimador, entre las que cabe citar: la consistencia uniforme débil y fuerte probada por Dabrowska (1989) para un contexto de diseño aleatorio con pesos del tipo Nadaraya-Watson; una representación casi segura del estimador condicional, también en diseño aleatorio pero con pesos del tipo k puntos próximos derivada por Akritas (1994); las representaciones casi seguras estudiadas

por González Manteiga y Cadarso, Suárez (1994) y Van Keilegom y Veraverbeke (1997a) cuyo orden de error es $O\left(\frac{\ln n}{nh}\right)^{3-4}$; para un contexto de diseño ...jo con pesos del tipo Gasser-Müller y, como consecuencia de ellas, la normalidad asintótica del estimador y la convergencia débil del proceso.

Antes de exponer con detalle estas propiedades, presentaremos el modelo de censura aleatoria por la derecha con covariables, y derivaremos una expresión general del estimador límite producto condicional.

La modelización de una situación de censura con covariables tiene en cuenta, además de las variables: Y tiempo de fallo, C variable censura, $Z = \min fY; Cg$ variable observada y $\pm = 1_{fY \cdot Cg}$ variable indicadora de fallo; otra variable X denominada covariable, cuya relación con Y se pretende analizar. Generalmente esta variable se observa al comienzo del estudio de seguimiento y por ello los modelos más habituales suponen que la censura no afecta a las covariables. Es por esto, que observaremos una muestra del tipo

$$\begin{matrix} f(X_1; Z_1; \pm_1); (X_2; Z_2; \pm_2); \dots; (X_n; Z_n; \pm_n)g \\ \text{donde } Z_i = \min fY_i; C_i g \text{ y } \pm_i = 1_{fY_i \cdot C_i g} \end{matrix}$$

En este modelo de censura con covariables se pretende de...nir un estimador de la función de distribución de Y condicionada a X , por ello debemos considerar las funciones de...nidas en el capítulo anterior para el modelo de censura, pero de modo condicionado a X . Así tenemos:

1) La función de distribución de la variable tiempo de fallo condicionada a $X = x$, $F(y | x) = P(Y \leq y | X = x)$ y su función de supervivencia o ...abilidad condicional, $S(y | x) = 1 - F(y | x) = P(Y > y | X = x)$:

2) La función de distribución de la variable de censura condicionada a $X = x$, $G(y | x) = P(C \leq y | X = x)$ y su función de supervivencia o ...abilidad condicional, $1 - G(y | x) = P(C > y | X = x)$:

3) La función de distribución de la variable observada Z condicionada a $X = x$, $H(y | x) = P(Z \leq y | X = x)$ y su función de supervivencia o ...abilidad condicional, $1 - H(y | x)$.

Bajo la hipótesis, habitual en este modelo, de

$$Y \text{ y } C \text{ condicionalmente independientes a } X \tag{2.8}$$

se tiene (de forma análoga a (1.1)) que

$$1 - H(y | x) = (1 - F(y | x))(1 - G(y | x)) \tag{2.9}$$

4) La subdistribución de la variable observada, Z , cuando es una observación no censurada condicionada a $X = x$

$$H_1(y | x) = P(Z \leq y; \pm = 1 | X = x):$$

Esta función bajo la hipótesis (2.8) se puede escribir (de forma análoga a (1.2)) de la forma

$$H_1(y | x) = \int_0^y \frac{dH_1(t | x)}{1 - G(t | x)} dF(t | x) : \quad (2.10)$$

5) La función razón de fallo acumulada condicionada a $X = x$

$$\Lambda(y | x) = \int_0^y \frac{dF(t | x)}{1 - F(t | x)}$$

que determina de forma única la función de distribución $F(y | x)$ mediante la relación:

$$1 - F(y | x) = e^{-\Lambda_c(y | x)} \prod_{a_i \leq y} (1 - \Delta \Lambda(a_i | x)) \quad (2.11)$$

$\Delta \Lambda(a_i | x) = \Lambda(a_i | x) - \Lambda(a_i^- | x)$

donde $\Lambda_c(y | x)$ denota la parte continua de $\Lambda(y | x)$, A el conjunto de puntos donde $\Lambda(y | x)$ tiene discontinuidades de salto y $\Delta \Lambda(a_i | x) = \Lambda(a_i | x) - \Lambda(a_i^- | x)$ la magnitud del salto de $\Lambda(y | x)$ en a_i :

Precisamente, la forma más habitual de definir el estimador límite producto de la función de distribución condicional es derivarlo a partir de un estimador de $\Lambda(y | x)$ y de la relación (2.11), mediante el procedimiento ya familiar de:

1. Expresar la función de azar acumulada condicional $\Lambda(y | x)$; en función de cantidades estimables: para cada x fijo, si $y < b_{H(\cdot | x)}$; se verifica que

$$\Lambda(y | x) = \int_0^y \frac{dH_1(t | x)}{1 - H(t | x)} : \quad (2.12)$$

Esta expresión se obtiene de forma totalmente análoga a la ecuación (1.7) del capítulo anterior, pero utilizando las relaciones del caso condicional dadas por (2.9) y (2.10) en lugar de (1.1) y (1.2), respectivamente.

2. Estimar $H(y | x)$; y $H_1(y | x)$ mediante estimadores no paramétricos de funciones de distribución condicionales, esto es:

$$\hat{H}_n(y | x) = \sum_{i=1}^n B_{ni}(x) 1_{F_{Z_i} \leq y}$$

$$\hat{H}_{1n}(y | x) = \sum_{i=1}^n B_{ni}(x) 1_{F_{Z_i} \leq y; \pm_i = 1g}$$

donde $B_{ni}(x)$ denotan los pesos de la estimación no paramétrica.

3. Sustituir $H(t | j, x)$ y $H_1(t | j, x)$ por $\hat{H}_n(t | j, x)$ y $\hat{H}_{1n}(t | j, x)$ en la expresión (2.12) para obtener un estimador de $\Lambda(y | j, x)$; cuya expresión será

$$\hat{\Lambda}_n^c(y | j, x) = \int_0^y \frac{d\hat{H}_{1n}(t | j, x)}{\hat{H}_n(t | j, x)} = \prod_{i=1}^n \frac{1_{f_{Z_i} \cdot y; \pm_i = 1} g_{B_{ni}}(x)}{\prod_{j=1}^n 1_{f_{Z_j} < Z_i} g_{B_{nj}}(x)}$$

4. Estimar $F(y | j, x)$ mediante

$$\begin{aligned} 1_{i=1} \hat{F}_n^c(y | j, x) &= \prod_{i=1}^n \frac{1_{f_{Z_i} \cdot y; \pm_i = 1} g_{B_{ni}}(x)}{1_{i=1} \hat{H}_n(Z_i | j, x)} = \\ &= \prod_{i=1}^n \frac{1_{f_{Z_i} \cdot y; \pm_i = 1} g_{B_{ni}}(x)}{\prod_{j=1}^n 1_{f_{Z_j} < Z_i} g_{B_{nj}}(x)} \end{aligned} \quad (2.13)$$

debido a la relación entre $F(y | j, x)$ y $\Lambda(y | j, x)$ expresada por la ecuación (2.11) y al hecho de que el estimador $\hat{\Lambda}_n^c(y | j, x)$ es una función en escalera con saltos en los puntos $f_{Z_i} = Y_i = \pm_i = 1_{g_{i=1}^n}$:

Aunque para nosotros $f_{B_{ni}}(x)g_{i=1}^n$ denota una sucesión de pesos núcleo del tipo Nadaraya-Watson, cualquier otro tipo de pesos no paramétricos pueden ser considerados, por ello la expresión (2.13) proporciona una formulación general del estimador de la función de distribución de Y condicionada a $X = x$ bajo censura, llamado generalmente estimador límite- producto condicional o generalizado (LPC).

A continuación detallamos las propiedades más interesantes del estimador LPC con censura, que como ya hemos dicho al principio de esta sección han sido derivadas en diversos contextos.

Propiedad 2.43 El estimador LPC definido en (2.13) coincide:

- a) con el estimador tipo núcleo de la función de distribución condicional (para datos completos) en el caso de ausencia de censura.
- b) con el estimador de Kaplan-Meier en el caso de ausencia de covariables.

A partir de la expresión (2.13) del estimador y utilizando que cuando no hay censura todos los \pm toman el valor 1 y las Z coinciden con los valores de Y se llega

a

$$\begin{aligned}
1_i \hat{F}_n^c(y_j | x) &= \prod_{i=1}^n \frac{\prod_{j=1}^n 1_{f_{Y_i} < y_j} B_{ni}(x)}{\prod_{j=1}^n 1_{f_{Y_j} < Y_i} B_{nj}(x)} = \\
&= \prod_{i=Y_{i:n} \cdot y} \frac{\prod_{j=1}^n 1_{f_{Y_j} < Y_i} B_{nj}(x)}{\prod_{j=1}^n 1_{f_{Y_j} < Y_i} B_{nj}(x)} = \\
&= \prod_{i=Y_{i:n} \cdot y} \frac{\prod_{j=1}^i B_{n[j]}(x)}{\prod_{j=1}^{i-1} B_{n[j]}(x)} = \\
&= \frac{1_i B_{n[1]}(x)}{1} \cdot \frac{1_i B_{n[1]}(x) \cdot B_{n[2]}(x)}{1_i B_{n[1]}(x)} \cdot \\
&\quad \cdot \dots \cdot \frac{1_i B_{n[1]}(x) \cdot \dots \cdot B_{n[k_i-1]}(x) \cdot B_{n[k_i]}(x)}{1_i B_{n[1]}(x) \cdot \dots \cdot B_{n[k_i-1]}(x)} \\
&= 1_i B_{n[1]}(x) \cdot \dots \cdot B_{n[k_i-1]}(x) \cdot B_{n[k_i]}(x) = \\
&= \prod_{i=1}^n 1_{f_{Y_i} < y_j} B_{ni}(x) = 1_i \hat{F}_n^c(y_j | x)
\end{aligned}$$

donde $B_{n[j]}(x)$ denota el peso correspondiente a $X_{[j:n]}$ que es el concomitante asociado al j -ésimo estadístico ordenado $Y_{j:n}$ y siendo trivialmente $k = \max \{i = Y_{i:n} \cdot y\}$. Esto prueba el apartado a).

El apartado b) se demuestra fácilmente teniendo en cuenta que en ausencia de covariables $B_{ni}(x) = \frac{1}{n}$ para todo i . De este modo la expresión (2.13) queda

$$\begin{aligned}
1_i \hat{F}_n^c(y_j | x) &= \prod_{i=1}^n \frac{\prod_{j=1}^n 1_{f_{Z_i} < y_j} \frac{1}{n}}{\prod_{j=1}^n 1_{f_{Z_j} < Z_i} \frac{1}{n}} = \\
&= \prod_{i=1}^n \frac{1_{f_{Z_i} < y_j} \frac{1}{n}}{n_i \text{ rango}(Z_i) + 1}
\end{aligned}$$

coincidiendo con (1.8), expresión del estimador de Kaplan-Meier.

Propiedad 2.44 El estimador LPC tiene forma escalonada con saltos en las observaciones no censuradas. La magnitud del salto en Z_i ; $d\hat{F}_n^c(Z_i | x)$; es igual a

$$d\hat{F}_n^c(Z_i | x) = \frac{\pm_i B_{ni}(x)}{1_i \hat{H}_n(Z_i | x)} \cdot 1_i \hat{F}_n^c(Z_i | x)$$

Esta expresión se obtiene al calcular $\hat{F}_n^c(Z_i | j, x)$ y $\hat{F}_n^c(Z_i^{(i)} | j, x)$ utilizando (2.13):

$$\begin{aligned} & \hat{F}_n^c(Z_i | j, x) - \hat{F}_n^c(Z_i^{(i)} | j, x) = \\ &= \prod_{i=1}^n \frac{\hat{F}_n^c(Z_i | j, x)}{\hat{F}_n^c(Z_i^{(i)} | j, x)} - \prod_{i=1}^n \frac{\hat{F}_n^c(Z_i | j, x)}{\hat{F}_n^c(Z_i^{(i)} | j, x)} = \\ &= \prod_{i=1}^n \frac{1_{f_{Z_j < Z_i g \pm j}} B_{nj}(x)}{1_{f_{Z_j < Z_i g \pm j}} \hat{H}_n(Z_j | j, x)} - \prod_{i=1}^n \frac{1_{f_{Z_j < Z_i g \pm j}} B_{nj}(x)}{1_{f_{Z_j < Z_i g \pm j}} \hat{H}_n(Z_j | j, x)} = \\ &= \prod_{i=1}^n \frac{1_{f_{Z_j < Z_i g \pm j}} B_{nj}(x)}{\hat{C}_n(Z_j | j, x)} - \prod_{i=1}^n \frac{1_{f_{Z_j < Z_i g \pm j}} B_{nj}(x)}{1_{f_{Z_j < Z_i g \pm j}} \hat{H}_n(Z_j | j, x)} = \\ &= \prod_{i=1}^n \hat{F}_n^{tc}(Z_i^{(i)} | j, x) \frac{1_{f_{Z_j < Z_i g \pm j}} B_{nj}(x)}{1_{f_{Z_j < Z_i g \pm j}} \hat{H}_n(Z_j | j, x)} \end{aligned}$$

Propiedad 2.45 El estimador LPC para la función de distribución de la variable censura C condicionada a X, $G(y | j, x)$; viene dado por

$$\prod_{i=1}^n \hat{G}_n^c(y | j, x) = \prod_{i=1}^n \frac{1_{f_{Z_i \cdot y; \pm_i = 0g}} B_{ni}(x)}{1_{f_{Z_i \cdot y; \pm_i = 0g}} \hat{H}_n(Z_i^{(i)} | j, x)}$$

Es inmediato sin más que tener en cuenta que una observación censurada de Y_i corresponde a una observación no censurada de C_i y viceversa.

Propiedad 2.46 El producto de los estimadores LPC para las funciones de supervivencia de las variables Y y C coincide con el estimador no paramétrico de la supervivencia de la variable Z, es decir

$$\prod_{i=1}^n \hat{F}_n^c(y) - \prod_{i=1}^n \hat{G}_n^c(y) = \prod_{i=1}^n \hat{H}_n(y)$$

Explicitando las fórmulas del primer miembro de la igualdad anterior se tiene

$$\begin{aligned} & \prod_{i=1}^n \frac{1_{f_{Z_i \cdot y; \pm_i = 1g}} B_{ni}(x)}{1_{f_{Z_i \cdot y; \pm_i = 1g}} \hat{H}_n(Z_i^{(i)} | j, x)} - \prod_{i=1}^n \frac{1_{f_{Z_i \cdot y; \pm_i = 0g}} B_{ni}(x)}{1_{f_{Z_i \cdot y; \pm_i = 0g}} \hat{H}_n(Z_i^{(i)} | j, x)} \\ &= \prod_{i=1}^n \frac{1_{f_{Z_i \cdot yg}} B_{ni}(x)}{1_{f_{Z_i \cdot yg}} \hat{H}_n(Z_i^{(i)} | j, x)} = \prod_{i=1}^n B_{ni}(x) 1_{f_{Z_i \cdot yg}} \end{aligned}$$

El último paso se obtiene de igual forma que la propiedad 2.43a).

Propiedad 2.47 Supóngase un contexto de diseño aleatorio donde X es unidimensional y absolutamente continua con función de densidad $m(x)$. Sea $\hat{F}_n^c(y|x)$ el estimador LPC con pesos del tipo Nadaraya-Watson.

Si se verifican las condiciones:

i) Sea $I = [x_1; x_2]$ un intervalo contenido en el soporte de m verificando que

$$0 < \delta = \inf [m(x) : x \in I_\pm] < \sup [m(x) : x \in I_\pm] = \Gamma < 1$$

para algún entorno $I_\pm = [x_1 \pm \delta; x_2 \pm \delta]$ con $\delta > 0$ y $0 < \delta\Gamma < 1$:

Además, existe T tal que $\inf [(1 - H(T|x)) : x \in I_\pm] \geq \mu > 0$:

ii) Las funciones $m(x)$; $H(y|x)$ y $H_1(y|x)$ tienen primera derivada respecto a x continua y acotada en $(y|x) \in [0; +1] \cap I_\pm$:

iii) Las funciones $m(x)$; $H(y|x)$ y $H_1(y|x)$ tienen segunda derivada respecto a x continua y acotada en $(y|x) \in [0; +1] \cap I_\pm$:

iv) La función núcleo, K , es una función de densidad simétrica, que vale 0 fuera del intervalo $(-1; 1)$ y su variación total está acotada por $\nu < +1$.

v) El parámetro ventana $h = h_n$ verifica: $h \rightarrow 0$ y $\ln n = o(nh)$.

Entonces:

$$\sup_{[0; T] \cap I} | \hat{F}_n^c(y|x) - F(y|x) | = o_p(0)$$

Si además imponemos $nh^5 = \ln n \rightarrow 0$ se tiene:

$$\sup_{[0; T] \cap I} | \hat{F}_n^c(y|x) - F(y|x) | = O((\ln n / (nh))^{1/2}) \quad \text{c.s.}$$

Este orden de convergencia uniforme fuerte se debe a Dabrowska (1989).

Propiedad 2.48 Supóngase un contexto de diseño aleatorio donde $(X; Y)$ es una v.a. absolutamente continua y M denota la distribución marginal de X . Sea $\hat{F}_n^c(y|x)$ el estimador LPC con pesos del tipo k -vecinos próximos con función núcleo $K(u) = 0.5 \cdot \mathbb{1}_{|u| < 1}$.

Si se verifican las condiciones:

i) Las densidades conjuntas $f(u; z)$ y $f(u; \pm; z)$ de $(M(X); Z)$ y de $(M(X); \pm; Z)$ respectivamente, son dos veces continuamente diferenciables con respecto a u , y

$$\begin{aligned} \sup_{-a}^a | f''(u; z) | &< 1 \\ \sup_{-a}^a | f''(u; \pm; z) | &< 1 : \end{aligned}$$

ii) La sucesión $h = (h_n)_{n \geq 2N}$ verifica que $nh^3 \int_0^1 \ln h \, d\mu_{3:5} \neq 0$ y $nh^5 \int_0^1 \log h \, d\mu_{1:1} \neq 0$.
Entonces para $y \cdot T_{X_i} < b_{H(t_j X_i)}$ se verifica que:

$$\hat{F}_n^c(y | X_i) = F(y | X_i) + \frac{1}{nh} \sum_{j=1}^n \omega_j(y; X_i) \int_{\hat{M}_j(X_i)}^{\hat{M}_j} f_j \, d\mu_j + r_n(y; X_i)$$

donde

$$\omega_j(y; X_i) = (1 - F(y | X_i)) \int$$

$$\int \frac{1_{FZ_j \cdot y; \pm_j = 1g}}{1 - H(Z_j | X_i)} \int_0^Z \frac{1_{ft \cdot Z_j g}}{(1 - H(t | X_i))^2} dH_1(t | X_i) ;$$

$$\sup_j |r_n(y; X_i)| = o_p \left(\frac{1}{n^{1/2}} \right) \quad c.s.$$

y \hat{M} es la distribución empírica de los X_i :

Este resultado, debido a Akritas (1994), constituye una representación casi segura del estimador LPC que permite escribir la diferencia entre dicho estimador y la función de distribución condicional como una suma de variables independientes ponderada por los pesos no paramétricos, más un término de error cuyo orden es conocido. El orden obtenido por Akritas es un poco pobre, pero suficiente para sus demostraciones, ya que el presente resultado forma parte, en su artículo, de otro más general: la definición, representación casi segura del estimador y convergencia débil del proceso asociados a la estimación de la función de distribución conjunta del vector $(X; Y)$:

Propiedad 2.49 Supóngase un contexto de diseño ...jo sobre el intervalo $[0; 1] : 0 \cdot x_1 \cdot \dots \cdot x_n \cdot 1$; con observaciones $f(x_i; Z_i; \pm_i)_{i=1}^n$. Sea $\hat{F}_n^c(y | x)$ el estimador LPC con pesos $B_{ni}(x)$ del tipo Gasser-Müller.

Si se verifican las condiciones:

i) $\max_i |x_i - x_{i-1}| = O(1/n)$:

ii) Las funciones $F(y | x_i)$ y $G(y | x_i)$; $i = 1; \dots; n$; pertenecen a las familias $fF(y | x) = x^2 [0; 1]g$ y $fG(y | x) = x^2 [0; 1]g$ respectivamente, que son continuamente derivables respecto a y para cada x y dos veces derivables con derivada segunda acotada (uniformemente en y) con respecto a cada x_i , $i = 1; 2; \dots; n$.

iii) La función núcleo, K , es una función de densidad, simétrica, positiva, de soporte compacto y Lipschitz-continua de orden 1.

iv) La sucesión $h = (h_n)_{n \geq 2N}$ verifica que: $h \neq 0$, $\ln n = (nh) \neq 0$ y $nh^5 = \ln n = O(1)$:

Entonces para $y \cdot T < b_{H(t|x)}$ se verifica que:

$$\hat{F}_n^c(y|x) = F(y|x) + \sum_{i=1}^n B_{ni}(x) \kappa_i(y|x) + r_n(y|x)$$

donde

$$\kappa_i(y|x) = (1 - F(y|x)) \frac{1_{f_{Z_i} \cdot y; \pm i = 1g}}{1 - H(Z_i|x)} \int_0^y \frac{1_{f_{t, Z_j} g}}{(1 - H(t|x))^2} dH_1(t|x);$$

y

$$\sup_{0 \leq y \leq T} |r_n(y|x)| = O\left(\frac{\tilde{A}_\mu \ln n}{nh} \mathbb{P}_{3=4}^!\right) \quad c:s:$$

Esta representación casi segura cuyo orden de error es $O\left(\frac{\ln n}{nh}\right)^{3=4}$ se puede ver en González Manteiga y Cadarso Suárez (1994) y Van Keilegom y Veraverbeke (1997a). Como consecuencias de dicha representación se pueden obtener: la consistencia fuerte uniforme y la normalidad asintótica del estimador así como la convergencia débil del proceso, propiedades que se enuncian a continuación.

Propiedad 2.50 En las mismas condiciones y con la misma notación de la propiedad anterior se tiene:

$$\sup_{0 \leq y \leq T} |\hat{F}_n^c(y|x) - F(y|x)| = O\left(\frac{\tilde{A}_\mu \ln n}{nh} \mathbb{P}_{1=2}^!\right) \quad c:s:$$

Este resultado de consistencia uniforme fuerte fue demostrado por Van Keilegom y Veraverbeke (1996).

Propiedad 2.51 En las mismas condiciones y con la misma notación de la Propiedad 2.49 se tiene:

a) Si $(\ln n)^3 = nh \neq 0$ y $nh^5 \neq 0$; entonces para $y \cdot T < b_{H(t|x)}$ se verifica que

$$(nh)^{1=2} \hat{F}_n^c(y|x) - F(y|x) \stackrel{i}{\underset{i}{\rightsquigarrow}} N(0; s(y|x))$$

b) Si $h = Cn^{1=5}$; entonces para $y \cdot T < b_{H(t|x)}$ se verifica que

$$(nh)^{1=2} \hat{F}_n^c(y|x) - F(y|x) \stackrel{i}{\underset{i}{\rightsquigarrow}} N(b(y|x); s(y|x))$$

donde

$$b(y|x) = C^{5-2} \frac{1}{2} (1 - F(y|x)) \int_{-\infty}^{\infty} z^2 K(z) dz \int_{-\infty}^{\infty} \int_0^y \frac{d\ddot{H}_1(t|x)}{1 - H(t|x)} \int_0^y \frac{1 - \ddot{H}(t|x)}{(1 - H(t|x))^2} dH_1(t|x) ;$$

$$s^2(y|x) = (1 - F(y|x))^2 \int_{-\infty}^{\infty} K^2(z) dz \int_0^y \frac{dH_1(t|x)}{(1 - H(t|x))^2} ;$$

y $\ddot{H}(t|x)$; $\ddot{H}_1(t|x)$ denotan las derivadas segundas respecto a la variable x de H y H_1 ; respectivamente.

Normalidad asintótica obtenida por González Manteiga y Cadarso Suárez (1994) y Van Keilegom y Veraverbeke (1997a).

Propiedad 2.52 En las mismas condiciones y con la misma notación del apartado a) de la propiedad anterior se tiene:

$$(nh)^{1-2} [\hat{F}_n(t|x) - F(t|x)] \overset{d}{\rightarrow} N(t|x) \quad \text{en } D[0; T]$$

donde $N(t|x)$ denota un proceso Gaussiano de media cero y función de covarianzas

$$\Gamma(y; t|x) = (1 - F(y|x))(1 - F(t|x)) \int_{-\infty}^{\infty} K^2(z) dz \int_0^{y \wedge t} \frac{dH_1(u|x)}{(1 - H(u|x))^2}$$

$y \wedge t = \min\{y, t\}$; tg y $D[0; T] = \{f \in C([0; T]; R) : f \text{ continua por la derecha y con discontinuidades, a lo sumo, de saltos, con la topología de Skorohod (Billingsley, pag.111)}\}$.

Se trata de la convergencia débil del proceso $(nh)^{1-2} [\hat{F}_n(t|x) - F(t|x)]$ y se puede ver en Van Keilegom y Veraverbeke (1997b).

2.3 Estimación no paramétrica de la función de distribución condicional con censura y truncamiento.

En primer lugar decir que la única referencia que hemos encontrado de un estimador no paramétrico de la función de distribución condicional con truncamiento es de LaValley y Akritas (1994). Su estimador tiene la siguiente expresión:

$$\hat{F}_n^t(y|x) = \frac{\sum_{i=1}^n \tilde{A}_i \mathbb{1}_{F_{Y_i} \cdot y} B_{ni}(x)}{\sum_{j=1}^n \mathbb{1}_{F_{T_j} \cdot Y_j} B_{nj}(x)} \quad (2.14)$$

donde $f_{B_{ni}}(x)g_{i=1}^n$ denota una sucesión de pesos no paramétricos del tipo k - puntos próximos. Para este estimador, los autores derivan una representación casi segura con un término de error de orden $O((\ln n)/(nh))^{3/4}$:

En este apartado definiremos un estimador de la función de distribución condicional para datos con censura y truncamiento cuando hay covariables y obtendremos una representación casi segura para dicho estimador como suma ponderada de variables independientes, que permitirá demostrar la consistencia uniforme fuerte, la normalidad asintótica y la convergencia débil del proceso. Como ya hemos mencionado anteriormente, trabajaremos con pesos no paramétricos de Nadaraya-Watson. Estos resultados han sido publicados en el Journal of Nonparametric Statistics (Iglesias Pérez y González Manteiga (1999)).

Comenzamos precisando el modelo de censura y truncamiento con covariables, que aporta al modelo de censura y truncamiento expuesto con detalle en el capítulo anterior, una nueva variable X , que denota la covariable relacionada con Y .

Sea entonces

| |
|--|
| $(X_i; Z_i; T_i; \delta_i) \quad i = 1; 2; \dots; n$ |
| donde $T_i \leq Z_i$ |

la muestra observada, y consideremos las siguientes funciones:

1) $g(x) = P(T \leq Z | X = x)$ que supondremos estrictamente positiva.

2) $M(x) = P(X \leq x)$ la función de distribución de X y también, $M^*(x) = P(X \leq x | T \leq Z)$ la función de distribución de X condicionada a que no ha ocurrido truncamiento. Cuando se suponga X absolutamente continua, denotaremos por m y m^* a las densidades asociadas a M y M^* ; respectivamente. Es interesante notar que

$$m^*(x) = \frac{g(x)}{g} m(x)$$

donde $g = P(T \leq Z)$; y por tanto $g(x)/g$ representa un índice de ausencia de truncamiento en el punto x :

3) $1 - F(y | x) = P(Y > y | X = x)$ la función de supervivencia de Y condicionada a $X = x$.

4) $1 - G(y | x) = P(C > y | X = x)$ la función de supervivencia de C condicionada a $X = x$.

5) $1 - L(y | x) = P(T > y | X = x)$ la función de supervivencia de T condicionada a $X = x$, y sea también $1 - L^*(y | x) = P(T > y | X = x; T \leq Z)$ la función de supervivencia de T condicionada a $X = x$ y a que no ha ocurrido truncamiento.

6) $1 - H(y | x) = P(Z > y | X = x)$ la función de supervivencia de Z condicionada a $X = x$. Nosotros supondremos que

$$\boxed{Y; T \text{ y } C \text{ son condicionalmente independientes a } X} \quad (2.15)$$

y por tanto se verificará que $1 - H(y | x) = (1 - F(y | x))(1 - G(y | x))$.

Como en el caso anterior definiremos $H^a(y | x) = P(Z > y | X = x; T > Z)$:

7) $H_1(y | x) = P(Z > y; \pm = 1 | X = x)$ la subdistribución de Z cuando no hay censura condicionada a $X = x$ y $H_1^a(y | x) = P(Z > y; \pm = 1 | X = x; T > Z)$ la subdistribución de Z cuando no hay censura condicionada a $X = x$ y a que no ha ocurrido truncamiento.

8) $H_0(y | x) = P(Z > y; \pm = 0 | X = x)$ la subdistribución de Z cuando hay censura condicionada a $X = x$ y también $H_0^a(y | x) = P(Z > y; \pm = 0 | X = x; T > Z)$ la subdistribución de Z cuando hay censura condicionada a $X = x$ y a que no ha ocurrido truncamiento.

9) La función razón de fallo acumulada de Y condicionada a $X = x$

$$\Lambda(y | x) = \int_0^y \frac{dF(t | x)}{1 - F(t | x)}$$

que determina de modo único $F(\cdot | x)$ mediante la relación

$$1 - F(y | x) = e^{-\int_0^y \Lambda_c(t | x) dt} \prod_{a_i \leq y} (1 - \Delta a_i | x) \quad (2.16)$$

$a_i \leq y$

donde $\Lambda_c(y | x)$ denota la parte continua de $\Lambda(y | x)$, A el conjunto de puntos donde $\Lambda(y | x)$ tiene discontinuidades de salto y $\Delta a_i | x = \Lambda(a_i | x) - \Lambda(a_i^- | x)$ la magnitud del salto de $\Lambda(y | x)$ en a_i :

2.3.1 Definición del estimador.

Definiremos el estimador de $F(\cdot | x)$ partiendo de un estimador para $\Lambda(\cdot | x)$ y utilizando la relación entre ambas funciones, especificada en (2.16). Para calcular un estimador de $\Lambda(\cdot | x)$ tendremos en cuenta que

$$\begin{aligned} H_1^a(y | x) &= P(Z > y; \pm = 1 | x; T > Z) = \\ &= \frac{P(Z > y; \pm = 1; T > Z | x)}{P(T > Z | x)} = \frac{P(Y > y; Y > C; T > Y | x)}{P(T > Z | x)} = \\ &= \int_0^y \frac{dF(t | x)}{1 - F(t | x)} \end{aligned}$$

de donde obtenemos

$$dH_1^a(y|x) = \int_0^y (x)^{i-1} P(T \cdot y \cdot C|x) dF(y|x); \quad (2.17)$$

Y utilizando esta expresión alcanzamos el objetivo de escribir $\Lambda(t|x)$ en función de expresiones estimables empíricamente:

$$\Lambda(y|x) = \int_0^y \frac{dF(t|x)}{1 - F(t|x)} = \quad (2.18)$$

$$= \int_0^y \frac{\int_0^t (x)^{i-1} P(T \cdot t \cdot C|x) dF(t|x)}{\int_0^t (x)^{i-1} P(T \cdot t \cdot C|x) (1 - F(t|x))} =$$

$$= \int_0^y \frac{dH_1^a(t|x)}{C(t|x)} \quad (2.19)$$

donde

$$C(t|x) = \int_0^t (x)^{i-1} P(T \cdot t \cdot C|x) (1 - F(t|x)) \quad (2.20)$$

$$= P(T \cdot t \cdot Z|x; T \cdot Z):$$

Antes de definir un estimador de $\Lambda(t|x)$ a partir de la expresión (2.19), es conveniente profundizar en la cadena de igualdades que conducen a ella.

En primer lugar, (2.18) nos indica que $t \in [a_F(t|x), y]$ y que $F(t|x)$ debe ser < 1 en dicho intervalo, lo cuál, debido al carácter creciente de $F(t|x)$; se verificará si $y < b_F(t|x)$:

Por otro lado, ...jándonos en el término (2.19) debemos imponer que $\int_{a_F(t|x)}^y \frac{dH_1^a}{C} =$

Bajo la hipótesis de que $T; Y$ y C sean condicionalmente independientes a $X = x$, podemos escribir

$$C(t|x) = \int_0^t (x)^{i-1} L(t|x) \int_0^t H(t|x) \quad (2.21)$$

que será distinto de cero para $H(t|x) < 1$, $L(t|x) > 0$ y $(x)^i > 0$: Es fácil ver que necesitamos la condición $a_{L(t|x)} \cdot b_{H(t|x)}$ para que $(x)^i \neq 0$: También se comprueba fácilmente que $a_{H(t|x)} \cdot a_{F(t|x)}$, $a_{H^a(t|x)} = \max(a_{H(t|x)}, a_{L(t|x)})$ y $b_{H^a(t|x)} = b_{H(t|x)} \cdot b_{F(t|x)}$:

Entonces, si $y < b_{H(t|x)}$ y $a_{L(t|x)} \cdot a_{H(t|x)}$ se verificará:

$$\int_{a_F(t|x)}^y \frac{dH_1^a}{C} = \int_{a_F(t|x)}^y \frac{dH_1^a}{C} = \int_{a_F(t|x)}^y \frac{L(1 - G)dF}{L(1 - H)} = \int_{a_F(t|x)}^y \frac{dF}{1 - F} =$$

$$= \int_{a_H(t|x)}^y \frac{dF}{1 - F} + \int_{a_F(t|x)}^y \frac{dF}{1 - F} = \int_{a_F(t|x)}^y \frac{dF}{1 - F} = \int_{a_F(t|x)}^y \frac{dF}{1 - F}:$$

Estas consideraciones nos llevarán a suponer en el modelo de censura y truncamiento con covariables, que

$$\boxed{a_{L(j|x)} \cdot a_{H(j|x)}; y < b_{H(j|x)}}$$

Bajo estas condiciones, la expresión (2.19) nos permite definir un estimador natural para $\Lambda(: j x)$ que es

$$\begin{aligned} \hat{\Lambda}_n^{tc}(y j x) &= \frac{\sum_{i=1}^n \frac{d\hat{H}_{1n}^a(t j x)}{\hat{C}_n(t j x)}}{\sum_{i=1}^n \frac{1_{f_{Z_i} \cdot y; \pm_i=1} g_{B_{ni}}(x)}{\hat{C}_n(Z_i j x)}} = \\ &= \frac{\sum_{i=1}^n \frac{1_{f_{Z_i} \cdot y; \pm_i=1} g_{B_{ni}}(x)}{\sum_{j=1}^n 1_{f_{T_j} \cdot Z_i \cdot Z_j} g_{B_{nj}}(x)}} \end{aligned}$$

donde $\hat{H}_{1n}^a(y j x)$ y $\hat{C}_n(y j x)$ son, respectivamente, los estimadores no paramétricos de $H_1^a(y j x)$ y $C(y j x)$ dados por

$$\hat{H}_{1n}^a(y j x) = \sum_{i=1}^n 1_{f_{Z_i} \cdot y; \pm_i=1} g_{B_{ni}}(x) \quad (2.22)$$

y

$$\hat{C}_n(y j x) = \sum_{i=1}^n 1_{f_{T_i} \cdot y \cdot Z_i} g_{B_{ni}}(x) \quad (2.23)$$

siendo $f_{B_{ni}}(x)$ los pesos de la estimación no paramétrica (Nadaraya-Watson, por ejemplo).

Finalmente, sin más que tener en cuenta la relación entre $F(: j x)$ con $\Lambda(: j x)$ (específicamente en (2.16)), definimos el estimador de la función de distribución condicional bajo censura y truncamiento de forma que:

$$\begin{aligned} \hat{F}_n^{tc}(y j x) &= \frac{\sum_{i=1}^n \frac{1_{f_{Z_i} \cdot y; \pm_i=1} g_{B_{ni}}(x)}{\hat{C}_n(Z_i j x)}}{\sum_{i=1}^n \frac{1_{f_{T_i} \cdot y \cdot Z_i} g_{B_{ni}}(x)}} = \\ &= \frac{\sum_{i=1}^n \frac{1_{f_{Z_i} \cdot y; \pm_i=1} g_{B_{ni}}(x)}{\sum_{j=1}^n 1_{f_{T_j} \cdot Z_i \cdot Z_j} g_{B_{nj}}(x)}} \quad (2.24) \end{aligned}$$

Éste resulta ser un estimador de tipo límite producto, pero el peso de cada observación Z_i está influido por la cercanía a x de su correspondiente valor X_i ; de ahí que también llamemos al estimador (2.24) "estimador límite producto condicional (LPC) con censura y truncamiento".

2.3.2 Propiedades del estimador: características generales.

En este primer apartado presentamos algunas observaciones sencillas, pero bastante útiles, que se derivan del modelo presentado y de la propia definición del estimador $\hat{F}_n^{tc}(y|x)$. Todas ellas parten de las hipótesis ya mencionadas de independencia condicional a $X = x$ entre Y , T y C así como $a_{L(j|x)} \cdot a_{H(j|x)}$ e $y < b_{H(j|x)}$:

Propiedad 2.53 El estimador de la función de distribución condicional (LPC) bajo censura y truncamiento, definido en (2.24), coincide:

- con el estimador límite producto condicional con censura en el caso de ausencia de truncamiento.
- con el estimador de la función de distribución condicional con truncamiento en el caso de ausencia de censura.
- con el estimador tipo núcleo de la función de distribución condicional (para datos completos) en el caso de ausencia de censura y truncamiento.
- con el estimador de Turnbull en el caso de ausencia de covariables.

A partir de la expresión (2.24) del estimador y utilizando que cuando no hay truncamiento todos los $f_{T_i} g_{i=1}^n$ toman el valor 0 y las $f(X_i; Z_i; \pm_i) g_{i=1}^n$ constituyen una m.a.s. de tamaño n de la variable $(X; Z; \pm)$ se obtiene que:

$$\begin{aligned} \prod_{i=1}^n \hat{F}_n^{tc}(y|x) &= \prod_{i=1}^n \frac{\prod_{j=1}^n 1_{f_{Z_i} \cdot y g_{\pm_i} B_{ni}(x)}}{\prod_{j=1}^n 1_{f_{T_j} \cdot Z_i \cdot Z_j g_{B_{nj}(x)}}} = \\ &= \prod_{i=1}^n \frac{\prod_{j=1}^n 1_{f_{Z_i} \cdot y; \pm_i = 1 g_{B_{ni}(x)}}}{\prod_{j=1}^n 1_{f_{Z_i} \cdot Z_j g_{B_{nj}(x)}}} : \end{aligned}$$

La última expresión es igual al estimador límite producto condicional con censura (ecuación (2.13)) debido a que $\prod_{j=1}^n B_{nj}(x) = 1$: Esto prueba el apartado a).

El apartado b) se prueba partiendo de la expresión (2.24) del estimador y utilizando que cuando no hay censura la muestra $(X_i; T_i; Z_i; \pm_i) = (X_i; T_i; Y_i; 1)$; $\delta_i = 1$; $i = 1, \dots, n$: Así, se llega al estimador con truncamiento (expresión (2.14)):

$$\begin{aligned} \prod_{i=1}^n \hat{F}_n^{tc}(y|x) &= \prod_{i=1}^n \frac{\prod_{j=1}^n 1_{f_{Z_i} \cdot y g_{\pm_i} B_{ni}(x)}}{\prod_{j=1}^n 1_{f_{T_j} \cdot Z_i \cdot Z_j g_{B_{nj}(x)}}} = \\ &= \prod_{i=1}^n \frac{\prod_{j=1}^n 1_{f_{Y_i} \cdot y g_{B_{ni}(x)}}}{\prod_{j=1}^n 1_{f_{T_j} \cdot Y_i \cdot Y_j g_{B_{nj}(x)}}} : \end{aligned}$$

Para probar c) se procede de forma análoga. Utilizando que $T_i = 0$, $(Z_i; \pm_i) = (Y_i; 1)$; $\delta_i = 1$; $i = 1, \dots, n$ y n no es aleatorio, sino el tamaño del total de la muestra, se

llega a:

$$\begin{aligned}
 1_i \hat{F}_n^{tc}(y_j | x) &= \frac{\prod_{i=1}^n \tilde{A}_i \prod_{j=1}^n 1_{f_{Z_i} \cdot y_{g \pm i}} B_{ni}(x)}{\prod_{i=1}^n \tilde{A}_i \prod_{j=1}^n 1_{f_{T_j} \cdot Z_i \cdot Z_j g} B_{nj}(x)} = \\
 &= \frac{\prod_{i=1}^n \tilde{A}_i \prod_{j=1}^n 1_{f_{Y_i} \cdot y_g} B_{ni}(x)}{\prod_{i=1}^n \tilde{A}_i \prod_{j=1}^n 1_{f_{Y_i} \cdot Y_j g} B_{nj}(x)} = \\
 &= \frac{\prod_{i=Y_{i \cdot} \cdot y} \tilde{A}_i \prod_{j=1}^n 1_{f_{Y_i} \cdot Y_j g} B_{nj}(x) \prod_{i=1}^n B_{ni}(x)}{\prod_{j=1}^n \prod_{i=Y_{i \cdot} \cdot y} 1_{f_{Y_i} \cdot Y_j g} B_{nj}(x)} : \quad (2.25)
 \end{aligned}$$

Teniendo en cuenta que el estimador tipo núcleo de la función de distribución condicional para datos completos es

$$\hat{F}_n(y_j | x) = \prod_{i=1}^n B_{ni}(x) 1_{f_{Y_i} \cdot y_g}; \quad \text{donde} \quad \prod_{i=1}^n B_{ni}(x) = 1$$

y que ello implica la relación

$$\hat{F}_n(Y_i | j | x) = \hat{F}_n(Y_i^i | j | x) + B_{ni}(x);$$

se puede escribir la última expresión de (2.25) como

$$\begin{aligned}
 &\frac{\prod_{i=Y_{i \cdot} \cdot y} \tilde{A}_i \prod_{j=1}^n 1_{f_{Y_i} \cdot y_g} B_{ni}(x)}{\prod_{i=1}^n \tilde{A}_i \prod_{j=1}^n 1_{f_{Y_i} \cdot Y_j g} B_{nj}(x)} = \\
 &= \frac{\prod_{i=Y_{i:n} \cdot y} \tilde{A}_i \prod_{j=1}^n 1_{f_{Y_i} \cdot y_g} B_{ni}(x)}{\prod_{i=1}^n \tilde{A}_i \prod_{j=1}^n 1_{f_{Y_i} \cdot Y_j g} B_{nj}(x)} = \\
 &= \frac{\prod_{i=1}^n \tilde{A}_i \prod_{j=1}^n 1_{f_{Y_i} \cdot y_g} B_{ni}(x)}{\prod_{i=1}^n \tilde{A}_i \prod_{j=1}^n 1_{f_{Y_i} \cdot Y_j g} B_{nj}(x)} = \frac{\prod_{i=1}^n \tilde{A}_i \prod_{j=1}^n 1_{f_{Y_i} \cdot y_g} B_{ni}(x)}{\prod_{i=1}^n \tilde{A}_i \prod_{j=1}^n 1_{f_{Y_i} \cdot Y_j g} B_{nj}(x)} = \\
 &= \prod_{i=1}^n \tilde{A}_i \prod_{j=1}^n 1_{f_{Y_i} \cdot y_g} B_{ni}(x) = \prod_{i=1}^n \tilde{A}_i \prod_{j=1}^n 1_{f_{Y_i} \cdot y_g} B_{ni}(x);
 \end{aligned}$$

donde $k = \max \{i = Y_{i:n} \cdot y\}$: Lo que prueba el resultado deseado.

El apartado d) se prueba partiendo nuevamente de la expresión (2.24) del estimador y utilizando que cuando no hay covariables las X_i desaparecen y los pesos no paramétricos $B_{ni}(x) = 1/n; \delta_i = 1; \dots; n$: Así, se llega al estimador de Turnbull (expresión (1.32)):

$$\begin{aligned}
 1_i \hat{F}_n^{tc}(y_j | x) &= \frac{\prod_{i=1}^n \tilde{A}_i \prod_{j=1}^n 1_{f_{Z_i} \cdot y_{g \pm i}} \frac{1}{n}}{\prod_{i=1}^n \tilde{A}_i \prod_{j=1}^n 1_{f_{T_j} \cdot Z_i \cdot Z_j g} \frac{1}{n}} = \\
 &= \frac{\prod_{i=1}^n \tilde{A}_i \prod_{j=1}^n 1_{f_{Z_i} \cdot y_{g \pm i}}}{\prod_{i=1}^n \tilde{A}_i \prod_{j=1}^n 1_{f_{T_j} \cdot Z_i \cdot Z_j g}} = 1_i \hat{F}_n^{tc}(y);
 \end{aligned}$$

Propiedad 2.54 El estimador LPC tiene forma escalonada con saltos en las observaciones no censuradas. La magnitud del salto en Z_i , $d\hat{F}_n^{tc}(Z_i | j, x)$; es igual a

$$d\hat{F}_n^{tc}(Z_i | j, x) = \frac{\pm_i B_{ni}(x)}{\hat{C}_n(Z_i | j, x)} \quad 1_i \hat{F}_n^{tc}(Z_i^{i-1} | j, x) :$$

Esta expresión se obtiene al calcular $\hat{F}_n^{tc}(Z_i | j, x)$ y $\hat{F}_n^{tc}(Z_i^{i-1} | j, x)$ utilizando (2.24):

$$\begin{aligned} & \hat{F}_n^{tc}(Z_i | j, x) - \hat{F}_n^{tc}(Z_i^{i-1} | j, x) = \\ &= \prod_{j=1}^i \frac{\hat{F}_n^{tc}(Z_i^{i-1} | j, x) - \hat{F}_n^{tc}(Z_i^{i-2} | j, x)}{\hat{C}_n(Z_j | j, x)} - \prod_{j=1}^{i-1} \frac{\hat{F}_n^{tc}(Z_i^{i-1} | j, x) - \hat{F}_n^{tc}(Z_i^{i-2} | j, x)}{\hat{C}_n(Z_j | j, x)} = \\ &= \prod_{j=1}^{i-1} \frac{1_{f_{Z_j} < Z_i} g_{\pm j} B_{nj}(x)}{\hat{C}_n(Z_j | j, x)} - \prod_{j=1}^{i-1} \frac{1_{f_{Z_j} < Z_i} g_{\pm j} B_{nj}(x)}{\hat{C}_n(Z_j | j, x)} = \\ &= \prod_{j=1}^{i-1} \frac{1_{f_{Z_j} < Z_i} g_{\pm j} B_{nj}(x)}{\hat{C}_n(Z_j | j, x)} \left(1_i + \frac{\pm_i B_{ni}(x)}{\hat{C}_n(Z_i | j, x)} \right) = \\ &= \prod_{j=1}^{i-1} \hat{F}_n^{tc}(Z_i^{i-1} | j, x) \frac{\pm_i B_{ni}(x)}{\hat{C}_n(Z_i | j, x)} : \end{aligned}$$

Propiedad 2.55 El estimador LPC propicia una definición inmediata de un estimador para la función de regresión $g(x) = E(Y | j, x)$ con censura y truncamiento dado por

$$\hat{g}_n(x) = \sum_{i=1}^n Z_i \frac{\pm_i B_{ni}(x)}{\hat{C}_n(Z_i | j, x)} \prod_{j=1}^i \hat{F}_n^{tc}(Z_i^{i-1} | j, x) : \quad (2.26)$$

Basta tener en cuenta que $g(x) = \int_R y dF(y | j, x)$ y por tanto un estimador natural de $g(x)$ viene dado por

$$\hat{g}_n(x) = \int_R y d\hat{F}_n(y | j, x) :$$

Finalmente, utilizando el valor de los saltos de $\hat{F}_n(y | j, x)$; ya calculados, en la propiedad anterior se llega a la expresión (2.26).

Propiedad 2.56 El estimador LPC permite construir un estimador de la mediana condicional $Me(Y | j, x) = \inf \{y : F(y | j, x) \geq 0.5\}$; definido como

$$\hat{Me}_n(Y | j, x) = \inf \{y : \hat{F}_n^{tc}(y | j, x) \geq 0.5\} :$$

De forma análoga se pueden definir estimadores de los percentiles condicionales.

Propiedad 2.57 El estimador de tipo límite producto para la función de distribución condicionada a X de la variable $Z = \min(Y; C_g)$; $H(y | x)$; viene dado por

$$\hat{H}_n^{tc}(y | x) = \prod_{i=1}^n \frac{1_{f_{Z_i} \cdot y} B_{ni}(x)}{\prod_{j=1}^n 1_{f_{T_j} \cdot Z_i \cdot Z_j} B_{nj}(x)} = \prod_{i=1}^n \frac{1_{f_{Z_i} \cdot y} B_{ni}(x)}{\hat{C}_n(Z_i | x)} \quad (2.27)$$

y el salto de $\hat{H}_n^{tc}(y | x)$ en Z_i ; que denotaremos por $d\hat{H}_n^{tc}(Z_i | x)$ es

$$d\hat{H}_n^{tc}(Z_i | x) = \frac{B_{ni}(x)}{\hat{C}_n(Z_i | x)} \hat{H}_n^{tc}(Z_i^- | x) :$$

Estamos trabajando bajo las hipótesis de $a_{L(j|x)} \cdot a_{H(j|x)}$ e Y ; T y C condicionalmente independientes a $X = x$. Por tanto $Z = \min(Y; C_g)$ y T son condicionalmente independientes a $X = x$: Estas hipótesis constituyen las condiciones de identidad de $H(y | x)$ bajo truncamiento (generalización al caso condicional de la Propiedad 1.30): el papel de $F(y | x)$ ahora lo juega $H(y | x)$ y por tanto el estimador de H se obtiene sustituyendo en la expresión del estimador límite producto condicional con truncamiento (ecuación (2.14)) las variables Y_i por las Z_i :

En cuanto al salto de $\hat{H}_n^{tc}(y | x)$ en Z_i ; se obtiene de forma totalmente análoga al de $\hat{F}_n^{tc}(y | x)$ en Z_i (calculado en la Propiedad 2.54).

Propiedad 2.58 Si $b_{L(j|x)} \cdot b_{H(j|x)}$ el estimador de tipo límite producto para la función de distribución condicional, $L(y | x)$; de la variable de truncamiento T viene dado por

$$\hat{L}_n^{tc}(y | x) = \prod_{i=1}^n \frac{1_{f_{T_i} > y} B_{ni}(x)}{\hat{C}_n(T_i | x)} = \prod_{i=1}^n \frac{1_{f_{T_i} > y} B_{ni}(x)}{\prod_{j=1}^n 1_{f_{T_j} \cdot T_i \cdot Z_j} B_{nj}(x)} \quad (2.28)$$

y el salto de $\hat{L}_n^{tc}(y | x)$ en T_i ; que denotaremos por $d\hat{L}_n^{tc}(T_i | x)$ es

$$d\hat{L}_n^{tc}(T_i | x) = \frac{B_{ni}(x)}{\hat{C}_n(T_i | x)} \hat{L}_n^{tc}(T_i^- | x) :$$

Las hipótesis de trabajo implican que $Z = \min(Y; C_g)$ y T son condicionalmente independientes a $X = x$ y $b_{L(j|x)} \cdot b_{H(j|x)}$: Estas hipótesis constituyen las condiciones de identidad de $L(y | x)$ bajo truncamiento cuando el papel de Y ; con

función de distribución condicional $F(y | j, x)$; lo juega Z ; con función de distribución condicional $H(y | j, x)$). Por tanto el estimador de $L(y | j, x)$ se obtiene de forma análoga al de $L(y)$ (expresión (1.34)) pero de forma condicional.

El salto de $\hat{L}_n^{tc}(y | j, x)$ en T_i se obtiene de forma análoga al de $\hat{F}_n^{tc}(y | j, x)$ en Z_i (calculado en la Propiedad 2.54).

Propiedad 2.59 El estimador de tipo límite producto condicional para la función de distribución G de la variable C viene dado por

$$\hat{G}_n^{tc}(y | j, x) = \prod_{i=1}^n \frac{1_{f_{Z_i, y; \pm i=0}} B_{ni}(x)}{\hat{C}_n(Z_i | j, x)} \quad \text{para } y < b_{H(\cdot | j, x)} \quad (2.29)$$

El proceso es análogo al seguido para obtener el estimador condicional $\hat{F}_n^{tc}(y | j, x)$; consistente en encontrar una expresión que caracterice $G(y | j, x)$ y que dependa de cantidades estimables no paramétricamente. Dicha expresión viene dada por $\Lambda_G(y | j, x) = \int_0^y \frac{dG(t | j, x)}{1 - G(t | j, x)}$; expresión que determina la distribución $G(y | j, x)$ de forma única mediante la relación

$$1 - G(y | j, x) = e^{-\Lambda_G^c(y | j, x)} \prod_{a_i \in A} (1 - \Lambda_G(a_i | j, x)) \quad (2.30)$$

$a_i \in A = a_i \cdot y$

donde $\Lambda_G^c(y | j, x)$ denota la parte continua de $\Lambda_G(y | j, x)$, A el conjunto de puntos donde $\Lambda_G(y | j, x)$ tiene discontinuidades de salto y $\Lambda_G(a_i | j, x) = \Lambda_G(a_i | j, x) - \Lambda_G(a_i^- | j, x)$ la magnitud del salto de $\Lambda_G(y | j, x)$ en a_i ;

Además, si $a_{L(\cdot | j, x)} < a_{H(\cdot | j, x)}$ y para $y < b_{H(\cdot | j, x)}$; $\int_0^y \frac{dG(t | j, x)}{1 - G(t | j, x)}$ se puede escribir como

$$\begin{aligned} \int_0^y \frac{dG(t | j, x)}{1 - G(t | j, x)} &= \frac{\int_0^y \frac{dG(t | j, x)}{1 - G(t | j, x)} \mathbb{1}_{(x)}(T \cdot t \cdot Y | j, x) \mathbb{1}_{(x)}(1 - G(t | j, x))}{\int_0^y \frac{dG(t | j, x)}{1 - G(t | j, x)} \mathbb{1}_{(x)}(T \cdot t \cdot Y | j, x) \mathbb{1}_{(x)}(1 - G(t | j, x))} \\ &= \frac{dH_0^a(t | j, x)}{C(t | j, x)} \end{aligned} \quad (2.31)$$

sin más que utilizar las expresiones siguientes de $C(t | j, x)$ y $dH_0^a(t | j, x)$:

$$\begin{aligned} C(t | j, x) &= \mathbb{1}_{(x)}(T \cdot t \cdot Y | j, x) \mathbb{1}_{(x)}(1 - G(t | j, x)) \\ &= \mathbb{1}_{(x)}(T \cdot t \cdot Y | j, x) \mathbb{1}_{(x)}(1 - G(t | j, x)) \end{aligned}$$

y

$$\begin{aligned}
 H_0^{\alpha}(y | j, x) &= P(Z \cdot y; \pm = 0 | j, x; T \cdot Z) \\
 &= \frac{P(Z \cdot y; \pm = 0; T \cdot Z | j, x)}{P(T \cdot Z | j, x)} = \frac{P(C \cdot y; C \cdot Y; T \cdot C | j, x)}{P(T \cdot Z | j, x)} \\
 &= \int_{y_1}^{Z \cdot y} \frac{dH_0^{\alpha}(t | j, x)}{P(T \cdot t \cdot Y | j, x)} dG(t | j, x)
 \end{aligned}$$

respectivamente.

Sustituyendo ahora, en el último término de la igualdad (2.31) $H_0^{\alpha}(y | j, x)$ y $C(y | j, x)$ por sus respectivos estimadores no paramétricos y utilizando el resultado (2.30) se obtiene el estimador $\hat{G}_n^{tc}(y | j, x)$; para $y < b_{H(\cdot | j, x)}$:

Propiedad 2.60 El producto de los estimadores de las funciones $1_{\pm} F(y | j, x)$ y $1_{\pm} G(y | j, x)$, es igual al estimador de la función $1_{\pm} H(y | j, x)$, esto es:

$$1_{\pm} \hat{F}_n^{tc}(y | j, x) \cdot 1_{\pm} \hat{G}_n^{tc}(y | j, x) = 1_{\pm} \hat{H}_n^{tc}(y | j, x).$$

Basta multiplicar las expresiones (2.24) y (2.29) y tener en cuenta que para cada i desde 1 hasta n se verifica que $\pm_i = 0$ o $\pm_i = 1$. Entonces:

$$\begin{aligned}
 &= \prod_{i=1}^n \frac{1_{\pm_i} \hat{F}_n^{tc}(y | j, x)}{\hat{A}_n} \cdot \prod_{i=1}^n \frac{1_{\pm_i} \hat{G}_n^{tc}(y | j, x)}{\hat{A}_n} \\
 &= \prod_{i=1}^n \frac{1_{\pm_i} \frac{1_{f_{Z_i \cdot y; \pm_i=1g}} B_{ni}(x)}{\hat{C}_n(Z_i | j, x)}}{\hat{A}_n} \cdot \prod_{i=1}^n \frac{1_{\pm_i} \frac{1_{f_{Z_i \cdot y; \pm_i=0g}} B_{ni}(x)}{\hat{C}_n(Z_i | j, x)}}{\hat{A}_n} \\
 &= \prod_{i=1}^n \frac{1_{\pm_i} \frac{1_{f_{Z_i \cdot yg}} B_{ni}(x)}{\hat{C}_n(Z_i | j, x)}}{\hat{A}_n} \cdot \prod_{i=1}^n \frac{1_{\pm_i} \frac{1_{f_{Z_i \cdot yg}} B_{ni}(x)}{\hat{C}_n(Z_i | j, x)}}{\hat{A}_n} \\
 &= \prod_{i=1}^n \frac{1_{\pm_i} \frac{1_{f_{Z_i \cdot yg}} B_{ni}(x)}{\hat{C}_n(Z_i | j, x)}}{\hat{A}_n} = 1_{\pm} \hat{H}_n^{tc}(y | j, x):
 \end{aligned}$$

Propiedad 2.61 Siempre que tengan sentido los cocientes de las expresiones enunciadas a continuación, se verifica:

a) El salto de la función $\hat{H}_n^{tc}(y | j, x)$; en $y = Z_j$ es igual a

$$d\hat{H}_n^{tc}(Z_j | j, x) = \frac{B_{nj}(x) = \hat{L}_n^{tc}(Z_j | j, x)}{\prod_{i=1}^n B_{ni}(x) = \hat{L}_n^{tc}(Z_i | j, x)} \cdot 1_{\pm_j}$$

b) El salto de la función $\hat{L}_n^{tc}(y | x)$ en $y = T_j$ es igual a

$$d\hat{L}_n^{tc}(T_j | x) = \frac{B_{nj}(x) = \prod_{i=1}^n \hat{H}_n^{tc}(T_j^i | x)}{\prod_{i=1}^n B_{ni}(x) = \prod_{i=1}^n \hat{H}_n^{tc}(T_i^i | x)};$$

c) En las condiciones anteriores, se tiene

$$\prod_{i=1}^n \frac{B_{ni}(x)}{\hat{L}_n^{tc}(Z_i | x)} = \prod_{i=1}^n \frac{B_{ni}(x)}{\prod_{i=1}^n \hat{H}_n^{tc}(T_i^i | x)}; \quad (2.32)$$

Nota: Si $B_{nj}(x) = 0$; las expresiones de $d\hat{H}_n^{tc}(Z_j | x)$ y $d\hat{L}_n^{tc}(T_j | x)$ son siempre 0.

Es interesante observar que el apartado a) es el reflejo empírico de la relación

$$dH(y | x) = \frac{dH^a(y | x)}{L(y | x)} \otimes(x); \quad (2.33)$$

relación inmediata sin más que tener en cuenta que

$$\begin{aligned} H^a(y | x) &= P(Z \cdot y | x; T \cdot Z) = \frac{P(Z \cdot y; T \cdot Z | x)}{P(T \cdot Z | x)} = \\ &= \int_{i=1}^Z y \int_{\otimes(x)^{i-1}} \frac{dH^a(t | x)}{L(t | x)} dH(t | x); \end{aligned}$$

En cuanto al apartado b) es el reflejo empírico de

$$dL(y | x) = \frac{dL^a(y | x)}{\prod_{i=1}^n H(y^i | x)} \otimes(x); \quad (2.34)$$

igualdad que se basa en la relación siguiente

$$\begin{aligned} L^a(y | x) &= P(T \cdot y | x; T \cdot Z) = \frac{P(T \cdot y; T \cdot Z | x)}{P(T \cdot Z | x)} = \\ &= \int_{i=1}^Z y \int_{\otimes(x)^{i-1}} \frac{dL^a(t | x)}{\prod_{i=1}^n H(t^i | x)} dL(t | x); \end{aligned}$$

Además, integrando en R las expresiones (2.33) y (2.34) se llega a que

$$\otimes(x)^{i-1} = \int_R \frac{dH^a(y | x)}{L(y | x)} = \int_R \frac{dL^a(y | x)}{\prod_{i=1}^n H(y^i | x)};$$

La igualdad (2.32) resulta ser la transcripción empírica de la expresión anterior y por tanto proporciona un estimador de $\theta(x)^i$.

Comenzamos demostrando el apartado b). Para ello basta probar los dos resultados siguientes:

1. La expresión

$$\frac{1_i \hat{H}_n^{tc}(T_i^i | j, x) d\hat{L}_n^{tc}(T_i | j, x)}{B_{ni}(x)} = \text{cte} \quad \forall i = 1; 2; \dots; n \quad (2.35)$$

donde $d\hat{L}_n^{tc}(T_i | j, x)$ denota el salto de la función $\hat{L}_n^{tc}(y | j, x)$ en T_i :

2. Dicha constante resulta ser

$$\text{cte} = \frac{B_{ni}(x)}{1_i \hat{H}_n^{tc}(T_i^i | j, x)} \quad (2.36)$$

De este modo, utilizando (2.35) escribimos

$$d\hat{L}_n^{tc}(T_i | j, x) = \frac{B_{ni}(x)}{1_i \hat{H}_n^{tc}(T_i^i | j, x)} \text{cte}$$

y utilizando (2.36) llegamos a

$$\begin{aligned} d\hat{L}_n^{tc}(T_i | j, x) &= \frac{B_{ni}(x)}{1_i \hat{H}_n^{tc}(T_i^i | j, x)} \frac{B_{nj}(x)}{1_i \hat{H}_n^{tc}(T_j^i | j, x)} = \\ &= \frac{B_{ni}(x)}{1_i \hat{H}_n^{tc}(T_i^i | j, x)} \frac{1_j \hat{H}_n^{tc}(T_j^j | j, x)}{1_j \hat{H}_n^{tc}(T_j^i | j, x)} \end{aligned}$$

que es exactamente lo que dice el apartado b).

Para demostrar 1 suponemos, por simplicidad, que en la muestra no hay empates y ordenamos la muestra respecto a los valores de T_i ; muestra que denotaremos por $X_{[i]}, T_{(i)}, Z_{[i]}, \pm_{[i]} \quad i=1, \dots, n$. En cuanto a $B_{n[i]}(x)$ denota el peso no paramétrico asociado a $X_{[i]}$. Procediendo por inducción se tiene:

-Para un i arbitrario entre 1 y n :

$$\begin{aligned} &\frac{1_i \hat{H}_n^{tc}(T_{(i)}^i | j, x) d\hat{L}_n^{tc}(T_{(i)} | j, x)}{B_{n[i]}(x)} \\ &= \frac{1_i \hat{H}_n^{tc}(T_{(i)}^i | j, x) B_{n[i]}(x)}{B_{n[i]}(x) \hat{C}_n(T_{(i)} | j, x)} \hat{L}_n^{tc}(T_{(i)} | j, x) = \\ &= \frac{1_i \hat{H}_n^{tc}(T_{(i)}^i | j, x)}{\hat{C}_n(T_{(i)} | j, x)} \end{aligned} \quad (2.37)$$

-Para $i + 1$:

$$\begin{aligned} & \frac{\int_0^{\infty} 1_{\{i\}} \hat{H}_n^{tc}(T_{(i+1)}^i | j, x) d\hat{L}_n^{tc}(T_{(i+1)} | j, x)}{B_{n[i+1]}(x)} \\ &= \int_0^{\infty} 1_{\{i\}} \hat{H}_n^{tc}(T_{(i+1)}^i | j, x) \frac{\hat{L}_n^{tc}(T_{(i+1)} | j, x)}{\hat{C}_n(T_{(i+1)} | j, x)}; \end{aligned} \quad (2.38)$$

Teniendo en cuenta que

$$\begin{aligned} \int_0^{\infty} 1_{\{i\}} \hat{H}_n^{tc}(T_{(i+1)}^i | j, x) &= \int_{Z_j < T_{(i+1)}} \prod_{\substack{\tilde{A} \\ \tilde{A}}} 1_{\{i\}} \frac{B_{nj}(x)}{\hat{C}_n(Z_j | j, x)} = \\ &= \int_0^{\infty} 1_{\{i\}} \hat{H}_n^{tc}(T_{(i)}^i | j, x) \prod_{T_{(i)} \cdot Z_j < T_{(i+1)}} \prod_{\tilde{A}} 1_{\{i\}} \frac{B_{nj}(x)}{\hat{C}_n(Z_j | j, x)} \end{aligned}$$

y

$$\begin{aligned} \int_0^{\infty} 1_{\{i\}} \hat{L}_n^{tc}(T_{(i+1)} | j, x) &= \int_{T_j > T_{(i+1)}} \prod_{\tilde{A}} 1_{\{i\}} \frac{B_{nj}(x)}{\hat{C}_n(T_j | j, x)} = \\ &= \frac{\int_0^{\infty} 1_{\{i\}} \hat{L}_n^{tc}(T_{(i)} | j, x)}{\int_0^{\infty} 1_{\{i\}} \frac{B_{n[i+1]}(x)}{\hat{C}_n(T_{(i+1)} | j, x)}} = \frac{\int_0^{\infty} 1_{\{i\}} \hat{L}_n^{tc}(T_{(i)} | j, x) \hat{C}_n(T_{(i+1)} | j, x)}{\int_0^{\infty} 1_{\{i\}} \frac{B_{n[i+1]}(x)}{\hat{C}_n(T_{(i+1)} | j, x)}} \end{aligned}$$

podemos escribir (2.38) de la forma

$$\begin{aligned} & \int_0^{\infty} 1_{\{i\}} \hat{H}_n^{tc}(T_{(i)}^i | j, x) \int_{T_{(i)} \cdot Z_j < T_{(i+1)}} \prod_{\tilde{A}} 1_{\{i\}} \frac{B_{nj}(x)}{\hat{C}_n(Z_j | j, x)} \\ & \int_0^{\infty} 1_{\{i\}} \frac{\hat{L}_n^{tc}(T_{(i)} | j, x)}{\hat{C}_n(T_{(i+1)} | j, x) \int_0^{\infty} 1_{\{i\}} \frac{B_{n[i+1]}(x)}{\hat{C}_n(T_{(i+1)} | j, x)}} = \\ &= \int_0^{\infty} 1_{\{i\}} \hat{H}_n^{tc}(T_{(i)}^i | j, x) \frac{\int_0^{\infty} 1_{\{i\}} \hat{L}_n^{tc}(T_{(i)} | j, x)}{\hat{C}_n(T_{(i)} | j, x)} \int_{T_{(i)} \cdot Z_j < T_{(i+1)}} \prod_{\tilde{A}} 1_{\{i\}} \frac{B_{nj}(x)}{\hat{C}_n(Z_j | j, x)} \frac{\hat{C}_n(T_{(i)} | j, x)}{\hat{C}_n(T_{(i+1)} | j, x) \int_0^{\infty} 1_{\{i\}} \frac{B_{n[i+1]}(x)}{\hat{C}_n(T_{(i+1)} | j, x)}}; \end{aligned}$$

Por tanto, para probar el resultado 1 (ecuación (2.35)) es suficiente demostrar que

$$\prod_{T_{(i)} \cdot Z_j < T_{(i+1)}} \frac{B_{nj}(x)}{\hat{C}_n(Z_j | j, x)} = \frac{\hat{C}_n(T_{(i)} | j, x)}{\hat{C}_n(T_{(i+1)} | j, x) \prod_{i+1}^n B_{n[i+1]}(x)} = 1 \quad (2.39)$$

Si analizamos el conjunto de $Z_j = T_{(i)} \cdot Z_j < T_{(i+1)}$ vemos que se trata justamente de los $Z_j \in [T_{(i)}, T_{(i+1)})$ y dichos Z_j sólo pueden estar entre $Z_{[1]}, Z_{[2]}, \dots, Z_{[i]}$. Supongamos que dicho conjunto está formado por k elementos ordenados:

$$T_{(i)} < Z_{[m_1]} < Z_{[m_2]} < \dots < Z_{[m_k]} < T_{(i+1)}$$

entonces por la definición de $\hat{C}_n(y | j, x)$ se tiene que:

$$\begin{aligned} \hat{C}_n(T_{(i)} | j, x) &= \prod_{j=1}^k 1_{[T_{(i)}, T_{(i+1)})}(Z_{[j]}) B_{n[j]}(x) \\ \hat{C}_n(Z_{[m_1]} | j, x) &= \hat{C}_n(T_{(i)} | j, x) \\ \hat{C}_n(Z_{[m_2]} | j, x) &= \hat{C}_n(T_{[m_1]} | j, x) \prod_{i+1}^n B_{n[m_1]}(x) \\ \hat{C}_n(Z_{[m_3]} | j, x) &= \hat{C}_n(Z_{[m_2]} | j, x) \prod_{i+1}^n B_{n[m_2]}(x) \\ &\vdots \end{aligned}$$

$$\begin{aligned} \hat{C}_n(Z_{[m_k]} | j, x) &= \hat{C}_n(Z_{[m_{(k-1)}]} | j, x) \prod_{i+1}^n B_{n[m_{(k-1)}]}(x) \\ \hat{C}_n(T_{(i+1)} | j, x) &= \hat{C}_n(Z_{[m_k]} | j, x) \prod_{i+1}^n B_{n[m_k]}(x) + \prod_{i+1}^n B_{n[i+1]} \end{aligned}$$

Utilizando estos resultados podemos escribir (2.39) de la siguiente forma:

$$\begin{aligned} & \frac{\hat{C}_n(Z_{[m_1]} | j, x) \prod_{i+1}^n B_{n[m_1]}(x)}{\hat{C}_n(Z_{[m_1]} | j, x)} \cdot \frac{\hat{C}_n(Z_{[m_2]} | j, x) \prod_{i+1}^n B_{n[m_2]}(x)}{\hat{C}_n(Z_{[m_2]} | j, x)} \cdot \dots \\ & \cdot \frac{\hat{C}_n(Z_{[m_k]} | j, x) \prod_{i+1}^n B_{n[m_k]}(x)}{\hat{C}_n(Z_{[m_k]} | j, x)} \cdot \frac{\hat{C}_n(T_{(i)} | j, x)}{\hat{C}_n(T_{(i+1)} | j, x) \prod_{i+1}^n B_{n[i+1]}(x)} \\ & = \frac{1}{\hat{C}_n(Z_{[m_1]} | j, x)} \hat{C}_n(T_{(i)} | j, x) = 1 \end{aligned}$$

con lo cual hemos demostrado(2.35).

-Finalmente probaremos (2.36), resultado que se obtiene teniendo en cuenta que (2.35) implica

$$\int_{i=1}^{\infty} d\hat{L}_n^{tc}(T_i | j, x) = cte \int_{i=1}^{\infty} \frac{B_{ni}(x)}{1 - \hat{H}_n^{tc}(T_i | j, x)} dx$$

$$(\quad) \quad \frac{1}{cte} = \int_{i=1}^{\infty} \frac{B_{ni}(x)}{1 - \hat{H}_n^{tc}(T_i | j, x)} dx$$

La demostración del apartado a) es totalmente análoga a la demostración del apartado b) y por ello sólo indicaremos los pasos a seguir. Primero se prueba que

$$\frac{\int_{i=1}^{\infty} \hat{L}_n^{tc}(Z_i | j, x) d\hat{H}_n^{tc}(Z_i | j, x)}{B_{ni}(x)} = cte^0 \quad \forall i = 1; 2; \dots; n \quad (2.40)$$

y después, que dicha constante es

$$cte^0 = \frac{\int_{i=1}^{\infty} \frac{B_{ni}(x)}{\hat{L}_n^{tc}(Z_i | j, x)} dx}{\int_{i=1}^{\infty} \hat{L}_n^{tc}(Z_i | j, x) dx} \quad (2.41)$$

De este modo, se tiene, sin más que despejar en (2.40), que

$$d\hat{H}_n^{tc}(Z_i | j, x) = \frac{B_{ni}(x)}{\hat{L}_n^{tc}(Z_i | j, x)} cte^0$$

y sustituyendo en esta expresión el valor cte^0 dado por (2.41), se prueba el apartado a):

$$\hat{H}_n^{tc}(Z_i | j, x) = \frac{\int_{i=1}^{\infty} \frac{B_{ni}(x) \hat{L}_n^{tc}(Z_i | j, x)}{B_{ni}(x) \hat{L}_n^{tc}(Z_i | j, x)} dx}{\int_{i=1}^{\infty} \frac{B_{ni}(x) \hat{L}_n^{tc}(Z_i | j, x)}{B_{ni}(x) \hat{L}_n^{tc}(Z_i | j, x)} dx}$$

En cuanto al apartado c) se puede probar, por ejemplo, teniendo en cuenta que, en virtud de (2.35) y (2.40), se tiene

$$\frac{\int_{i=1}^{\infty} \frac{1 - \hat{H}_n^{tc}(T_i | j, x) d\hat{L}_n^{tc}(T_i | j, x)}{B_{ni}(x)} = \frac{cte^0 \int_{i=1}^{\infty} \frac{B_{ni}(x)}{1 - \hat{H}_n^{tc}(T_i | j, x)} dx}{\int_{i=1}^{\infty} \frac{B_{ni}(x) \hat{L}_n^{tc}(Z_i | j, x)}{B_{ni}(x) \hat{L}_n^{tc}(Z_i | j, x)} dx} \quad \forall i, j$$

Tomando el par $(T_{(n)}; Z_{[n]})$, la igualdad anterior equivale a

$$\frac{\prod_{i=1}^n \hat{H}_n^{tc}(T_{(n)}^i | j, x) \cdot \hat{L}_n^{tc}(T_{(n)} | j, x)}{\hat{C}_n(T_{(n)} | j, x)} = \frac{\prod_{i=1}^n \hat{L}_n^{tc}(Z_{[n]}^i | j, x) \cdot \prod_{i=1}^n \hat{H}_n^{tc}(Z_{[n]}^i | j, x)}{\hat{C}_n(Z_{[n]} | j, x)} \quad (2.42)$$

$$(\) \quad \frac{\hat{C}_n(Z_{[n]} | j, x)}{\hat{C}_n(T_{(n)} | j, x)} = \frac{\prod_{i=1}^n \hat{H}_n^{tc}(Z_{[n]}^i | j, x)}{\prod_{i=1}^n \hat{H}_n^{tc}(T_{(n)}^i | j, x)}$$

Para comprobar que la última igualdad se verifica estudiamos sus dos miembros por separado.

Denotando por $(T_{(n)}; Z_{[m_1]}, Z_{[m_2]}, \dots, Z_{[m_k]} | Z_{[n]}^c)$ y teniendo en cuenta la definición de $\hat{C}_n(y | j, x)$ se tiene que:

$$\begin{aligned} \hat{C}_n(Z_{[m_1]} | j, x) &= \hat{C}_n(T_{(n)} | j, x) \\ \hat{C}_n(Z_{[m_2]} | j, x) &= \hat{C}_n(Z_{[m_1]} | j, x) \cdot B_{n[m_1]}(x) = \hat{C}_n(T_{(n)} | j, x) \cdot B_{n[m_1]}(x) \\ &\vdots \end{aligned} \quad (2.43)$$

$$\begin{aligned} \hat{C}_n(Z_{[m_k]} | j, x) &= \hat{C}_n(Z_{[m_{(k-1)}]} | j, x) \cdot B_{n[m_{(k-1)}]}(x) = \\ &= \hat{C}_n(T_{(n)} | j, x) \cdot B_{n[m_1]}(x) \cdot \dots \cdot B_{n[m_{(k-1)}]}(x) \end{aligned}$$

$$\begin{aligned} \hat{C}_n(Z_{[n]} | j, x) &= \hat{C}_n(Z_{[m_k]} | j, x) \cdot B_{n[m_k]}(x) = \\ &= \hat{C}_n(T_{(n)} | j, x) \cdot B_{n[m_1]}(x) \cdot \dots \cdot B_{n[m_k]}(x) \end{aligned}$$

y por tanto el primer miembro de (2.42) es igual a

$$\frac{\hat{C}_n(Z_{[n]} | j, x)}{\hat{C}_n(T_{(n)} | j, x)} = \prod_{i=1}^k \frac{B_{n[m_i]}(x)}{\hat{C}_n(T_{(n)} | j, x)} \quad (2.44)$$

En cuanto al segundo miembro de (2.42), se puede escribir de la forma:

$$\frac{\prod_{i=1}^n \hat{H}_n^{tc}(Z_{[n]}^i | j, x)}{\prod_{i=1}^n \hat{H}_n^{tc}(T_{(n)}^i | j, x)} = \frac{\prod_{Z_j \in T_{(n)}} \prod_{Z_j \in Z_{[n]}} \frac{B_{nj}(x)}{\hat{C}_n(Z_j | j, x)}}{\prod_{Z_j \in Z_{[n]}} \frac{B_{nj}(x)}{\hat{C}_n(Z_j | j, x)}} =$$

$$\begin{aligned}
&= \prod_{j=1}^m \frac{B_{nj}(x)}{\hat{C}_n(Z_j | x)} \\
&= \frac{B_{n[m_1]}(x)}{\hat{C}_n(Z_{[m_1]} | x)} \frac{B_{n[m_2]}(x)}{\hat{C}_n(Z_{[m_2]} | x)} \cdots \frac{B_{n[m_k]}(x)}{\hat{C}_n(Z_{[m_k]} | x)}
\end{aligned}$$

que, utilizando nuevamente las relaciones de (2.43) es

$$\frac{B_{n[m_k]}(x)}{\hat{C}_n(Z_{[m_k]} | x)} = \frac{\hat{C}_n(Z_{[m_k]} | x)}{\hat{C}_n(Z_{[m_1]} | x)} = \frac{\hat{C}_n(Z_{[n]} | x)}{\hat{C}_n(T_{(n)} | x)} \quad (2.45)$$

La igualdad entre (2.44) y (2.45) prueba (2.42) y, por tanto, el apartado c).

2.3.3 Propiedades del estimador: aspectos asintóticos.

En este apartado obtenemos una representación casi segura del estimador $\hat{F}_n^{tc}(y | x)$ como suma ponderada de variables independientes más un término de error de orden $O\left(\frac{\ln n}{nh}\right)^{3-4}$. A partir de ella demostramos varias propiedades importantes: la consistencia uniforme fuerte y la normalidad asintótica del estimador, la convergencia débil del proceso y, también, algunos resultados de convergencia uniforme para las aproximaciones normales. Estos resultados están enunciados en forma de teoremas y se verifican bajo algunas de las hipótesis que a continuación enunciamos.

Hipótesis.

(H1) $X; Y; T$ y C son variables aleatorias reales y absolutamente continuas. Además, $Y; T$ y C se suponen no negativas.

(H2)a) Sea $I = [x_1; x_2]$ un intervalo contenido en el soporte de m^{\square} , verificando que existe un entorno $I_{\pm} = [x_1 \mp \epsilon; x_2 + \epsilon]$ con $\epsilon > 0$ tal que:

$$0 < \epsilon = \inf [m^{\square}(x) : x \in I_{\pm}] < \sup [m^{\square}(x) : x \in I_{\pm}] = \Gamma < 1$$

y $0 < \pm \Gamma < 1$. Y para todo $x \in I_{\pm}$ las variables $T; Y$ y C son condicionalmente independientes a $X = x$:

b) Además:

i) $a_{L(\cdot|x)} \cdot a_{H(\cdot|x)}, b_{L(\cdot|x)} \cdot b_{H(\cdot|x)}; \forall x \in I_{\pm}$:

ii) Existen $a, b \in \mathbb{R}$; con $a < b$ de forma que

$$\inf_{x \in I_{\pm}} \frac{f^{(i)}(x) (1 - H(b|x)) L(a|x)}{C(y|x)} > 0; \mu > 0;$$

(Recordemos que si $a_{L(j|x)} < y < b_{H(j|x)}$ entonces $C(y|x) = f^{(i)}(x) (1 - H(y|x)) L(y|x) > 0$, por ello la condición ii) dice que $C(y|x) > 0$ de modo uniforme en $[a; b] \in I_{\pm}$).

(H3) Las funciones $m(x)$ y $f^{(i)}(x)$ tienen primera derivada respecto a x continua en I_{\pm} y las funciones $L(y|x)$; $H(y|x)$ y $H_1(y|x)$ tienen primera derivada respecto a x continua y acotada en $(y|x) \in [0; 1] \in I_{\pm}$:

(H4) Las funciones $m(x)$ y $f^{(i)}(x)$ tienen segunda derivada respecto a x continua en I_{\pm} y las funciones $L(y|x)$; $H(y|x)$ y $H_1(y|x)$ tienen segunda derivada respecto a x continua y acotada en $(y|x) \in [0; 1] \in I_{\pm}$:

(H5) Las funciones $L(y|x)$; $H(y|x)$ y $H_1(y|x)$ tienen primera derivada respecto a y continua en $(y|x) \in [a; b] \in I_{\pm}$:

(H6) Las funciones $L(y|x)$; $H(y|x)$ y $H_1(y|x)$ tienen segunda derivada respecto a y continua en $(y|x) \in [a; b] \in I_{\pm}$:

(H7) Las funciones $L(y|x)$; $H(y|x)$ y $H_1(y|x)$ tienen derivada primero respecto a x y segundo respecto a y continua en $(y|x) \in [a; b] \in I_{\pm}$:

(H8) Las correspondientes funciones de densidad l , h y h_1 (impropia) de las funciones de distribución $L(y)$; $H(y)$ y de la subdistribución $H_1(y)$ están acotadas y son estrictamente positivas en $[a; b]$:

Trabajaremos con pesos no paramétricos del tipo Nadaraya-Watson, para los cuales se suponen:

(H9) La función núcleo K , es una función de densidad simétrica, que vale 0 fuera del intervalo $(-1; 1)$ y su variación total está acotada por $\int_{-1}^1 K(u) du < +1$:

(H10) El parámetro ventana $h = (h_n)$ verifica: $h \rightarrow 0$; $\ln n = (nh) \rightarrow 0$ y $nh^5 = \ln n = O(1)$:

Resultados.

Teorema 2.1 Bajo las hipótesis H1-H10 se verifica, para $x \in I_{\pm}$ y $y \in [a; b]$; que

a)

$$\begin{aligned} & \frac{h}{n} \sum_{i=1}^n \hat{\Lambda}_n^{tc}(y|x) \frac{1}{\Lambda(y|x)} - \frac{h}{n} \sum_{i=1}^n \hat{\Lambda}_n^{tc}(a|x) \frac{1}{\Lambda(a|x)} \\ &= \sum_{i=1}^n B_{ni}(x) \omega_a(Z_i; T_i; \pm; y; x) + R_{na}(y|x) \end{aligned}$$

donde

$$\omega(Z; T; \pm; y; x) = \frac{1_{fZ \cdot y; \pm=1g}}{C(Z|x)} \int_0^Z \frac{1_{fT \cdot u \cdot Zg}}{(C(u|x))^2} dH_1^n(u|x)$$

$$\mathbb{E}_a(Z; T; \pm; y; x) = \mathbb{E}(Z; T; \pm; y; x) \mid \mathbb{E}(Z; T; \pm; a; x)$$

y

$$\sup_{a, y, b: x \in I} |R_n(y; x)| = O\left(\frac{\tilde{A}_\mu \ln n}{nh}\right) \quad \text{c.s.}$$

b) Si $a < a_{H(j|x)}$, $x \in I$ entonces a) se reduce a la expresión

$$\hat{\Lambda}_n^{tc}(y; x) \mid \Lambda(y; x) = \sum_{i=1}^K B_{ni}(x) \mathbb{E}(Z_i; T_i; \pm; y; x) + R_n(y; x)$$

donde

$$\sup_{[a; b] \in I} |R_n(y; x)| = O\left(\frac{\tilde{A}_\mu \ln n}{nh}\right) \quad \text{c.s.}$$

c) Si $a < a_{H(j|x)}$, $x \in I$ entonces

$$\begin{aligned} \hat{F}_n^{tc}(y; x) \mid F(y; x) &= \\ &= (1 - F(y; x)) \sum_{i=1}^K B_{ni}(x) \mathbb{E}(Z_i; T_i; \pm; y; x) + R_n^0(y; x) \end{aligned}$$

donde

$$\sup_{[a; b] \in I} |R_n^0(y; x)| = O\left(\frac{\tilde{A}_\mu \ln n}{nh}\right) \quad \text{c.s.}$$

Teorema 2.2 Bajo las hipótesis del Teorema 2.1c) se verifica que

$$\sup_{[a; b] \in I} |\hat{F}_n^{tc}(y; x) \mid F(y; x)| = O\left(\frac{\tilde{A}_\mu \ln n}{nh}\right) \quad \text{c.s.}$$

Teorema 2.3 Bajo las hipótesis del Teorema 2.1c) se tiene para $x \in I$ e $y \in [a; b]$:

a) Si $nh^5 \rightarrow 0$ y $(\ln n)^3 = nh \rightarrow 0$ entonces

$$(nh)^{1-2} [\hat{F}_n^{tc}(y; x) \mid F(y; x)] \stackrel{d}{\rightarrow} N(0; s^2(y; x))$$

donde

$$s^2(y; x) = (1 - F(y; x))^2 \int_0^{\mu_Z} K^2(z) dz + \int_0^{\mu_Z} y \frac{dH_1^a(u; x)}{(C(u; x))^2} = m^a(x)$$

b) Si $h = Cn^{1-5}$ entonces

$$(nh)^{1-2} [\hat{F}_n^{tc}(y|x) - F(y|x)] \xrightarrow{d} N(b(y|x); s^2(y|x))$$

donde

$$b(y|x) = C^{5-2} (1 - F(y|x)) \frac{\int_0^R z^2 K(z) dz}{2m^2(x)} \int_0^1 \Phi^{(0)}(x) m^2(x) + 2\Phi^{(0)}(x) m^{(0)}(x)$$

y

$$\begin{aligned} \Phi(u) &= E \left(\int_0^T \frac{dH_1^2(t|u)}{C(t|x)} \cdot Z; X = u \right) \\ &= \int_0^y \frac{dH_1^2(t|u)}{C(t|x)} \int_0^y \frac{C(s|x)}{(C(s|x))^2} dH_1^2(s|x); \end{aligned}$$

Teorema 2.4 Bajo las hipótesis del Teorema 2.1c) y si $nh^5 \rightarrow 0$ y $(\ln n)^3 = nh \rightarrow 0$ se verifica para $x \in I$:

$$(nh)^{1-2} \hat{F}_n^{tc}(t|x) - F(t|x) \xrightarrow{d} N(t|x) \quad \text{en } D[a; b]$$

donde $N(t|x)$ denota un proceso Gaussiano de media cero y función de covarianzas

$$\Gamma(y; t|x) = (1 - F(y|x))(1 - F(t|x)) \frac{\int_0^R K^2(z) dz}{m^2(x)} \int_0^y \frac{dH_1^2(u|x)}{(C(u|x))^2}$$

con $y \wedge t = \min\{y, t\}$ y $D[a; b] = \{f \in C([a; b]; R) : f \text{ continua por la derecha y con discontinuidades, a lo sumo, de salto}\}$; con la topología de Skorohod (Billingsley, pag.111).

Teorema 2.5 Bajo las hipótesis del Teorema 2.1c) y tomando $h = Cn^{1-5}$ se verifica, para $x \in I$ e $y \in [a; b]$, que:

$$\begin{aligned} \text{a) } & \sup_{t \in R} P^{Y|X} \left[(nh)^{1-2} (1 - F(y|x)) \sum_{i=1}^n B_{ni}(x) \int_0^t \frac{dH_1^2(s|x)}{C(s|x)} \right. \\ & \left. - \int_0^y \frac{dH_1^2(s|x)}{C(s|x)} \right] = O \left(\frac{(\ln n)^{1-2}}{n^{1-5}} \right) \quad c.s. \end{aligned}$$

donde $P^{Y|X}$ denota la probabilidad condicionada a las observaciones $X_1; X_2; \dots; X_n$ y Φ_N denota la función de distribución de una variable $N(0; 1)$.

$$\text{b) } \sup_{t \in R} P^{Y|X} \left[(nh)^{1-2} \hat{F}_n^{tc}(y|x) - F(y|x) - \int_0^y \frac{dH_1^2(s|x)}{C(s|x)} \right] \rightarrow 0$$

2.3.4 Lemas y demostraciones de los principales resultados.

Hemos organizado este apartado de acuerdo con el siguiente esquema:

1. Enunciados de los lemas utilizados en la demostración del Teorema 2.1.
2. Demostración del Teorema 2.1.
3. Demostración de los Teoremas 2.2, 2.3 y 2.4.
4. Demostración de los lemas previos al Teorema 2.1.
5. Lemas utilizados en la demostración del Teorema 2.5.
6. Demostración del Teorema 2.5.

1. Enunciados de los lemas utilizados en la demostración del Teorema 2.1.

Lema 2.1 Bajo las hipótesis H1, H2a), H3, H4, H9 y H10 se verifica, con probabilidad 1, que:

$$\sup_{y \in (0; x_2]} \left| \hat{H}_{1n}^{\alpha}(y | x) - H_1^{\alpha}(y | x) \right| = O \left(\frac{\tilde{A} \mu_{\ln n}^{\alpha}}{nh} \right) + O(h^2) = O \left(\frac{\tilde{A} \mu_{\ln n}^{\alpha}}{nh} \right)$$

y

$$\sup_{y \in (0; x_2]} \left| \hat{C}_n(y | x) - C(y | x) \right| = O \left(\frac{\tilde{A} \mu_{\ln n}^{\alpha}}{nh} \right) + O(h^2) = O \left(\frac{\tilde{A} \mu_{\ln n}^{\alpha}}{nh} \right)$$

Lema 2.2 Sea Y una v.a. continua y no negativa con función de supervivencia condicionada a $X = x$, $S(t | x) = P(Y > t | X = x)$, donde X es una v.a. absolutamente continua con función de densidad m . Se define el estimador núcleo con pesos Nadaraya-Watson de $S(t | x)$ como $\hat{S}_n(t | x) = \hat{m}_n^{-1}(x) \hat{A}_n(t; x)$ donde

$$\hat{A}_n(t; x) = \frac{1}{nh} \sum_{i=1}^n 1_{\{Y_i > t\}} K \left(\frac{X_i - x}{h} \right)$$

y

$$\hat{m}_n(x) = \frac{1}{nh} \sum_{j=1}^n K \left(\frac{X_j - x}{h} \right)$$

Si se verifican las condiciones:

(h1) Sea $I = [x_1; x_2]$ un intervalo contenido en el soporte de m verificando que

$$0 < \gamma = \inf [m(x) : x \in I_{\pm}] < \sup [m(x) : x \in I_{\pm}] = \Gamma < 1$$

para algún entorno $I_{\pm} = [x_1 \pm \epsilon; x_2 \pm \epsilon]$ con $\epsilon > 0$ y $0 < \pm\Gamma < 1$:

(h2) Las funciones $m(x)$ y $S(y | x)$ tienen primera derivada respecto a x continua y acotada en $(y; x) \in [0; +1) \in I_{\pm}$:

(h3) Las funciones $m(x)$ y $S(y | x)$ tienen segunda derivada respecto a x continua y acotada en $(y; x) \in [0; +1) \in I_{\pm}$:

(h4) La función núcleo, K , es una función de densidad simétrica, que vale 0 fuera del intervalo $(-1; 1)$ y su variación total está acotada por $\int K < +1$:

(h5) El parámetro ventana $h = h_n$ verifica: $h \rightarrow 0$ y $\ln n = o(nh)$:

(h6) La función $S(y | x)$ tiene primera derivada respecto a y ; $S^{(0)}$; continua y acotada en $[0; +1) \in I_{\pm}$:

(h7) La función $S(y | x)$ tiene segunda derivada respecto a y ; $S^{(0)}$; continua y acotada en $[0; +1) \in I_{\pm}$:

(h8) La función $S^{(0)}$ admite primera derivada respecto a x , continua y acotada en $[0; +1) \in I_{\pm}$:

(h9) La función de densidad de Y , f_Y ; está acotada en $[0; +1)$ de forma que $0 < c \cdot f_Y \cdot d$

entonces

$$\begin{aligned} & \sup_{s,t \in [0; +1)} \sup_{|s-t| \leq b_n} \left| \hat{S}_n(s | x) - \hat{S}_n(t | x) - S(s | x) + S(t | x) \right| \quad (2.46) \\ & = O\left(\frac{\ln n}{nh}\right) + O\left(\frac{1}{b_n^2}\right) + O(b_n h) \end{aligned}$$

de modo uniforme en el intervalo I , donde $(b_n)_{n \geq 2N}$ es una sucesión con límite 0 y $b_n \frac{\ln n}{nh} \rightarrow \Delta > 0$.

Además si $nh^5 = \ln n = O(1)$ y $b_n \gg \frac{\ln n}{nh}$ se verifica, uniformemente en I y con probabilidad 1, que

$$\sup_{s,t \in [0; +1)} \sup_{|s-t| \leq b_n} \left| \hat{S}_n(s | x) - \hat{S}_n(t | x) - S(s | x) + S(t | x) \right| = O\left(\frac{\ln n}{nh}\right)$$

Lema 2.3 Bajo las hipótesis H1, H2a), H3-H10 se verifica, de modo uniforme en el intervalo I y con probabilidad 1, que:

$$\sup_{s,t \in [0; +1)} \sup_{|s-t| \leq b_n} \left| \hat{H}_n^{\alpha}(s | x) - \hat{H}_n^{\alpha}(t | x) - H_1^{\alpha}(s | x) + H_1^{\alpha}(t | x) \right| = O\left(\frac{\ln n}{nh}\right)$$

y

$$\sup_{s: 2[0; +1)} \sup_{j: s_j \leq t_j \leq b_n} \frac{|\hat{C}_n(s; j; x) - \hat{C}_n(t; j; x) - C(s; j; x) + C(t; j; x)|}{nh} = O\left(\frac{\ln n}{nh}\right)$$

donde $b_n \gg \frac{\ln n}{nh}$:

2. Demostración del Teorema 2.1.

a) Partimos de

$$\begin{aligned} & \int_a^h \frac{\hat{\Lambda}_n^{tc}(y; j; x) - \Lambda(y; j; x)}{\hat{C}_n(s; j; x)} dy - \int_a^h \frac{\hat{\Lambda}_n^{tc}(a; j; x) - \Lambda(a; j; x)}{C(s; j; x)} dy \\ &= \int_a^y \frac{d\hat{H}_{1n}^a(s; j; x)}{\hat{C}_n(s; j; x)} - \int_a^y \frac{dH_1^a(s; j; x)}{C(s; j; x)} \end{aligned} \quad (2.47)$$

sumando y restando $\int_a^y \frac{d\hat{H}_{1n}^a(s; j; x)}{C(s; j; x)}$, $\int_a^y \frac{dH_1^a(s; j; x)}{\hat{C}_n(s; j; x)}$ y agrupando términos (2.47) es igual a

$$\begin{aligned} & \int_a^y \frac{1}{\hat{C}_n(s; j; x)} - \frac{1}{C(s; j; x)} d\hat{H}_{1n}^a(s; j; x) + \int_a^y \frac{1}{C(s; j; x)} d\hat{H}_{1n}^a(s; j; x) - \int_a^y \frac{1}{C(s; j; x)} dH_1^a(s; j; x) + \\ & + \int_a^y \frac{1}{\hat{C}_n(s; j; x)} - \frac{1}{C(s; j; x)} dH_1^a(s; j; x) : \end{aligned}$$

Utilizando que

$$\frac{1}{\hat{C}_n(s; j; x)} - \frac{1}{C(s; j; x)} = \frac{C(s; j; x) - \hat{C}_n(s; j; x)}{(C(s; j; x))^2} + \frac{(C(s; j; x) - \hat{C}_n(s; j; x))^2}{(C(s; j; x))^2 \hat{C}_n(s; j; x)} \quad (2.48)$$

obtenemos que (2.47) es igual a

$$\begin{aligned} & \int_a^y \frac{d\hat{H}_{1n}^a(s; j; x)}{C(s; j; x)} - \int_a^y \frac{\hat{C}_n(s; j; x)}{(C(s; j; x))^2} dH_1^a(s; j; x) + \\ & + \int_a^y \frac{(C(s; j; x) - \hat{C}_n(s; j; x))^2}{(C(s; j; x))^2 \hat{C}_n(s; j; x)} dH_1^a(s; j; x) + \end{aligned}$$

$$+ \int_a^Z \frac{1}{\hat{C}_n(s; j; x)} \int \frac{1}{C(s; j; x)} d\hat{H}_{1n}^a(s; j; x) - \int dH_1^a(s; j; x) :$$

Los dos primeros sumandos corresponden a $\sum_{i=1}^n B_{ni}(x) \gg_a(Z_i; T_i; \pm_i; y; x)$, con la notación enunciada en el Teorema 2.1, y los dos últimos sumandos constituyen $R_{na}(y; j; x)$.

Estudiaremos por separado el orden de los dos sumandos de $R_{na}(y; j; x)$, que denotaremos $R_{1na}(y; j; x)$ y $R_{2na}(y; j; x)$, respectivamente.

$$\sup_{[a; b] \in I} |R_{1na}(y; j; x)| \leq \sup_{[a; b] \in I} |C(y; j; x) - \hat{C}_n(y; j; x)| \leq \epsilon$$

$$\leq \int_a^Z \frac{1}{(C(s; j; x))^2 \hat{C}_n(s; j; x)} dH_1^a(s; j; x) :$$

Nótese que para $a \leq y \leq b$ se verifica uniformemente que $C(y; j; x) \geq \mu > 0$ y además sabemos por el Lema 2.1 que

$$\sup_{[a; b] \in I} |\hat{C}_n(y; j; x) - C(y; j; x)| = O\left(\frac{\mu_{\ln n}^{-1}}{nh}\right) \quad c.s. :$$

lo cual nos dice que para un n suficientemente grande $|\hat{C}_n(y; j; x) - C(y; j; x)| \leq \mu$ uniformemente en $[a; b]$ de forma casi segura. Entonces

$$\sup_{[a; b] \in I} |R_{1na}(y; j; x)| \leq \frac{1}{\mu^3} \sup_{[a; b] \in I} |C(s; j; x) - \hat{C}_n(s; j; x)| = O\left(\frac{\mu_{\ln n}^{-1}}{nh}\right) \quad c.s. : \quad (2.49)$$

Trataremos ahora el término mas difícil, $R_{2na}(y; j; x)$.

Siguiendo la demostración de Lo y Singh (1986) obtenemos uniformemente en $x \in I$ que

$$\sup_{a \leq y \leq b} |R_{2na}(y; j; x)| \leq \epsilon \quad (2.50)$$

$$\leq \frac{k_n}{\mu^2} \sup_{[a; b] \in I} |C(y; j; x) - \hat{C}_n(y; j; x)| \leq \epsilon \quad (2.51)$$

$$E \max_{1 \leq i \leq k_n} j [\hat{H}_{1n}^{\alpha}(y_{i+1} | j, x) - H_1^{\alpha}(y_{i+1} | j, x)] + [\hat{H}_{1n}^{\alpha}(y_i | j, x) - H_1^{\alpha}(y_i | j, x)] j + \quad (2.52)$$

$$+ 2 \max_{1 \leq i \leq k_n} \sup_{y_i \leq y \leq y_{i+1}} \left| \frac{1}{\hat{C}_n(y | j, x)} - \frac{1}{C(y | j, x)} \right| + \left| \frac{1}{\hat{C}_n(y_i | j, x)} - \frac{1}{C(y_i | j, x)} \right| \quad (2.53)$$

donde $f[y_i; y_{i+1}]_{i=1}^{k_n}$ es una partición del intervalo $[a; b]$ en k_n intervalos donde $y_1 = a$, $y_{k_n+1} = b$ y $k_n \gg \frac{\ln n}{nh} \rightarrow 1$. Entonces:

$$(2.51) = O(1) \quad (2.54)$$

En cuanto al término (2.52) se puede acotar por

$$\max_{1 \leq i \leq k_n} \sup_{y_i \leq y \leq y_{i+1}} j [\hat{H}_{1n}^{\alpha}(y | j, x) - H_1^{\alpha}(y | j, x)] + [\hat{H}_{1n}^{\alpha}(y_i | j, x) - H_1^{\alpha}(y_i | j, x)] \quad (2.55)$$

y utilizando el Lema 2.3 obtenemos:

$$(2.55) = O \left(\frac{\mu_{\ln n}^{\alpha}}{nh} \right) \quad \text{c.s.} \quad (2.56)$$

Estudiamos, ahora, el orden de (2.53).

Utilizando (2.48) y que

$$\sup_{[a; b] \in I} \left| \frac{1}{\hat{C}_n(y | j, x)} - \frac{1}{C(y | j, x)} \right| = O \left(\frac{\mu_{\ln n}^{\alpha}}{nh} \right) \quad \text{c.s.}$$

obtenemos:

$$(2.53) \leq 2 \max_{1 \leq i \leq k_n} \sup_{y_i \leq y \leq y_{i+1}} \left| \frac{\hat{C}_n(y | j, x) - C(y | j, x)}{(C(y | j, x))^2} \right| + \left| \frac{\hat{C}_n(y_i | j, x) - C(y_i | j, x)}{(C(y_i | j, x))^2} \right| + O \left(\frac{\mu_{\ln n}^{\alpha}}{nh} \right) \quad \text{c.s.}$$

Además:

$$\begin{aligned} & \frac{\hat{C}_n(y | j, x) - C(y | j, x)}{(C(y | j, x))^2} + \frac{\hat{C}_n(y_i | j, x) - C(y_i | j, x)}{(C(y_i | j, x))^2} \\ &= \frac{\hat{C}_n(y | j, x) - C(y | j, x) + \hat{C}_n(y_i | j, x) - C(y_i | j, x)}{(C(y_i | j, x))^2} + \\ &+ \frac{\hat{C}_n(y | j, x) - C(y | j, x)}{(C(y | j, x))^2} + \frac{\hat{C}_n(y | j, x) - C(y | j, x)}{(C(y_i | j, x))^2} \end{aligned}$$

y este último término

$$= \frac{\frac{\hat{C}_n(y_j | x) - C(y_j | x)}{(C(y_j | x))^2} - \frac{\hat{C}_n(y_i | x) - C(y_i | x)}{(C(y_i | x))^2}}{(C(y_j | x))^2 - (C(y_i | x))^2} \frac{\hat{C}_n(y_j | x) - C(y_j | x)}{(C(y_j | x))^2} - \frac{\hat{C}_n(y_i | x) - C(y_i | x)}{(C(y_i | x))^2} :$$

Por tanto, (2.53) se puede acotar por

$$\begin{aligned} & \frac{2}{\mu^2} \max_{1 \leq i \leq k_n} \sup_{y_i \leq y \leq y_{i+1}} \left| \frac{\hat{C}_n(y_j | x) - C(y_j | x)}{(C(y_j | x))^2} - \frac{\hat{C}_n(y_i | x) - C(y_i | x)}{(C(y_i | x))^2} \right| + \\ & + \frac{2}{\mu^4} \sup_{[a; b] \in I} \left| \frac{\hat{C}_n(y_j | x) - C(y_j | x)}{(C(y_j | x))^2} - \frac{\hat{C}_n(y_i | x) - C(y_i | x)}{(C(y_i | x))^2} \right| \max_{1 \leq i \leq k_n} \sup_{y_i \leq y \leq y_{i+1}} \left| \frac{C(y_i | x)}{(C(y_j | x))^2} - \frac{C(y_j | x)}{(C(y_i | x))^2} \right| \\ & + O\left(\frac{\ln n}{nh}\right) \quad c.s. \end{aligned}$$

Aplicando nuevamente los Lemas 2.1 y 2.3, y las condiciones sobre la partición $f[y_i; y_{i+1}]_{g_{i=1}^{k_n}}$; se llega a

$$(2.53) = O\left(\frac{\tilde{\mu}_{\ln n}}{nh}\right)_{\mathbb{P}_{3=4}} + O\left(\frac{\mu_{\ln n}}{nh}\right)_{\mathbb{P}} = O\left(\frac{\tilde{\mu}_{\ln n}}{nh}\right)_{\mathbb{P}_{3=4}} \quad c.s.$$

Este resultado junto con (2.54) y (2.56) prueban que

$$(2.50) = O\left(\frac{\tilde{\mu}_{\ln n}}{nh}\right)_{\mathbb{P}_{3=4}} \quad c.s.$$

Y éste orden y (2.49) prueban el apartado a) del teorema.

b) Por la hipótesis H2 se tiene que $a_{L(t_j|x)} < a; \quad \forall x \in I$; Si además $a < a_{H(t_j|x)}, \quad \forall x \in I$; entonces $a_{L(t_j|x)} < a < a_{H(t_j|x)}; \quad \forall x \in I$; De esta forma el apartado b) es un caso particular del apartado a) para esta elección de a .

c) Denotemos por $\tilde{F}(y_j | x) = 1 - \exp(-\hat{\Lambda}_n^{tc}(y_j | x))$, entonces:

$$\hat{F}_n^{tc}(y_j | x) - F(y_j | x) = \hat{F}_n^{tc}(y_j | x) - \tilde{F}(y_j | x) + \tilde{F}(y_j | x) - F(y_j | x) = I + II$$

Teniendo en cuenta la relación expresada en (2.11) y utilizando un desarrollo de Taylor podemos escribir:

$$\begin{aligned}
 II &= \int_a^b \exp(i \hat{\Lambda}_n^{tc}(y|x)) - \exp(i \Lambda(y|x)) dy = \\
 &= \int_a^b \exp(i \hat{\Lambda}_n^{tc}(y|x)) - \exp(i \Lambda(y|x)) dy = \\
 &= \int_a^b \exp(i \Lambda(y|x)) \left[\exp(i (\hat{\Lambda}_n^{tc}(y|x) - \Lambda(y|x))) - 1 \right] dy \\
 &= \int_a^b \exp(i \Lambda(y|x)) \left[\exp(i \Lambda^{\square}(y|x)) - 1 \right] dy \\
 &= \int_a^b \exp(i \Lambda(y|x)) \left[\exp(i \Lambda^{\square}(y|x)) - 1 \right] dy \\
 &= II_1 + II_2
 \end{aligned}$$

donde $\Lambda^{\square}(y|x)$ está entre $\Lambda(y|x)$ y $\hat{\Lambda}_n^{tc}(y|x)$.

Resumiendo:

$$\hat{F}_n^{tc}(y|x) - F(y|x) = \int_a^b \exp(i \hat{\Lambda}_n^{tc}(y|x)) - \exp(i \Lambda(y|x)) dy + II_1 + II_2$$

De este modo el apartado c) se demuestra teniendo en cuenta:

1) $\int_a^b \exp(i \Lambda(y|x)) dy$ está acotado para $a \leq y \leq b$:

2) $\sup_{a \leq y \leq b} |II_2| = O\left(\frac{\ln n}{nh}\right)$ c.s.:

Consecuencia de

$$\sup_{[a;b] \in I} \left| \hat{\Lambda}_n^{tc}(y|x) - \Lambda(y|x) \right| = O\left(\frac{\tilde{A}_\mu \ln n}{nh}\right) \quad \text{c.s.}$$

Y esto se debe a que

$$\left[\hat{\Lambda}_n^{tc}(y|x) - \Lambda(y|x) \right] = \sum_{i=1}^{\infty} B_{ni}(x) \phi(Z_i; T_i; \pm_i; y; x) + R_n(y|x)$$

donde

$$\sup_{a \leq y \leq b; x \in I} |R_n(y|x)| = O\left(\frac{\tilde{A}_\mu \ln n}{nh}\right) \quad \text{c.s.}$$

junto con el hecho de que

$$\sup_{[a;b] \in I} \left| \sum_{i=1}^{\infty} B_{ni}(x) \phi_a(Z_i; T_i; \pm_i; y; x) \right| = O\left(\frac{\tilde{A}_\mu \ln n}{nh}\right) \quad \text{c.s.} \quad (2.57)$$

Esta última acotación se obtiene aplicando el Lema 2.1 repetidamente en la expresión:

$$\begin{aligned}
 & \int_{a_Z}^{a_Y} B_{ni}(x) \mathbb{1}_{(Z_i; T_i; \pm_i; y; x)} \\
 = & \int_{a_Z}^{a_Y} \frac{d\hat{H}_{1n}^a(s; j; x)}{C(s; j; x)} \int_{a_Z}^{a_Y} \frac{\hat{C}_n(s; j; x)}{(C(s; j; x))^2} dH_1^a(s; j; x) + \\
 & + \int_{a_Z}^{a_Y} \frac{1}{C(s; j; x)} dH_1^a(s; j; x) \int_{a_Z}^{a_Y} \frac{1}{C(s; j; x)} dH_1^a(s; j; x) \\
 = & \int_{a_Z}^{a_Y} \frac{d\hat{H}_{1n}^a(s; j; x)}{C(s; j; x)} \int_{a_Z}^{a_Y} \frac{\hat{C}_n(s; j; x)}{(C(s; j; x))^2} dH_1^a(s; j; x) = \\
 = & \int_{a_Z}^{a_Y} \frac{\hat{H}_{1n}^a(s; j; x)}{C(s; j; x)} \int_{a_Z}^{a_Y} \frac{\hat{H}_{1n}^a(s; j; x)}{(C(s; j; x))^2} dC(s; j; x) + \\
 & \int_{a_Z}^{a_Y} \frac{\hat{C}_n(s; j; x)}{(C(s; j; x))^2} dH_1^a(s; j; x)
 \end{aligned}$$

y tomando a tal que $a_{L(t; x)} < a < a_{H(t; x)}$:

$$3) \sup_{[a; b] \in I} |j| = O(1/nh) \quad c: s:$$

$$\begin{aligned}
 | & = \int_{a_Z}^{a_Y} \hat{F}_n^{tc}(y; j; x) \int_{a_Z}^{a_Y} \ln |1 - \exp(i \hat{\Lambda}_n^{tc}(y; j; x))| = \\
 & = \int_{a_Z}^{a_Y} \exp \ln |1 - \hat{F}_n^{tc}(y; j; x)| \int_{a_Z}^{a_Y} \exp(i \hat{\Lambda}_n^{tc}(y; j; x)) = \\
 & = \exp(i \Lambda^{aa}(y; j; x)) \int_{a_Z}^{a_Y} \ln |1 - \hat{F}_n^{tc}(y; j; x)| \int_{a_Z}^{a_Y} \hat{\Lambda}_n^{tc}(y; j; x)
 \end{aligned}$$

donde $\Lambda^{aa}(y; j; x)$ está entre $\ln |1 - \hat{F}_n^{tc}(y; j; x)|$ y $\hat{\Lambda}_n^{tc}(y; j; x)$.

Estudiamos:

$$\begin{aligned}
 & \int_{a_Z}^{a_Y} \ln |1 - \hat{F}_n^{tc}(y; j; x)| \int_{a_Z}^{a_Y} \hat{\Lambda}_n^{tc}(y; j; x) \\
 = & \int_{a_Z}^{a_Y} \ln |1 - \frac{\sum_{j=1}^n \mathbb{1}_{f_{T_j} \cdot Z_i \cdot Z_j g} B_{nj}(x)}{\sum_{j=1}^n \mathbb{1}_{f_{T_j} \cdot Z_i \cdot Z_j g} B_{nj}(x)}| \int_{a_Z}^{a_Y} \frac{\sum_{j=1}^n \mathbb{1}_{f_{T_j} \cdot Z_i \cdot Z_j g} B_{nj}(x)}{\sum_{j=1}^n \mathbb{1}_{f_{T_j} \cdot Z_i \cdot Z_j g} B_{nj}(x)} \\
 = & \int_{a_Z}^{a_Y} \mathbb{1}_{f_{Z_i} \cdot y g \pm i} \int_{a_Z}^{a_Y} \ln |1 - \frac{1}{\sum_{j=1}^n \mathbb{1}_{f_{T_j} \cdot Z_i \cdot Z_j g} B_{nj}(x)}| \int_{a_Z}^{a_Y} \frac{1}{\sum_{j=1}^n \mathbb{1}_{f_{T_j} \cdot Z_i \cdot Z_j g} B_{nj}(x)}
 \end{aligned} \tag{2.58}$$

Utilizando la desigualdad:

$$0 < \ln \frac{1}{x} \leq \frac{1}{x} \quad \text{para } x > 1$$

para $x = \frac{1}{n} \sum_{j=1}^n 1_{F_{T_j} \cdot Z_i \cdot Z_j} B_{nj}(x) = B_{ni}(x)$ cuando $B_{ni}(x) > 0$, obtenemos

$$(2.58) \quad \frac{\sum_{i=1}^n 1_{F_{Z_i} \cdot y \geq z_i} h^3 \sum_{j=1}^n 1_{F_{T_j} \cdot Z_i \cdot Z_j} B_{nj}(x)}{\sum_{j=1}^n 1_{F_{T_j} \cdot Z_i \cdot Z_j} B_{nj}(x)} \leq \frac{(B_{ni}(x))^2}{B_{ni}(x)}$$

donde el denominador es un \hat{C}_n corregido, es decir

$$\begin{aligned} \tilde{C}_n(Z_i; j; x) &= \sum_{j=1}^n 1_{F_{T_j} \cdot Z_i \cdot Z_j} B_{nj}(x) A_i B_{ni}(x) = \\ &= \sum_{j=1}^n 1_{F_{T_j} < Z_i} B_{ni}(x) \quad \sum_{j=1}^n 1_{F_{Z_j} < Z_i} B_{ni}(x): \end{aligned}$$

De este modo, y razonando de la misma forma que para obtener el orden de $\sup_y R_{1na}(y; j; x)$, se llega a

$$\sup_{a \leq y \leq b} (1) \cdot \text{cte} \max_{1 \leq i \leq n} B_{ni}(x) \int_a^z \frac{d\hat{H}_{1n}^{\mu}(s; j; x)}{\tilde{C}_n(s; j; x)} = O\left(\frac{1}{nh}\right) \quad \text{c.s.}$$

3. Demostración de los Teoremas 2.2, 2.3 y 2.4.

3.1. Demostración del Teorema 2.2.

El apartado c) del Teorema 2.1 dice que, para $x \in I$ e $y \in [a; b]$:

$$\hat{F}_n^{tc}(y; j; x) - F(y; j; x) = (1 - F(y; j; x)) \sum_{i=1}^n B_{ni}(x) \mu(Z_i; T_i; \pm_i; y; x) + R_n^0(y; j; x)$$

donde

$$\sup_{[a; b] \in I} |R_n^0(y; j; x)| = O\left(\frac{\ln n}{nh}\right) \quad \text{c.s.}$$

Por otro lado en la demostración de dicho apartado se ha probado que

$$\sup_{[a; b] \in I} |(1 - F(y; j; x)) \sum_{i=1}^n B_{ni}(x) \mu(Z_i; T_i; \pm_i; y; x)| = O\left(\frac{\ln n}{nh}\right) \quad \text{c.s.}$$

(ver (2.57) y tener en cuenta que $a_{L(tjx)} < a < a_{H(tjx)}$, $8 \times 2 \text{ I}$). De este modo se obtiene el resultado deseado de:

$$\sup_{[a;b] \in I} \hat{F}_n^{tc}(y_j | x) - F(y_j | x) = O\left(\frac{\tilde{A} \mu_{\ln n}^{11=2}}{nh}\right) \quad \text{c:s:}$$

3.2. Demostración del Teorema 2.3.

a) Estamos en condiciones de aplicar el apartado c) del Teorema 2.1, esto es, para $x \geq 1$ e $y \in [a;b]$:

$$\hat{F}_n^{tc}(y_j | x) - F(y_j | x) = (1 - F(y_j | x)) \sum_{i=1}^{\infty} B_{ni}(x) \mu(Z_i; T_i; \pm_i; y; x) + R_n^0(y_j | x)$$

donde

$$\sup_{a \leq y \leq b} \sum_{j=1}^{\infty} R_n^0(y_j | x) = O\left(\frac{\tilde{A} \mu_{\ln n}^{11=3=4}}{nh}\right) \quad \text{c:s:}$$

Como $(nh)^{1=2} \frac{\ln n}{nh} \mu_{\ln n}^{11=3=4} = \frac{\ln^3 n}{nh} \mu_{\ln n}^{1=4} \rightarrow 0$, la distribución asintótica de

$$\frac{1}{nh} \sum_{i=1}^{\infty} B_{ni}(x) \mu(Z_i; T_i; \pm_i; y; x)$$

coincide con la de

$$\begin{aligned} & \frac{1}{nh} (1 - F(y_j | x)) \sum_{i=1}^{\infty} B_{ni}(x) \mu(Z_i; T_i; \pm_i; y; x) \\ &= \frac{1}{nh} (1 - F(y_j | x)) \sum_{i=1}^{\infty} \frac{1}{nh} K \left(\frac{x_i - X_i}{h} \right) \mu(Z_i; T_i; \pm_i; y; x) = \hat{m}^p(x) \end{aligned}$$

donde el denominador denota el estimador de Parzen-Rosenblatt de la función de densidad de X condicionada a $T \cdot Z$. Es conocido que $\hat{m}^p(x)$ converge en probabilidad a $m^p(x)$ y por tanto, sólo resta estudiar la distribución asintótica del numerador que podemos descomponer de la forma:

$$\begin{aligned} & \frac{1}{nh} (1 - F(y_j | x)) \sum_{i=1}^{\infty} \frac{1}{nh} E \left[\mu \left(\frac{x_i - X_i}{h} \right) \mu(Z_i; T_i; \pm_i; y; x) \right] \\ &= \frac{1}{nh} (1 - F(y_j | x)) \sum_{i=1}^{\infty} \frac{1}{nh} E \left[\mu \left(\frac{x_i - X_i}{h} \right) \mu(Z_i; T_i; \pm_i; y; x) \right] + \\ &+ \frac{1}{nh} (1 - F(y_j | x)) \sum_{i=1}^{\infty} \frac{1}{nh} E \left[\mu \left(\frac{x_i - X_i}{h} \right) \mu(Z_i; T_i; \pm_i; y; x) \right] = I + II: \end{aligned}$$

Comenzamos estudiando el segundo sumando.

En primer lugar se tiene que:

$$\begin{aligned}
 & E \left[\frac{1}{nh} \sum_{i=1}^n K \left(\frac{x_i - X}{h} \right) \mu_{X_i} \right] \mathbb{1}_{\{Z \in [Z_j, Z_{j+1})\}} \\
 &= \frac{1}{h} E \left[\sum_{i=1}^n K \left(\frac{x_i - X}{h} \right) \mu_{X_i} \right] \mathbb{1}_{\{Z \in [Z_j, Z_{j+1})\}} \\
 &= \frac{1}{h} E \left[\sum_{i=1}^n K \left(\frac{x_i - X}{h} \right) \mu_{X_i} \right] \mathbb{1}_{\{Z \in [Z_j, Z_{j+1})\}} \\
 &= \frac{1}{h} \int K \left(\frac{x_i - u}{h} \right) E \left[\mu_{X_i} \mathbb{1}_{\{Z \in [Z_j, Z_{j+1})\}} \mid X = u \right] m^{\mu}(u) du \quad (2.59)
 \end{aligned}$$

donde

$$\begin{aligned}
 & E \left[\mu_{X_i} \mathbb{1}_{\{Z \in [Z_j, Z_{j+1})\}} \mid X = u \right] \\
 &= E \left[\frac{1_{FZ} \cdot y_{j+1} - 1_{FZ} \cdot y_j}{C(Z_j, X)} \int_{Z_j}^y \frac{1_{FT} \cdot s - Z_j}{(C(s, X))^2} dH_1^{\mu}(s, X) \mid X = u \right] \\
 &= \int_{Z_j}^y \frac{dH_1^{\mu}(t, u)}{C(t, X)} \int_{Z_j}^y \frac{C(s, X)}{(C(s, X))^2} dH_1^{\mu}(s, X) \quad (2.60)
 \end{aligned}$$

y por tanto $\Phi(u) = E \left[\mu_{X_i} \mathbb{1}_{\{Z \in [Z_j, Z_{j+1})\}} \mid X = u \right]$ es una función 2 veces derivable y con derivada segunda acotada respecto a u en un entorno de x (se escribe mediante operaciones de funciones que por la hipótesis H4 verifican las anteriores restricciones). Además $\Phi(x) = 0$; sin más que aplicar (2.60) con $u = x$: Teniendo en cuenta estas propiedades y utilizando cálculos estándar para derivar el sesgo de un estimador núcleo con pesos de Nadaraya-Watson, (2.59) se puede escribir de la forma:

$$\begin{aligned}
 & \frac{1}{h} \int K \left(\frac{x_i - u}{h} \right) \Phi(u) m^{\mu}(u) du = \int K(z) \Phi(x_j + hz) m^{\mu}(x_j + hz) dz = \\
 &= \Phi(x) m^{\mu}(x) + \frac{h^2}{2} \int z^2 K(z) dz \left[\Phi''(x) m^{\mu}(x) + \Phi''(x) m^{\mu}(x) + 2\Phi'(x) m^{\mu'}(x) \right] + \\
 &+ o(h^2) \\
 &= \frac{h^2}{2} \int z^2 K(z) dz \left[\Phi''(x) m^{\mu}(x) + 2\Phi'(x) m^{\mu'}(x) \right] + o(h^2) \quad (2.61)
 \end{aligned}$$

y entonces

$$\begin{aligned} II &= \frac{1}{nh} \sum_{i=1}^n (1 - F(y_j | x)) \int_{-\infty}^{\infty} z^2 K(z) dz \int_{-\infty}^{\infty} \Phi^{(0)}(x) m^{\alpha}(x) + 2\Phi^{(0)}(x) m^{\alpha(0)}(x) + o(h^2) \\ &= O\left(\frac{1}{nh^2}\right) = O\left((nh^5)^{1/2}\right) \end{aligned} \quad (2.62)$$

Bajo la hipótesis $nh^5 \rightarrow 0$ se tiene que $II = o(1)$ y entonces la distribución asintótica de $I + II$ coincide con la de I :

Es inmediato que

$$I = \sum_{i=1}^n \hat{\epsilon}_{i;n}(x; y)$$

donde las variables $\hat{\epsilon}_{i;n}(x; y)$ vienen dadas por

$$\hat{\epsilon}_{i;n}(x; y) = \frac{1}{nh} \sum_{j=1}^n (1 - F(y_j | x)) \int_{-\infty}^{\infty} z^2 K(z) dz \int_{-\infty}^{\infty} \Phi^{(0)}(x) m^{\alpha}(x) + 2\Phi^{(0)}(x) m^{\alpha(0)}(x) + o(h^2) \quad (2.63)$$

Para cada n , $\hat{\epsilon}_{i;n}(x; y)$ son n variables independientes de media 0. Si además $\text{Var} \hat{\epsilon}_{i;n}(x; y) = \frac{1}{n} \sigma_{i;n}^2 < +\infty$ para $i = 1, \dots, n$ y $\frac{1}{n} \sum_{i=1}^n \text{Var} \hat{\epsilon}_{i;n}(x; y)$ es positiva, estaremos en condiciones de aplicar el Teorema Central del Límite para disposiciones triangulares (teorema 7.2 de Billingsley, pag. 42) y obtener:

$$\frac{\sum_{i=1}^n \hat{\epsilon}_{i;n}(x; y)}{\sqrt{\sum_{i=1}^n \text{Var} \hat{\epsilon}_{i;n}(x; y)}} \xrightarrow{d} N(0, 1)$$

siempre que la condición de Lindeberg se verifique.

Primero, demostraremos que la varianza es finita:

$$\text{Var} \hat{\epsilon}_{i;n}(x; y) = (1 - F(y_j | x))^2 \frac{1}{n} \text{Var} \left[\int_{-\infty}^{\infty} z^2 K(z) dz \int_{-\infty}^{\infty} \Phi^{(0)}(x) m^{\alpha}(x) + 2\Phi^{(0)}(x) m^{\alpha(0)}(x) + o(h^2) \right]$$

y, razonando como en el cálculo de (2.61), se obtiene que

$$\begin{aligned}
 & \text{Var} \left[\frac{1}{h} K \left(\frac{X_i - X}{h} \right) \mid (Z; T; \pm; y; x) \right] \\
 &= E \left[\frac{1}{h^2} K^2 \left(\frac{X_i - X}{h} \right) \mid (Z; T; \pm; y; x) \right] \\
 &= \int_Z \frac{1}{h^2} K^2 \left(\frac{x_i - u}{h} \right) E \left[\mathbb{1}_{(Z; T; \pm; y; x)} \mid Z; X = u \right] m^{\pm}(u) du + O(h^4) \\
 &= \frac{1}{h} \int_Z K^2(z) dz E \left[\mathbb{1}_{(Z; T; \pm; y; x)} \mid Z; X = x \right] m^{\pm}(x) + O(h) + O(h^4) \\
 &= \frac{1}{h} \int_Z K^2(z) dz \text{Var} \left[\mathbb{1}_{(Z; T; \pm; y; x)} \mid Z; X = x \right] m^{\pm}(x) + O(h) \quad (2.64)
 \end{aligned}$$

Por tanto

$$\begin{aligned}
 & \text{Var} \left[\hat{F}_{i;n}(x; y) \right] \\
 &= \frac{1}{n} (1 - F(y \mid x))^2 \int_Z K^2(z) dz \text{Var} \left[\mathbb{1}_{(Z; T; \pm; y; x)} \mid Z; X = x \right] m^{\pm}(x) + \\
 &+ O \left(\frac{1}{n} \right) \quad (2.65)
 \end{aligned}$$

Finalmente

$$\begin{aligned}
 & \text{Var} \left(\mathbb{1}_{(Z; T; \pm; y; x)} \mid Z; X = x \right) \\
 &= E \left[\frac{1_{FZ \cdot y; \pm = 1g}}{C(Z \mid x)} \int_0^y \frac{1_{FT \cdot u \cdot Zg}}{(C(u \mid x))^2} dH_1^{\pm}(u \mid x) \mid Z; X = x \right]^2 \\
 &= E \left[\frac{1_{FZ \cdot y; \pm = 1g}}{(C(Z \mid x))^2} \mid Z; X = x \right] + \\
 &+ E \left[\int_0^y \frac{1_{FT \cdot u \cdot Zg}}{(C(u \mid x))^2} dH_1^{\pm}(u \mid x) \mid Z; X = x \right]^2 \\
 &= A + B + 2C:
 \end{aligned}$$

Veremos, al final de esta demostración que $B_j - 2C = 0$ y entonces

$$\begin{aligned} \text{Var}(\mu(Z; T; \pm; y; x) | T \cdot Z; X = x) &= A = E \frac{1_{FZ \cdot y; \pm=1g}}{(C(Z; j; x))^2} | X = x; T \cdot Z = \int_0^Z \frac{dH_1^{\pm}(u; j; x)}{(C(u; j; x))^2} \end{aligned} \quad (2.66)$$

Los resultados (2.65) y (2.66) llevan a que

$$\begin{aligned} \frac{3}{4}n^2 &= \sum_{i=1}^n \text{Var} f_{i;n}(x; y) \\ &= \int_0^Z \frac{1_{F(y; j; x)}^2}{(C(u; j; x))^2} K^2(z) dz \int_0^Z \frac{dH_1^{\pm}(u; j; x)}{(C(u; j; x))^2} m^{\pm}(x) + o(1) \\ &= \frac{3}{4}n^2 + o(1) \end{aligned}$$

lo cual prueba que la varianza $\frac{3}{4}n^2$ es finita.

En segundo lugar, verificaremos la condición de Lindeberg, esto es

$$\frac{1}{\frac{3}{4}n^2} \sum_{i=1}^n \int_{|f_{i;n}(x; y)| > \frac{3}{4}ng} f_{i;n}^2(x; y) dP_i \rightarrow 0 \quad (2.67)$$

Si definimos la función indicadora:

$$\begin{aligned} I_{i;n} &= 1_{|f_{i;n}(x; y)| > \frac{3}{4}ng} \\ &= 1^{\frac{1}{2}} \frac{1_{(nh)^i \cdot 1_{F(y; j; x)}^2 K^2(\frac{x_i; X_i}{n}) \gg (Z_i; T_i; \pm; y; x)_i E K^2(\frac{x_i; X_i}{n}) \gg (Z_i; T_i; \pm; y; x)_i > \frac{3}{4}n^2}}{2} \end{aligned}$$

(2.67) puede escribirse como

$$(2.67) = \frac{1}{\frac{3}{4}n^2} E \sum_{i=1}^n \int_{\frac{3}{4}ng}^{\frac{3}{4}ng} f_{i;n}^2(y; x) I_{i;n} dz$$

Debido a que $(nh)^i \rightarrow 0$ y las funciones K y \gg son acotadas, se tiene:

$$\begin{aligned} & \lim_{n \rightarrow \infty} \sum_{i=1}^n \int_{\frac{3}{4}ng}^{\frac{3}{4}ng} f_{i;n}^2(y; x) I_{i;n} dz = 0; \quad \forall w \text{ and } \forall i \in \{1, 2, \dots, ng\} \\ & \lim_{n \rightarrow \infty} \sum_{i=1}^n \int_{\frac{3}{4}ng}^{\frac{3}{4}ng} f_{i;n}^2(y; x) I_{i;n} dz = 0; \quad \forall w \\ & \lim_{n \rightarrow \infty} \sum_{i=1}^n \int_{\frac{3}{4}ng}^{\frac{3}{4}ng} f_{i;n}^2(y; x) I_{i;n} dz = 0 \end{aligned}$$

y por tanto

$$\lim_{n \rightarrow \infty} \frac{1}{n} E(\hat{\tau}_n) = 0$$

lo cual demuestra la condición de Lindeberg.

Finalmente, hemos probado que

$$\sqrt{n}(\hat{F}_n^{tc}(y|x) - F(y|x)) \xrightarrow{d} N(0; s^2(y|x))$$

donde

$$s^2(y|x) = (1 - F(y|x))^2 \int_0^y \frac{dH_1^u(u|x)}{(C(u|x))^2} = m^u(x):$$

Veremos ahora que $B = 2C = 0$:

El término B es igual a

$$\begin{aligned} & E \int_0^y \frac{1_{F_T \cdot u} \cdot Z_g}{(C(u|x))^2} dH_1^u(u|x) \int_0^y \frac{1_{F_T \cdot v} \cdot Z_g}{(C(v|x))^2} dH_1^v(v|x) | X = x; T \cdot Z \\ = & E \int_0^y \int_0^y \frac{1_{F_T \cdot u} \cdot Z_g}{(C(u|x))^2} \frac{1_{F_T \cdot v} \cdot Z_g}{(C(v|x))^2} dH_1^u(u|x) dH_1^v(v|x) | X = x; T \cdot Z \\ = & \int_0^y \int_0^y E \frac{1_{F_T \cdot u} \cdot Z_g 1_{F_T \cdot v} \cdot Z_g}{(C(u|x))^2 (C(v|x))^2} | X = x; T \cdot Z \end{aligned}$$

Estudiamos:

$$\begin{aligned} & E \frac{1_{F_T \cdot u} \cdot Z_g 1_{F_T \cdot v} \cdot Z_g}{(C(u|x))^2 (C(v|x))^2} | X = x; T \cdot Z \\ = & \int_{u \wedge v}^{u \vee v} \int_1^1 P(T \cdot u \wedge v; Z \cdot u \vee v; T \cdot Z | X = x) \\ = & \int_1^1 dH(s|x) dL(r|x) \\ = & \int_1^1 H((u \vee v)^i | x) L(u \wedge v | x) \end{aligned} \tag{2.68}$$

donde $u \wedge v = \min\{u, v\}$; $u \vee v = \max\{u, v\}$:

Volviendo a B, tenemos que

$$B = \int_0^y \int_0^y \frac{1_{F_T \cdot u} \cdot Z_g 1_{F_T \cdot v} \cdot Z_g}{(C(u|x))^2 (C(v|x))^2} dH_1^u(u|x) dH_1^v(v|x)$$

expresión que podemos descomponer en dos sumandos, integrando sobre los recintos $f_{u;v} \geq 0; y = u \cdot v$ y $f_{u;v} \geq 0; y = u > v$, respectivamente. Así obtenemos:

$$\begin{aligned} B &= \int_0^Z \frac{L(u; j, x)}{(C(u; j, x))^2} \int_0^u \frac{(1 - H(v; j, x))}{(C(v; j, x))^2} dH_1^a(v; j, x) dH_1^a(u; j, x) + \\ &+ \int_0^Z \frac{L(v; j, x)}{(C(v; j, x))^2} \int_u^Z \frac{(1 - H(u; j, x))}{(C(u; j, x))^2} dH_1^a(u; j, x) dH_1^a(v; j, x) \\ &= 2 \int_0^Z \frac{L(u; j, x)}{(C(u; j, x))^2} \int_0^u \frac{(1 - H(v; j, x))}{(C(v; j, x))^2} dH_1^a(v; j, x) dH_1^a(u; j, x): \end{aligned}$$

Estudiamos ahora el sumando C :

$$\begin{aligned} C &= \int_0^Z \frac{1_{fZ; y; \pm=1g}}{C(Z; j, x)} \int_0^Z \frac{1_{fT; u; Zg}}{(C(u; j, x))^2} dH_1^a(u; j, x) j X = x; T \cdot Z \\ &= \int_0^Z \frac{1_{fZ; y; \pm=1g} 1_{fT; u; Zg}}{C(Z; j, x)} j X = x; T \cdot Z \frac{dH_1^a(u; j, x)}{(C(u; j, x))^2} \\ &= \int_0^Z \frac{1_{fu; Z; y; \pm=1g} 1_{fT; ug}}{C(Z; j, x)} j X = x; T \cdot Z \frac{dH_1^a(u; j, x)}{(C(u; j, x))^2} \\ &= \int_0^Z \frac{L(u; j, x)}{(C(u; j, x))^2} \int_0^u \frac{dH_1(v; j, x)}{C(v; j, x)} \frac{dH_1^a(u; j, x)}{(C(u; j, x))^2} \\ &= \int_0^Z \frac{L(u; j, x)}{(C(u; j, x))^2} \int_0^u \frac{dH_1(v; j, x)}{C(v; j, x)} dH_1^a(u; j, x): \end{aligned}$$

Para que $B = 2C$ basta comprobar que

$$\int_0^Z \frac{(1 - H(v; j, x))}{(C(v; j, x))^2} dH_1^a(v; j, x) = \int_0^Z \frac{dH_1(v; j, x)}{C(v; j, x)}: \quad (2.69)$$

Las expresiones (2.17) y (2.20) indican que

$$\begin{aligned} dH_1^a(v; j, x) &= \int_0^v (1 - F(v; j, x)) dF(v; j, x) \\ C(v; j, x) &= \int_0^v (1 - F(v; j, x)) dF(v; j, x) \end{aligned}$$

y además, bajo independencia condicional a X de las variables Y y C ; se tiene:

$$\begin{aligned} \int_0^v (1 - F(v; j, x)) dF(v; j, x) &= \int_0^v (1 - F(v; j, x)) dF(v; j, x) \\ dH_1(v; j, x) &= \int_0^v (1 - F(v; j, x)) dF(v; j, x): \end{aligned}$$

Por tanto:

$$\begin{aligned} & \int_y^z \frac{(1 - H(v|j;x))}{(C(v|j;x))^2} dH_1^a(v|j;x) \\ &= \int_y^z \frac{(1 - F(v|j;x))(1 - G(v|j;x))^{(x)} P(T \cdot v \cdot C|j;x) dF(v|j;x)}{C(v|j;x)^{(x)} P(T \cdot v \cdot C|j;x) (1 - F(v|j;x))} \\ &= \int_y^z \frac{(1 - G(v|j;x)) dF(v|j;x)}{C(v|j;x)} = \int_y^z \frac{dH_1(v|j;x)}{C(v|j;x)} \end{aligned}$$

lo que prueba (2.69).

Una observación ...nal, es que chequeando A, B y C se ve que

$$E \left[\int_y^z \frac{dH_1^a(v|j;x)}{C(v|j;x)} \mid T \cdot Z; X = u \right]$$

es una función dos veces derivable con derivada segunda acotada en un entorno de $u = x$; lo cuál justifica el razonamiento utilizado para escribir (2.64).

b) Este apartado se obtiene usando el apartado anterior y sustituyendo h por $C_n^{i-1=5}$ en la expresión del término II (ecuación (2.62)). Además

$$\Phi^0(x) = \int_0^y \frac{dH_1^a(t|j;x)}{C(t|j;x)} + \int_0^z \frac{\dot{C}(s|j;x)}{(C(s|j;x))^2} dH_1^a(s|j;x)$$

y

$$\Phi^{00}(x) = \int_0^y \frac{d\ddot{H}_1^a(t|j;x)}{C(t|j;x)} + \int_0^z \frac{\ddot{C}(s|j;x)}{(C(s|j;x))^2} dH_1^a(s|j;x)$$

donde $\dot{H}_1^a(y|j;x)$ y $\dot{C}(y|j;x)$ denotan las derivadas de $H_1^a(y|j;x)$ y $C(y|j;x)$ con respecto a x , valoradas en el punto x ; y $\ddot{H}_1^a(y|j;x)$ y $\ddot{C}(y|j;x)$ denotan las segundas derivadas de $H_1^a(y|j;x)$ y $C(y|j;x)$ con respecto a x , valoradas en el punto x ; respectivamente.

3.3. Demostración del Teorema 2.4.

Podemos dar un paso más y considerar I un elemento aleatorio de $D[a; b] = \mathcal{F}^2 F([a; b]; R) = \mathcal{F}$ continua por la derecha y con discontinuidades, a lo sumo de saltos, es decir

$$\begin{aligned} I &= I_n : \Omega \rightarrow D[a; b] \\ w & \rightarrow I_n(w) : [a; b] \rightarrow R \\ & \text{y } [I_n(w)](y) = I_n(w; y) = \\ &= \sum_{i=1}^n (nh)^{i-1} z_{in}(y; x)(w) \end{aligned}$$

Cálculos totalmente análogos a los realizados en la demostración del teorema anterior para el estudio de $\text{Var}(\mathbb{1}(Z; T; \pm; y; x) | T \cdot Z; X = x)$ llevan a

$$E(\mathbb{1}(Z; T; \pm; t_j; x) \mathbb{1}(Z; T; \pm; t_k; x) | T \cdot Z; X = x) = \int_0^{t_j \wedge t_k} \frac{dH_1^\pm(u; j; x)}{(C(u; j; x))^2} \quad (2.72)$$

Para comprobarlo no hay más que desarrollar:

$$\begin{aligned} & E(\mathbb{1}(Z; T; \pm; t_j; x) \mathbb{1}(Z; T; \pm; t_k; x) | T \cdot Z; X = x) \\ &= E \left[\frac{1_{fZ \cdot t_j; \pm=1g}}{C(Z; j; x)} \int_0^{t_j} \frac{1_{fT \cdot u \cdot Zg}}{(C(u; j; x))^2} dH_1^\pm(u; j; x) \right] \\ &= E \left[\frac{1_{fZ \cdot t_k; \pm=1g}}{C(Z; j; x)} \int_0^{t_k} \frac{1_{fT \cdot u \cdot Zg}}{(C(u; j; x))^2} dH_1^\pm(u; j; x) \right] | T \cdot Z; X = x \\ &= E \left[\frac{1_{fZ \cdot t_j; \pm=1g} 1_{fZ \cdot t_k; \pm=1g}}{(C(Z; j; x))^2} | T \cdot Z; X = x \right] + \\ &+ E \left[\int_0^{t_j} \frac{1_{fT \cdot u \cdot Zg}}{(C(u; j; x))^2} dH_1^\pm(u; j; x) \int_0^{t_k} \frac{1_{fT \cdot u \cdot Zg}}{(C(u; j; x))^2} dH_1^\pm(u; j; x) | T \cdot Z; X = x \right] \\ &= E \left[\frac{1_{fZ \cdot t_j; \pm=1g}}{C(Z; j; x)} \int_0^{t_k} \frac{1_{fT \cdot u \cdot Zg}}{(C(u; j; x))^2} dH_1^\pm(u; j; x) | T \cdot Z; X = x \right] \\ &+ E \left[\frac{1_{fZ \cdot t_k; \pm=1g}}{C(Z; j; x)} \int_0^{t_j} \frac{1_{fT \cdot u \cdot Zg}}{(C(u; j; x))^2} dH_1^\pm(u; j; x) | T \cdot Z; X = x \right] \\ &= A + B \quad C_1 \quad C_2 \end{aligned}$$

y ver que $B = C_1 + C_2$ (probado al ...nal de la presente demostración de i)), junto con el resultado inmediato de

$$A = E \left[\frac{1_{fZ \cdot t_j; \pm=1g} 1_{fZ \cdot t_k; \pm=1g}}{(C(Z; j; x))^2} | T \cdot Z; X = x \right] = \int_0^{t_j \wedge t_k} \frac{dH_1^\pm(u; j; x)}{(C(u; j; x))^2}$$

Entonces, razonando como en el paso de (2.59) a (2.61) y utilizando que

$$\Phi(x) = E(\mathbb{1}(Z; T; \pm; t_j; x) | T \cdot Z; X = x) = 0$$

y (2.72), se obtiene:

$$\begin{aligned}
 & E \left(\sum_{i=1}^3 \ln(t_j; X) \sum_{i=1}^3 \ln(t_k; X) \right) \\
 &= (1 - F(t_j; X)) (1 - F(t_k; X)) E \\
 &= E \left[K^2 \frac{X_i - X_{i-1}}{h} \right] \left(Z_i; T_i; \pm_i; t_j; X \right) \left(Z_i; T_i; \pm_i; t_k; X \right) \\
 &= E \left[K \frac{X_i - X_{i-1}}{h} \right] \left(Z_i; T_i; \pm_i; t_j; X \right) E \left[K \frac{X_i - X_{i-1}}{h} \right] \left(Z_i; T_i; \pm_i; t_k; X \right) \\
 &= (1 - F(t_j; X)) (1 - F(t_k; X)) E \\
 &= h \int_0^Z K^2(u) du m^a(X) E \left(\left(Z; T; \pm; t_j; X \right) \left(Z; T; \pm; t_k; X \right) \mid T \cdot Z; X = X \right) \\
 &+ O(h^3) + O(h^3)^a \\
 &= (1 - F(t_j; X)) (1 - F(t_k; X)) h \int_0^Z K^2(u) du m^a(X) \int_0^{t_j \wedge t_k} \frac{dH_1^a(u; X)}{(C(u; X))^2} \\
 &+ O(h^3): \tag{2.73}
 \end{aligned}$$

Ahora, recordando (2.71) y teniendo en cuenta (2.73), se llega a

$$\begin{aligned}
 & \sum_{i=1}^3 E \left(X_{ni}^{(j)} X_{ni}^{(k)} \right) \\
 &= \sum_{i=1}^3 (nh)^{i-1} (1 - F(t_j; X)) (1 - F(t_k; X)) h \int_0^Z K^2(u) du m^a(X) E \\
 &= \sum_{i=1}^3 \int_0^{t_j \wedge t_k} \frac{dH_1^a(u; X)}{(C(u; X))^2} + \sum_{i=1}^3 (nh)^{i-1} O(h^3) \\
 &= (1 - F(t_j; X)) (1 - F(t_k; X)) \int_0^Z K^2(u) du \int_0^{t_j \wedge t_k} \frac{dH_1^a(u; X)}{(C(u; X))^2} m^a(X) + \\
 &+ O(h^2)
 \end{aligned}$$

es decir $\sum_{i=1}^3 E \left(X_{ni}^{(j)} X_{ni}^{(k)} \right) = a_{jk} \neq 0$ donde

$$\begin{aligned}
 a_{jk} &= (1 - F(t_j; X)) (1 - F(t_k; X)) \int_0^Z K^2(u) du E \\
 &= \int_0^{t_j \wedge t_k} \frac{dH_1^a(u; X)}{(C(u; X))^2} m^a(X): \tag{2.74}
 \end{aligned}$$

Veamos, antes de verificar ii), que $B = C_1 + C_2$. Respecto a B se tiene que:

$$E \int_0^Z \int_0^{t_j} \frac{1_{fT \cdot u} \cdot Z_g}{(C(u, j, x))^2} dH_1^a(u, j, x) \int_0^Z \int_0^{t_k} \frac{1_{fT \cdot u} \cdot Z_g}{(C(u, j, x))^2} dH_1^a(u, j, x) \mid X = x; T \cdot Z \\ = \int_0^Z \int_0^{t_j} \int_0^{t_k} \frac{1_{fT \cdot u} \cdot Z_g \cdot 1_{fT \cdot v} \cdot Z_g}{(C(u, j, x))^2 (C(v, j, x))^2} dH_1^a(u, j, x) dH_1^a(v, j, x):$$

Aplicando (2.68), la expresión anterior se puede escribir de la forma:

$$\int_0^Z \int_0^{t_j} \int_0^{t_k} \frac{1_{i_1} H^i(u, v) \cdot L(u \wedge v, j, x)}{(C(u, j, x))^2 (C(v, j, x))^2} dH_1^a(u, j, x) dH_1^a(v, j, x)$$

donde, una vez más, $u \wedge v = \min\{u, v\}$ y $u \vee v = \max\{u, v\}$. Descomponiendo ahora, la integral en dos sumandos: uno para los puntos donde $u \leq v$ y otro para aquellos con $u > v$, se tiene que:

$$B = \int_0^Z \int_0^{t_j} \int_0^{t_k} \frac{L(u, j, x)}{(C(u, j, x))^2} \int_0^Z \int_0^{t_k} \frac{(1_{i_1} H(v, j, x))}{(C(v, j, x))^2} dH_1^a(v, j, x) dH_1^a(u, j, x) + \\ + \int_0^Z \int_0^{t_j} \int_0^{t_k} \frac{L(v, j, x)}{(C(v, j, x))^2} \int_0^Z \int_0^{t_k} \frac{(1_{i_1} H(u, j, x))}{(C(u, j, x))^2} dH_1^a(u, j, x) dH_1^a(v, j, x):$$

Estudiamos ahora el sumando C_1 :

$$C_1 = E \int_0^Z \int_0^{t_j} \int_0^{t_k} \frac{1_{fZ \cdot t_j; \pm=1g}}{C(Z, j, x)} \int_0^Z \int_0^{t_k} \frac{1_{fT \cdot u} \cdot Z_g}{(C(u, j, x))^2} dH_1^a(u, j, x) \mid X = x; T \cdot Z \\ = \int_0^Z \int_0^{t_k} E \frac{1_{fZ \cdot t_j; \pm=1g} \cdot 1_{fT \cdot u} \cdot Z_g}{C(Z, j, x)} \mid X = x; T \cdot Z \int_0^Z \frac{dH_1^a(u, j, x)}{(C(u, j, x))^2} \\ = \int_0^Z \int_0^{t_k} E \frac{1_{f_u \cdot Z \cdot t_j; \pm=1g} \cdot 1_{fT \cdot u} \cdot Z_g}{C(Z, j, x)} \mid X = x; T \cdot Z \int_0^Z \frac{dH_1^a(u, j, x)}{(C(u, j, x))^2} \\ = \int_0^Z \int_0^{t_k} \int_0^{t_j} \frac{L(u, j, x)}{C(Z, j, x)} \int_0^Z \frac{1_{f_u \cdot v} \cdot Z_g}{C(v, j, x)} \int_0^Z \frac{dH_1^a(v, j, x)}{(C(v, j, x))^2} \\ = \int_0^Z \int_0^{t_k} \int_0^{t_j} \frac{L(u, j, x)}{(C(u, j, x))^2} \int_0^Z \frac{dH_1^a(v, j, x)}{C(v, j, x)} dH_1^a(u, j, x):$$

Análogamente

$$C_2 = \int_0^Z \int_0^{t_k} \int_0^{t_j} \frac{L(u, j, x)}{(C(u, j, x))^2} \int_0^Z \frac{dH_1^a(v, j, x)}{C(v, j, x)} dH_1^a(u, j, x):$$

Finalmente, teniendo en cuenta (2.69) se tiene que $B = C_1 + C_2$.

Ahora estudiamos la condición ii).

Definiendo la siguiente función indicadora:

$$J_{ni} = \int_{f_j(x_{ni}) > g} = 1^{1/2} P_{j=1}^p \int_{x_{ni}^{(j)} > g} =$$

$$= 1^{1/2} (nh)^{i-1} P_{j=1}^p \int K \left(\frac{x_i - x_{ni}^{(j)}}{h} \right) \gg (Z_i; T_i; \delta_i; t_j; x) E \int K \left(\frac{x_i - x_{ni}^{(j)}}{h} \right) \gg (Z_i; T_i; \delta_i; t_j; x) \int_{x_{ni}^{(j)} > g} =$$

la condición ii) equivale a probar la convergencia hacia 0, $g > 0$; de la expresión

$$\int_{i=1}^n \int_{f_j(x_{ni}) > g} \int_{x_{ni}^{(j)} > g} dP = \int_{i=1}^n \int_{j=1}^p \int_{x_{ni}^{(j)} > g} (nh)^{i-1} \int_{x_{ni}^{(j)} > g} dP =$$

$$= (nh)^{i-1} E \int_{j=1}^p \int_{x_{ni}^{(j)} > g} \int_{x_{ni}^{(j)} > g} dP$$

Y como $(nh)^{i-1} \rightarrow 0$ y las funciones K y \gg son acotadas $\delta_j = 1; 2; \dots; p$, se tiene que

$$\int_{j=1}^p \int_{x_{ni}^{(j)} > g} \int_{x_{ni}^{(j)} > g} dP = 0; \int_{j=1}^p \int_{x_{ni}^{(j)} > g} \int_{x_{ni}^{(j)} > g} dP = 0; \int_{j=1}^p \int_{x_{ni}^{(j)} > g} \int_{x_{ni}^{(j)} > g} dP = 0$$

lo cual implica que

$$\lim_{n \rightarrow \infty} (nh)^{i-1} E \int_{j=1}^p \int_{x_{ni}^{(j)} > g} \int_{x_{ni}^{(j)} > g} dP = 0$$

que es la verificación de la condición ii).

Como consecuencia de las condiciones i) e ii) hemos probado que

$$(I_n(t_1); I_n(t_2); \dots; I_n(t_p)) \xrightarrow{d} N(0; A)$$

donde $A = (a_{jk})_{p \times p}$ con a_{jk} de acuerdo a (2.74).

Para analizar la presente demostración debemos probar el tightness del proceso I_n .

Una condición suficiente a tal efecto es comprobar que

$$E (I_n(t_2) - I_n(t_1))^2 (I_n(t_2) - I_n(t_1))^2 \leq C [G_n(t_2) - G_n(t_1)]^2$$

para $t_1 \leq t \leq t_2$ y donde G_n es una sucesión de funciones no decrecientes que convergen a una función G no decreciente y continua (ver la particularización al caso de procesos empíricos unidimensionales del Teorema 3 de Bickel y Wichura (1971)).

La demostración de que dicha condición se verifica para el proceso que nos ocupa, se obtiene utilizando argumentos análogos a los que utiliza Stute (1997) en su Teorema 1.1 y Lema 5.1.

4. Demostración de los lemas previos al Teorema 2.1.

4.1. Demostración del Lema 2.1.

Este lema es consecuencia de la Propiedad 2.42, propiedad que también se verifica para el estimador de Nadaraya-Watson de cualquier subdistribución de una v.a. Y ; del tipo $P(Y \leq t; \pm = k | X = x)$; siendo \pm una v.a. indicadora ($k = 0$ o 1), sin más que transformar la variable \pm en una variable continua (Dabrowska, 1989).

La función \hat{H}_{1n}^{\pm} es un caso particular de los estimadores descritos para la subdistribución $H_1^{\pm}(y | x) = P(Z \leq y; \pm = 1 | x; T \leq Z)$ y las hipótesis de este lema implican las de la Propiedad 2.42 para $H_1^{\pm}(y | x)$: Así se obtiene la primera parte.

La función $C(t | x) = P(T \leq t | Z | x; T \leq Z)$ se puede escribir de la forma

$$C(t | x) = P(T \leq t | x; T \leq Z) | P(Z < t | x; T \leq Z)$$

y es, por tanto, diferencia de dos funciones de distribución. Su función estimadora \hat{C}_n verifica:

$$\hat{C}_n(t | x) = \sum_{i=1}^n 1_{f_{T_i} \leq t} B_{ni}(x) | \sum_{i=1}^n 1_{f_{Z_i} < t} B_{ni}(x)$$

y entonces

$$\hat{C}_n(t | x) | C(t | x) = [\hat{L}_n^{\pm}(t | x) | L^{\pm}(t | x)] | [\hat{H}_n^{\pm}(t | x) | H^{\pm}(t | x)] \quad (2.75)$$

donde \hat{L}_n^{\pm} y \hat{H}_n^{\pm} son, respectivamente, estimadores del tipo descrito de las funciones $L^{\pm}(t | x) = P(T \leq t | x; T \leq Z)$ y $H^{\pm}(t | x) = P(Z < t | x; T \leq Z)$.

Las hipótesis de este lema implican las de la Propiedad 2.42 para las funciones $L^{\pm}(t | x)$ y $H^{\pm}(t | x)$. Esto, junto con la desigualdad $\sup_i |a_i - b_i| \leq \sup_i |a_i| + \sup_i |b_i|$ prueba la segunda parte del presente lema.

4.2. Demostración del Lema 2.2.

Como en la demostración de la Propiedad 2.42 descomponemos la cantidad de interés de la forma:

$$\begin{aligned}
 & \hat{S}_n(s; x) \wedge \hat{S}_n(t; x) \wedge S(s; x) + S(t; x) \\
 = & \frac{\hat{A}_n(s; x)}{\hat{m}_n(x)} \wedge \frac{\hat{A}_n(t; x)}{\hat{m}_n(x)} \wedge \frac{A(s; x)}{m(x)} + \frac{A(t; x)}{m(x)} = \\
 = & \frac{\hat{A}_n(s; x) \wedge \hat{A}_n(t; x)}{\hat{m}_n(x)} \wedge \frac{A(s; x) \wedge A(t; x)}{m(x)} + \\
 & + \frac{A(s; x) \wedge A(t; x)}{\hat{m}_n(x)} \wedge \frac{A(s; x) \wedge A(t; x)}{\hat{m}_n(x)} \\
 = & \frac{\hat{A}_n(s; x) \wedge \hat{A}_n(t; x) \wedge A(s; x) + A(t; x)}{\hat{m}_n(x)} + \frac{A(s; x) \wedge A(t; x) \cdot m(x) \wedge \hat{m}_n(x)}{m(x) \cdot \hat{m}_n(x)} = \\
 = & S_1 + S_2
 \end{aligned}$$

Comenzamos estudiando S_2 :

En primer lugar, se tiene, de modo uniforme para $x \in I$, que

$$|m(x) \wedge \hat{m}_n(x)| = O\left(\frac{\mu_{\ln n}^{1/2}}{nh}\right) + O(h^2) \quad c:s:$$

(este resultado se prueba de forma implícita en la demostración de la Propiedad 2.42). Entonces utilizando (h1), para un n suficientemente grande, se verá que $\hat{m}_n(x) > 0$:

En segundo lugar:

$$\frac{A(s; x) \wedge A(t; x)}{m(x)} = jS(s; x) \wedge S(t; x) \cdot S^{(j)}(s; t) = O(b_n)$$

de modo uniforme en $[0; 1] \in I$; por la hipótesis (h6), y de este modo

$$|jS_2| \cdot \frac{|A(s; x) \wedge A(t; x)|}{m(x)} = O\left(\frac{\mu_{\ln n}^{1/2}}{b_n nh}\right) + O(b_n h^2) \quad c:s:$$

Pasamos al estudio del primer sumando, cuyo orden coincide con el del numerador:

$$\begin{aligned}
 & \hat{A}_n(s; x) \wedge \hat{A}_n(t; x) \wedge A(s; x) + A(t; x) = \\
 = & \hat{A}_n(s; x) \wedge \hat{A}_n(t; x) \wedge E\hat{A}_n(s; x) + E\hat{A}_n(t; x) + \\
 & + E\hat{A}_n(s; x) \wedge E\hat{A}_n(t; x) \wedge A(s; x) + A(t; x) \\
 = & S_{11} + S_{12}:
 \end{aligned}$$

Razonando como en (2.3) tenemos que S_{11} es igual a:

$$\begin{aligned} & \frac{1}{h} \int_{(y;u)}^Z K \frac{\mu_{x_i|u}}{h} 1_{fy>sg} dM_n(y;u) + \frac{1}{h} \int_{(y;u)}^Z K \frac{\mu_{x_i|u}}{h} 1_{fy>tg} dM_n(y;u) \\ & + \frac{1}{h} \int_{(y;u)}^Z K \frac{\mu_{x_i|u}}{h} 1_{fy>sg} dM(y;u) + \frac{1}{h} \int_{(y;u)}^Z K \frac{\mu_{x_i|u}}{h} 1_{fy>tg} dM(y;u) \\ = & \frac{1}{h} \int_{(y;u)}^Z K \frac{\mu_{x_i|u}}{h} 1_{fy>sg} 1_{fy>tg} (dM_n(y;u) + dM(y;u)) \\ = & \frac{1}{h} \int_{(y;u)}^Z K \frac{\mu_{x_i|u}}{h} 1_{fy>sg} 1_{fy>tg} (dM_n(y;u) + dM(y;u)) \\ = & \frac{1}{h} \int_{(y;u)}^Z \cdot \int_{B_s}^Z \int_t^u \frac{\mu_{x_i|u}}{h} (dM_n(y;u) + dM(y;u)) dK \frac{\mu_{x_i|u}}{h} \end{aligned}$$

donde B representa el intervalo comprendido entre x y y y y se restringe al intervalo $[x_i - h; x_i + h]$. Llamando

$$R(s; t; x; u) = [\min(s; t); \max(s; t)] \cap [\min(x; u); \max(x; u)]$$

obtenemos que S_{11} es igual a:

$$\begin{aligned} & \frac{1}{h} \int_{(y;u)}^Z \int_{B_s}^Z \int_t^u R(s; t; x; u) dK \frac{\mu_{x_i|u}}{h} \\ = & \frac{1}{h} \int_w^Z \int_n^u (T(M_1(s); M_1(t); M_2(x); w)) dK \frac{\mu_{x_i|u}}{h} \end{aligned}$$

donde $w \in [M_2(x_i - h); M_2(x); M_2(x + h); M_2(x)]$.

De este modo

$$jS_{11}j \cdot \sup_{T \leq J} j \int_n^u (S) j \frac{dw}{h}$$

donde $J = fT = [u_1; u_2] \cap [v_1; v_2] = [ju_1; u_2] \cap [v_1; v_2] \cdot Cb_n; [v_1; v_2] \cdot 2\Gamma h$; puesto que

$$ju_1; u_2j = jM_1(s); M_1(t)j \cdot jf_Y(\cdot)jjs; tj \cdot Cb_n$$

y

$$\begin{aligned} & [v_1; v_2] \cdot [M_2(x) + w(\max); M_2(x) + w(\min)]j = \\ & = [M_2(x + h); M_2(x); M_2(x - h) + M_2(x)]j \\ & = [M_2(x + h); M_2(x - h)]j = jm(\cdot)j2hj \cdot 2\Gamma h \end{aligned}$$

Llamando $k_k = \sup_{f \in \mathcal{F}; t_j, 0 \leq j \leq i} t_j \cdot b_n g \in I_j$, se tiene entonces que

$$k_{S_{11}k} \cdot \sup_{T \in \mathcal{T}_J} j_n^{\otimes n}(T) j \frac{z}{h}$$

y utilizando el resultado dado por (2.2) se llega a

$$P(k_{S_{11}k} > \epsilon) \leq P\left(\sup_{T \in \mathcal{T}_J} j_n^{\otimes n}(T) j \frac{z}{h} > \epsilon\right) = P\left(\sup_{T \in \mathcal{T}_J} j_n^{\otimes n}(T) j > \frac{\epsilon h}{c} \cdot \frac{1}{\min\{z, \frac{1}{2}\}}\right) \cdot \frac{c_1}{[\min(Cb_n, 2\Gamma h)]^2} \exp\left\{-\frac{c_2 n h^{2\alpha} \epsilon^2}{2\Gamma h C b_n}\right\} \quad (2.76)$$

donde

$$\begin{aligned} \frac{1}{2}(a) &= \sup_{T \in \mathcal{T}_J} \int_{\mathcal{A}} dC(u; v) = \\ &= \sup_{f; v \in \mathcal{T} = \mathcal{T}_J g} c(u; v) \quad 2\Gamma h C b_n = \Gamma_0 2\Gamma h C b_n \end{aligned}$$

y Γ_0 es una constante (ya que (h1) y (h9) llevan a que la densidad c es acotada).
Entonces:

$$\begin{aligned} (2.76) &\leq \frac{c_1}{\min(C^2 b_n^2, 4\Gamma^2 h^2)} \exp\left\{-\frac{c_2 n h^{2\alpha} \epsilon^2}{2\Gamma_0 2\Gamma h C b_n}\right\} = \\ &= \frac{c_1}{\min(h^2, b_n^2)} \exp\left\{-\frac{c_2 n h^{2\alpha} \epsilon^2}{b_n}\right\} \quad (2.77) \end{aligned}$$

siempre que $2 \cdot \frac{\epsilon h}{c} \leq \frac{\epsilon h}{c} \cdot e_0 \cdot \frac{1}{2}(a)$ para alguna constante positiva e_0 :

Analicemos las dos últimas condiciones.

Tomando $\epsilon = c^\alpha \left(\frac{\ln n}{nh}\right)^{\frac{1}{1-2\alpha}} b_n$ la primera restricción se verifica puesto que:

$$2 \cdot \frac{\epsilon h}{c} \leq 2 \cdot \frac{c^\alpha h}{c} \left(\frac{\ln n}{nh}\right)^{\frac{1}{1-2\alpha}} b_n = \frac{2 c^{\alpha-1} h}{c} \left(\frac{\ln n}{nh}\right)^{\frac{1}{1-2\alpha}} b_n = \frac{2}{c^\alpha} \cdot (nh b_n \ln n)^{1-2\alpha}$$

y teniendo en cuenta que por hipótesis $b_n \left(\frac{\ln n}{nh}\right)^{\frac{1}{1-2\alpha}} > \Delta > 0$ se tiene, para n suficientemente grande,

$$nh b_n \ln n = \frac{nh (\ln n)^2 b_n}{\ln n} = \frac{nh b_n}{\ln n} (\ln n)^2 > \Delta (\ln n)^2 > \frac{2}{c^\alpha} \cdot \frac{1}{2}$$

La segunda restricción equivale a

$$\begin{aligned}
 & \zeta \cdot e_o \rho_{\bar{n}^{1/2}(a)}, \rho_{\bar{n}''} \frac{h}{\bar{n}} \cdot e_o \rho_{\bar{n}} \Gamma_0 2 \Gamma h C b_n \\
 & , \rho_{\bar{n}-c^a} \mu_{\frac{\ln n}{nh} b_n} \mathbb{1}_{1=2} \cdot e_o \rho_{\bar{n}} \Gamma_0 2 \Gamma h C b_n \\
 & , \mu_{\frac{\ln n}{nh} b_n} \mathbb{1}_{1=2} \cdot e_o \Gamma_0 2 \Gamma C b_n \frac{1}{c^a} \\
 & , \frac{\ln n}{nh b_n} \cdot e_o^2 \Gamma_0 2 \Gamma C \frac{1}{c^a} \mathbb{1}_2, e_o^2 \frac{(\ln n) = nh b_n}{\Gamma_0 2 \Gamma C \frac{1}{c^a}} \quad (2.78)
 \end{aligned}$$

pues bajo las hipótesis de este lema $\frac{1}{2}(a) = \Gamma_0 2 \Gamma h C b_n$:

Entonces, tomando $e_o^2 \frac{1}{\Phi(2i_{oi} C \frac{1}{c^a})^2}$ y utilizando $b_n i \frac{\ln n}{nh} \zeta_{i-1} > \Delta > 0$ obtenemos (2.78).

Ahora, estamos en condiciones de aplicar (2.77) para $\mu = c^a i \frac{\ln n}{nh} b_n \zeta_{1=2}$; y así obtener:

$$\begin{aligned}
 (2:77) & = \frac{c_1^0}{\min(h^2; b_n^2)} \exp \frac{1}{2} i \frac{c_2^0 nh c^{a2} b_n \ln n}{nh b_n} \frac{3}{4} = \frac{c_1^0}{\min(h^2; b_n^2)} \exp \circ i c_2^0 c^{a2} \ln n^a = \\
 & = \frac{c_1^0}{\min(h^2; b_n^2) n^{c_2^0 c^{a2}}}
 \end{aligned}$$

que es sumable sin más que tomar un c^a apropiado.

Aplicando el lema de Borel-Cantelli se tiene que

$$\|kS_{11}k = O \frac{\tilde{A} \mu_{\frac{\ln n}{nh} b_n} \mathbb{1}_{1=2}}{c:S}$$

Para terminar la demostración estudiamos el orden de $kS_{12}k$:

$$\begin{aligned}
 S_{12} & = E \hat{A}_n(s; x) i E \hat{A}_n(t; x) i A(s; x) + A(t; x) = \\
 & = \frac{1}{h} S(s; j u) K \frac{x i u}{h} m(u) du i \frac{1}{h} S(t; j u) K \frac{x i u}{h} m(u) du i \\
 & i S(s; j x) m(x) + S(t; j x) m(x) \\
 & = \frac{1}{h} (S(s; j u) i S(t; j u)) K \frac{x i u}{h} m(u) du i \\
 & i (S(s; j x) i S(t; j x)) m(x):
 \end{aligned}$$

Además utilizando las hipótesis (h6) y (h7) podemos escribir:

$$S(s_j | x)_i - S(t_j | x) = S^0(t_j | x)(s_i - t) + \frac{1}{2} S^{00}(\gg_j | x)(s_i - t)^2$$

y por tanto, $|jS_{12j}|$ se puede acotar por la siguiente suma:

$$\begin{aligned} & \left| \frac{1}{h} \int_{s_i}^Z S^0(t_j | u)(s_i - t) \left[K \frac{\mu_{x_i | u}}{h} m(u) du - S^0(t_j | x) m(x) \right] \right| + \\ & \left| \frac{1}{h} \int_{s_i}^Z \frac{1}{2} S^{00}(\gg_j | u)(s_i - t)^2 \left[K \frac{\mu_{x_i | u}}{h} m(u) du - \frac{1}{2} S^{00}(\gg_j | x)(s_i - t)^2 m(x) \right] \right| \\ & = |jS_{121j}| + |jS_{122j}| \end{aligned}$$

Para el primer sumando se tiene:

$$|jS_{121j}| \leq |jS_{i-} | t_j| \left| \frac{1}{h} \int_{s_i}^Z K \frac{\mu_{x_i | u}}{h} S^0(t_j | u) m(u) du - S^0(t_j | x) m(x) \right|$$

y como S^0 es derivable respecto a x con derivada acotada y m también, es un resultado conocido de la estimación núcleo con pesos Nadaraya-Watson que:

$$\left| \frac{1}{h} \int_{s_i}^Z K \frac{\mu_{x_i | u}}{h} S^0(t_j | u) m(u) du - S^0(t_j | x) m(x) \right| = O(h):$$

Así se obtiene de modo uniforme en $[0; +1) \in I$ que

$$|jS_{121j}| = O(b_n h):$$

De la definición y acotación de las funciones involucradas, se cumple:

$$|jS_{122j}| = O(b_n^2) \quad \text{en } [0; +1) \in I:$$

Resumiendo, hemos demostrado que

$$\begin{aligned} \|kS_{2k}\| &= O \left(\frac{\tilde{A} \mu_{\ln n}}{nh} \right)_{l=2} + O \left(\frac{1}{b_n h^2} \right) \quad \text{c.s.} \\ \|kS_{11k}\| &= O \left(\frac{\tilde{A} \mu_{\ln n}}{nh} b_n \right)_{l=2} \quad \text{c.s.} \\ \|kS_{12k}\| &= O(b_n h) + O \left(\frac{1}{b_n^2} \right) \end{aligned}$$

y estos resultados junto con

$$\frac{b_n \left(\frac{\ln n}{nh}\right)^{1/2}}{b_n \left(\frac{\ln n}{nh}\right)^{1/2}} = b_n^{1/2} = o(1)$$

prueban la primera parte del lema.

Bajo las hipótesis $nh^5 \ln n = O(1)$ y $b_n \gg \left(\frac{\ln n}{nh}\right)^{1/2}$ se tiene:

i) $O\left(\frac{\ln n}{nh} b_n\right)^{1/2} = O\left(\frac{\ln n}{nh}\right)^{3/4}$

ii) $O\left(b_n^2\right)^{1/2} = O\left(\frac{\ln n}{nh}\right)^{1/2}$

iii) $O(b_n h) = O\left(h^2 \frac{\ln n}{nh}\right)^{1/2}$ y $h \left(\frac{\ln n}{nh}\right)^{1/2} = \left(\frac{\ln n}{nh}\right)^{3/4} = \frac{nh^5}{\ln n}^{1/4} = O(1)$

que prueban la segunda parte.

4.3. Demostración del Lema 2.3.

Es consecuencia del Lema 2.2 aplicado a las funciones $H_1^\alpha(t | x)$, $H^\alpha(t | x)$ y $L^\alpha(t | x)$ y sus respectivos estimadores $\hat{H}_{1n}^\alpha(t | x)$, $\hat{H}_n^\alpha(t | x)$ y $\hat{L}_n^\alpha(t | x)$, dándose cuenta que las consideraciones hechas para la demostración del Lema 2.1 tienen cabida en la presente prueba.

5. Lemas utilizados en la demostración del Teorema 2.5.

Lema 2.4 (desigualdad de Bernstein). Sean $W_1; W_2; \dots; W_n$ v.a. independientes verificando que $P\{jW_i | E(W_i)\} \leq m$, para cada $i = 1; 2; \dots; n$; donde $m < 1$. Entonces, para $t > 0$;

$$P\left\{\sum_{i=1}^n W_i - \sum_{i=1}^n E(W_i) \geq nt\right\} \leq 2 \exp\left\{-\frac{nt^2}{2 \sum_{i=1}^n \text{Var}(W_i) + \frac{2}{3}mnt}\right\};$$

para todo $n = 1; 2; \dots$

Lema 2.5 (desigualdad de Berry-Essén). Si $V_1; V_2; \dots; V_n$ son v.a. independientes verificando $EV_i = 0; E V_i^2 = \sigma_i^2 > 0$ y $E |V_i|^3 < 1; \forall i$, entonces existe una constante A (universal) tal que

$$\sup_{z \in \mathbb{R}} \left| P\left\{\sum_{i=1}^n \frac{V_i}{\sigma_i} \leq z\right\} - \Phi_N(z)\right| \leq A \frac{\sum_{i=1}^n E |V_i|^3}{\sum_{i=1}^n \sigma_i^3}$$

donde Φ_N denota la función de distribución de la $N(0; 1)$:

Lema 2.6 Sean $\alpha_1, \alpha_2, \beta_1 > 0$ y $\beta_2 > 0$ números reales. Denotando por Φ_N y \hat{A}_N las funciones de distribución y densidad de la normal estándar, respectivamente, se veri...ca que

$$\sup_{z \in \mathbb{R}} \left| \Phi_N \left(\frac{z - \mu}{\sigma} \right) - \hat{A}_N \left(\frac{z - \mu}{\sigma} \right) \right| \leq k \hat{A}_N \left(\frac{z - \mu}{\sigma} \right) \left[\frac{1}{\sigma} + \frac{1}{\sigma^2} \max(\beta_1, \beta_2) + \frac{1}{\sigma^3} \right] + kt \hat{A}(t) k_1 \left(\frac{z - \mu}{\sigma} \right) \left[\frac{1}{\sigma} + \frac{1}{\sigma^2} \max(\beta_1, \beta_2) \right]$$

Demostración. Es el Lema 2.4 de la Tesis de Cao Abad (1990).

Lema 2.7 Bajo las hipótesis del Teorema 2.1c) y tomando $h = Cn^{-1/5}$ se veri...ca, para $x \in I$ e $y \in [a; b]$, que:

$$\sup_{t \in \mathbb{R}} \left| P^{Y|X} \left(\frac{t - \tilde{A}}{v_n^{1/2}(y|x)} \right) - \Phi_N \left(\frac{t - \tilde{A}}{v_n^{1/2}(y|x)} \right) \right| = O(n^{-2/5}) \quad c:s:$$

donde

$$b_n(y|x) = (nh)^{1/2} (1 - F(y|x)) \prod_{i=1}^n B_{ni}(x) E^{Y|X} \left(Z_i; T_i; \pm_i; y; x \right);$$

$$v_n(y|x) = nh (1 - F(y|x))^2 \prod_{i=1}^n (B_{ni}(x))^2 \text{Var}^{Y|X} \left(Z_i; T_i; \pm_i; y; x \right)$$

y $P^{Y|X}$; $E^{Y|X}$ y $\text{Var}^{Y|X}$ denotan, respectivamente, las probabilidades, esperanzas y varianzas de ciertas variables dependientes condicionadas a las observaciones $X_1; X_2; \dots; X_n$:

Demostración.

Aplicaremos la desigualdad de Berry- Essén a las variables

$$V_i = \frac{(nh)^{1/2} (1 - F(y|x)) (B_{ni}(x) \prod_{j=1}^n B_{nj}(x))}{(nh)^{1/2} (1 - F(y|x)) \prod_{j=1}^n B_{nj}(x)} \left(Z_i; T_i; \pm_i; y; x \right) - E^{Y|X} \left(Z_i; T_i; \pm_i; y; x \right)$$

$$= (nh)^{1/2} (1 - F(y|x)) \prod_{j=1}^n B_{nj}(x) \left(Z_i; T_i; \pm_i; y; x \right) - E^{Y|X} \left(Z_i; T_i; \pm_i; y; x \right)$$

que abreviadamente escribiremos como

$$V_i = (nh)^{1-2} (1 - F(y_j | x)) B_{ni}(x) \mu_{i-1}^{Y|X}(y; x) E^{Y|X} \mu_{i-1}^{Y|X}(y; x) :$$

Así obtenemos que el supremo del enunciado se acota por

$$\begin{aligned} & \frac{(nh)^{3-2} (1 - F(y_j | x))^3 \prod_{i=1}^n (B_{ni}(x))^3 E^{Y|X} \mu_{i-1}^{Y|X}(y; x) E^{Y|X} \mu_{i-1}^{Y|X}(y; x)^{-3}}{nh (1 - F(y_j | x))^2 \prod_{i=1}^n (B_{ni}(x))^2 \text{Var}^{Y|X} \mu_{i-1}^{Y|X}(y; x)} \\ &= A \frac{\prod_{i=1}^n (B_{ni}(x))^3 E^{Y|X} \mu_{i-1}^{Y|X}(y; x) E^{Y|X} \mu_{i-1}^{Y|X}(y; x)^{-3}}{\prod_{i=1}^n (B_{ni}(x))^2 \text{Var}^{Y|X} \mu_{i-1}^{Y|X}(y; x)^{3-2}} : \end{aligned} \tag{2.79}$$

A continuación veremos que $E^{Y|X} \mu_{i-1}^{Y|X}(y; x) E^{Y|X} \mu_{i-1}^{Y|X}(y; x)^{-3}$ está acotada de modo uniforme δ_i y como consecuencia, el numerador de la cota (2.79) tendrá orden $O(\prod_{i=1}^n (B_{ni}(x))^3)$ que es

$$\prod_{i=1}^n (B_{ni}(x))^3 \cdot \max_i B_{ni}(x) \prod_{i=1}^n B_{ni}(x) = O\left(\frac{1}{nh}\right) \quad \text{c.s.}$$

Estudiamos entonces:

$$\begin{aligned} & E^{Y|X} \mu_{i-1}^{Y|X}(y; x) E^{Y|X} \mu_{i-1}^{Y|X}(y; x)^{-3} = \\ &= E^{Y|X} \mu_{i-1}^{Y|X}(y; x)^3 + 3 \mu_{i-1}^{Y|X}(y; x) E^{Y|X} \mu_{i-1}^{Y|X}(y; x)^2 + \\ &+ 3 \mu_{i-1}^{Y|X}(y; x) E^{Y|X} \mu_{i-1}^{Y|X}(y; x)^2 + E^{Y|X} \mu_{i-1}^{Y|X}(y; x)^3 \\ &\cdot E^{Y|X} j_{i-1}^2(y; x) + 3 E^{Y|X} \mu_{i-1}^{Y|X}(y; x) E^{Y|X} (j_{i-1}(y; x) j) + \\ &+ 3 E^{Y|X} (j_{i-1}(y; x) j) E^{Y|X} (j_{i-1}(y; x) j) + E^{Y|X} (j_{i-1}(y; x) j)^3 \\ &\cdot E^{Y|X} j_{i-1}^3(y; x) + 3 E^{Y|X} \mu_{i-1}^{Y|X}(y; x) E^{Y|X} (j_{i-1}(y; x) j) + 4 E^{Y|X} (j_{i-1}(y; x) j)^3 \\ &\cdot 4 E^{Y|X} j_{i-1}^3(y; x) + E^{Y|X} (j_{i-1}(y; x) j)^3 \end{aligned}$$

(en la última desigualdad se ha utilizado que $j_{i-1}(y; x) j$ y $j_{i-1}(y; x) j^2$ tienen covarianza positiva y por tanto $E^{Y|X} j_{i-1}(y; x) j j_{i-1}(y; x) j^2 \geq E^{Y|X} (j_{i-1}(y; x) j) E^{Y|X} j_{i-1}(y; x) j^2$).

Como por otra parte, se obtiene de modo uniforme δ_i que:

$$|j_{\gg_i}(y; x)j| = \frac{\int_0^y \frac{1_{fT_i \cdot u \cdot Z_i g}}{(C(u; j; x))^2} dH_1^{\mu}(u=x)}{C(Z_i; j; x)} \cdot \left[\sup_{s2[a; y]} \frac{1}{C(s; j; x)} + \sup_{s2[a; y]} \frac{1}{(C(s; j; x))^2} \cdot 2 \sup_{s2[a; y]} \frac{1}{(C(s; j; x))^2} \cdot \frac{2}{\mu^2} \right]$$

se llega al resultado deseado de

$$\begin{aligned} & E^{YjX} |j_{\gg_i}(y; x)j| \leq E^{YjX} |j_{\gg_i}(y; x)j|^3 \\ & \leq 4 E^{YjX} |j_{\gg_i}(y; x)j|^3 + E^{YjX} (j_{\gg_i}(y; x)j)^3 \\ & \leq 4 \frac{\tilde{\mu}^3}{\mu^2} + \frac{\mu^3}{\mu^2} = 8 \frac{\mu^3}{\mu^2} \quad c:s: \end{aligned}$$

Continuamos estudiando el denominador de la cota (2.79) dado por

$$\tilde{\mu} \sum_{i=1}^3 (B_{ni}(x))^2 \text{Var}^{YjX} |j_{\gg_i}(y; x)j| \quad c:s:$$

Como consecuencia del Lema 2.8 (que se prueba a continuación) se obtiene

$$\sum_{i=1}^3 (B_{ni}(x))^2 \text{Var}^{YjX} |j_{\gg_i}(y; x)j| = \frac{s^2(y; j; x)}{(1 - F(y; j; x))^2 nh} + O\left(\frac{(\ln n)^{1=2}}{nh(nh)^{1=2}}\right) \quad c:s:$$

lo que, de forma inmediata, demuestra que

$$\tilde{\mu} \sum_{i=1}^3 (B_{ni}(x))^2 \text{Var}^{YjX} |j_{\gg_i}(y; x)j| = O\left(\frac{1}{(nh)^{3=2}}\right) \quad c:s:$$

y por tanto la cota en (2.79) tiene orden

$$O\left(\frac{(\ln n)^{1=2}}{(nh)^{3=2}}\right) = O\left(\frac{1}{(nh)^{1=2}}\right) \quad c:s:$$

Tomando $h = Cn^{1=5}$ se obtiene la tesis $O\left(\frac{1}{n^{2=5}}\right)$:

Lema 2.8 Bajo las hipótesis del Teorema 2.1c) y tomando $h = Cn^{-1/5}$ se verifica, para $x \in I$ e $y \in [a; b]$, que:

a)

$$b_n(y|x) = b(y|x) + O\left(\frac{(\ln n)^{1/2}}{n^{1/5}}\right) \quad \text{c.s.}$$

b)

$$v_n(y|x) = s^2(y|x) + O\left(\frac{(\ln n)^{1/2}}{n^{2/5}}\right) \quad \text{c.s.}$$

donde $b_n(y|x)$ y $v_n(y|x)$ están definidos en el enunciado del lema anterior,

$$b(y|x) = C^{5/2} (1 - F(y|x)) \int_{-\infty}^{\infty} \frac{z^2 K(z) dz}{2m^{\mu}(x)} \int_{-\infty}^{\infty} \Phi^{\mu}(x) m^{\mu}(x) + 2\Phi^0(x) m^{\mu_0}(x) dz$$

$$\begin{aligned} \Phi(u) &= E\left(\sum_{Z=y}^{\infty} \mathbb{1}_{\{Z; T; \pm; y; x\}} \cdot Z; X = u\right) = \\ &= \int_0^{\infty} \frac{dH_1^{\mu}(s|x)}{C(s|x)} \int_0^{\infty} \frac{C(s|u)}{(C(s|x))^2} dH_1^{\mu}(s|x) \end{aligned}$$

y

$$s^2(y|x) = (1 - F(y|x))^2 \int_{-\infty}^{\infty} \frac{K^2(z) dz}{m^{\mu}(x)} \int_0^{\infty} \frac{dH_1^{\mu}(u|x)}{(C(u|x))^2} dz$$

Demostración.

a) Recordemos que

$$\begin{aligned} b_n(y|x) &= (nh)^{1/2} (1 - F(y|x)) \sum_{i=1}^n B_{ni}(x) E^{Y|X}(\mathbb{1}_{\{Z_i; T_i; \pm_i; y; x\}}) = \\ &= (nh)^{1/2} (1 - F(y|x)) \sum_{i=1}^n B_{ni}(x) \Phi(X_i); \end{aligned}$$

Teniendo en cuenta que $B_{ni}(x) = \frac{1}{nh} K\left(\frac{x_i - x}{h}\right)$ y que

$$\sup_{x \in I} |m_n^{\mu}(x) - m^{\mu}(x)| = O\left(\frac{(\ln n)^{1/2}}{nh}\right) \quad \text{c.s.}$$

se obtiene

$$b_n(y|x) = \frac{1}{(nh)^{1-2}} \frac{1}{m^{\alpha}(x)} \sum_{i=1}^n \frac{1}{nh} K\left(\frac{x_i - x}{h}\right) \Phi(x_i) + O\left(\frac{(\ln n)^{1-2}}{nh}\right) \quad c:s:$$

que tomando $h = Cn^{-1/5}$ equivale a

$$b_n(y|x) - \tilde{b}_n(y|x) = O\left(\frac{(\ln n)^{1-2}}{n^{2/5}}\right) \quad c:s: \quad (2.80)$$

Para el análisis del término $\tilde{b}_n(y|x)$ utilizamos que $\Phi(x) = 0$ y efectuamos cálculos estándar en el análisis del sesgo de los estimadores de Nadaraya-Watson (similares al desarrollo del término (2.7)). Así, llegamos a

$$\begin{aligned} E \tilde{b}_n(y|x) &= \frac{1}{m^{\alpha}(x)} \sum_{i=1}^n \frac{1}{nh} K\left(\frac{x_i - x}{h}\right) (\Phi(x_i) - \Phi(x)) \\ &= \frac{1}{m^{\alpha}(x)h} \int_{-\infty}^{\infty} K\left(\frac{x - u}{h}\right) (\Phi(u) - \Phi(x)) m^{\alpha}(u) du \\ &= \frac{h^2}{2m^{\alpha}(x)} \int_{-\infty}^{\infty} z^2 K(z) dz \left[\Phi^{(0)}(x) m^{\alpha}(x) + 2\Phi^{(1)}(x) m^{\alpha}(x) \right] + O(h^3) \end{aligned}$$

(siempre que $h \neq 0$; lo cual nos permite trabajar en un entorno conveniente de $x \in I$). De esta forma, $E \tilde{b}_n(y|x)$ se puede escribir como

$$\frac{(nh)^{1-2} h^2}{2m^{\alpha}(x)} \frac{1}{(1 - F(y|x))} \int_{-\infty}^{\infty} z^2 K(z) dz \left[\Phi^{(0)}(x) m^{\alpha}(x) + 2\Phi^{(1)}(x) m^{\alpha}(x) \right] + O((nh)^{1-2} h^3)$$

y tomando $h = Cn^{-1/5}$ se tiene

$$E \tilde{b}_n(y|x) = b(y|x) + O(n^{-1/5}):$$

De este modo

$$\begin{aligned} \tilde{b}_n(y|x) - b(y|x) &= \tilde{b}_n(y|x) - E \tilde{b}_n(y|x) + E \tilde{b}_n(y|x) - b(y|x) = \\ &= \tilde{b}_n(y|x) - E \tilde{b}_n(y|x) + O(n^{-1/5}) \end{aligned} \quad (2.81)$$

Continuamos esta demostración probando que

$$\| \tilde{b}_n(y|x) - E \tilde{b}_n(y|x) \| = O \left(\frac{\tilde{A} \mu^{1/2}}{n^{2.5}} \right) = O \left(\frac{\tilde{A} (\ln n)^{1/2}}{n^{1.5}} \right) \quad \text{c.s.} \quad (2.82)$$

resultado que equivale a probar

$$\sum_{i=1}^n \| W_{ni} - E W_{ni} \| = O \left(\frac{\tilde{A} \mu^{1/2}}{n^{2.5}} \right) \quad \text{c.s.}$$

para las variables

$$W_{ni} = \frac{1}{m^{\mu}(x)} (nh)^{i-1/2} K \left(\frac{x_i - X_i}{h} \right) \Phi(X_i)$$

Para ello aplicaremos la desigualdad de Bernstein a las variables $W_{ni} - E W_{ni}$.

En primer lugar, recordando la expresión de $\Phi(u)$ (dada en el enunciado de este lema) y utilizando las hipótesis sobre acotación de las funciones involucradas en ella, junto con la de que el núcleo K es una función acotada (por ser de variación acotada) y tomando h suficientemente pequeño, podemos acotar la expresión:

$$\begin{aligned} |W_{ni} - E W_{ni}| &\leq \frac{1}{m^{\mu}(x)} (nh)^{i-1/2} \sup_{z \in \mathbb{R}} K(z) \frac{4}{\mu^2} = \\ &= O \left((nh)^{i-1/2} \right) = O \left(n^{i-2.5} \right); \end{aligned}$$

Por otro lado y teniendo en cuenta las mismas consideraciones efectuadas para la acotación anterior, se tiene

$$\begin{aligned} &\sum_{i=1}^n \text{Var}(W_{ni} - E W_{ni}) \\ &= \sum_{i=1}^n E W_{ni}^2 - \left(\sum_{i=1}^n E W_{ni} \right)^2 \\ &= \sum_{i=1}^n \frac{1}{m^{\mu}(x)} \frac{1}{h} \int_{\mathbb{R}} K^2 \left(\frac{x_i - u}{h} \right) (\Phi(u))^2 m^{\mu}(u) du \end{aligned}$$

donde, utilizando Taylor y que $\Phi(x) = 0$; la última integral es

$$\begin{aligned} & \frac{1}{h} \int_{-\infty}^{\infty} K^2\left(\frac{x_i - u}{h}\right) \Phi(u)^2 m^{\alpha}(u) du \\ &= \int_{-\infty}^{\infty} K^2(z) \Phi(x_i - hz)^2 m^{\alpha}(x_i - hz) dz \\ &= \int_{-\infty}^{\infty} K^2(z) \left[\Phi^0(x) + h^2 z^2 \frac{\Phi^{(2)}(x)}{2} + o(h^2) \right]^2 m^{\alpha}(x_i - hz) dz \\ &= \int_{-\infty}^{\infty} K^2(z) z^2 dz \left[\Phi^0(x)^2 m^{\alpha}(x) + o(h^2) \right] = O(h^2); \end{aligned}$$

De esta forma se llega a

$$\text{Var}(W_{ni}) = O(h^2) = O(n^{-2-5});$$

Por tanto, la desigualdad de Bernstein nos dice que

$$P\left(\sum_{i=1}^n (W_{ni} - EW_{ni}) > \epsilon\right) \leq 2 \exp\left\{-\frac{\epsilon^2}{2c_1 n^{1-5} + \frac{2}{3} c_2 \epsilon n^{1-5}}\right\} \leq 2 \exp\left\{-\frac{\epsilon^2}{2c_1 + \frac{2}{3} c_2 \epsilon}\right\} \leq 2 \exp\left\{-\frac{\epsilon^2}{c^0}\right\}$$

y para $\epsilon = c^{\alpha} \frac{\ln n}{n^{1-5}}$ queda

$$2 \exp\left\{-\frac{2n^{2-5}}{c^0}\right\} \cdot 2 \exp\left\{-\frac{c^{\alpha 2} \ln n}{c^0}\right\} = 2n^i \frac{c^{\alpha 2}}{c^0}$$

que es una serie sumable, sin más que elegir c^{α} de forma conveniente.

Aplicando el lema de Borel Cantelli, se obtiene:

$$\sum_{i=1}^{\infty} P\left(\sum_{j=1}^i (W_{nj} - EW_{nj}) > \frac{\ln n}{n^{1-5}}\right) = O\left(\frac{\ln n}{n^{1-5}}\right) \quad c:s:$$

Finalmente, como consecuencia de (2.80), (2.81) y (2.82) se llega al resultado deseado de

$$b_n(y | x) - b(y | x) = O\left(\frac{(\ln n)^{1-2}}{n^{1-5}}\right) \quad c:s:$$

b) Procedemos de modo análogo al del apartado a).

Recordamos la expresión

$$v_n(y|x) = nh(1 - F(y|x))^2 \sum_{i=1}^{\infty} (B_{ni}(x))^2 \text{Var}^{Y|X}(Z_i; T_i; \pm_i; y; x)$$

que podemos escribir, mediante cálculos sencillos y utilizando que

$$\sup_i |m_n^{\pm}(x) - m^{\pm}(x)| = O\left(\frac{\mu_{\ln n}^{(1)}(x)}{nh}\right) \quad \text{c.s.};$$

de la forma siguiente

$$v_n(y|x) = \tilde{v}_n(y|x) + O\left(\frac{\mu_{\ln n}^{(1)}(x)}{nh}\right) \quad \text{c.s.} \quad (2.83)$$

donde

$$\tilde{v}_n(y|x) = nh(1 - F(y|x))^2 \frac{1}{(nh)^2 m^{\pm 2}(x)} \mathbb{E} \sum_{i=1}^{\infty} K^2 \frac{\mu_{X_i} \mu_{X_i}}{h} \text{Var}^{Y|X}(Z_i; T_i; \pm_i; y; x)$$

A continuación estudiamos la esperanza de $\tilde{v}_n(y|x)$:

$$\begin{aligned} \mathbb{E}(\tilde{v}_n(y|x)) &= \frac{(1 - F(y|x))^2}{hm^{\pm 2}(x)} \mathbb{E} \sum_{i=1}^{\infty} K^2 \frac{\mu_{X_i}}{h} \int \text{Var}^{Y|X}(Z; T; \pm; y; x) \big|_{X=u; T=Z} m^{\pm}(u) du \\ &= \frac{(1 - F(y|x))^2}{m^{\pm 2}(x)} \sum_{i=1}^{\infty} K^2 \int (z) V(x; hz) m^{\pm}(x; hz) dz \end{aligned}$$

y aplicando Taylor a la última integral (en un entorno conveniente de $x \pm h$) se obtiene:

$$\begin{aligned} &\sum_{i=1}^{\infty} K^2 \int (z) V(x; hz) m^{\pm}(x; hz) dz \\ &= \sum_{i=1}^{\infty} K^2 \int \left[V(x; hz) m^{\pm}(x) + \frac{h^2 z^2}{2} V''(x) m^{\pm}(x) + \dots \right] dz \\ &= \sum_{i=1}^{\infty} K^2 \int V(x) m^{\pm}(x) dz + O(h^2) \end{aligned}$$

De esta forma

$$\begin{aligned} E(\tilde{v}_n(y_j|x)) &= \frac{(1 - F(y_j|x))^2}{m^{\alpha}(x)} \int K^2(z) dz V(x) + O(h^2) = \\ &= s^2(y_j|x) + O(h^2) = s^2(y_j|x) + O(n^{-2-5}): \end{aligned} \quad (2.84)$$

Continuamos esta demostración probando que

$$|j\tilde{v}_n(y_j|x) - E\tilde{v}_n(y_j|x)| = O\left(\frac{\tilde{A} \mu_{1=2}}{nh}\right) = O\left(\frac{\tilde{A} (\ln n)^{1=2}}{n^{2=5}}\right) \quad c:s: \quad (2.85)$$

Es claro que podemos escribir

$$|j\tilde{v}_n(y_j|x) - E\tilde{v}_n(y_j|x)| = \sum_{i=1}^n (W_{ni} - EW_{ni})$$

donde

$$W_{ni} = (1 - F(y_j|x))^2 \frac{1}{nhm^{\alpha}(x)} K^2\left(\frac{x_i - X_i}{h}\right) V(X_i)$$

Aplicaremos la desigualdad de Bernstein a las variables $W_{ni} - EW_{ni}$: Por tanto calculamos:

$$|jW_{ni} - EW_{ni}| \leq \frac{(1 - F(y_j|x))^2}{m^{\alpha}(x)} (nh)^{-1} \sup_{z \in (1;1)} K^2(z) \frac{4}{\mu^4} = O(nh)^{-1}$$

(donde la acotación para $V(X_i)$ se obtiene en un entorno de x ($h \rightarrow 0$), chequeando su expresión a lo largo de la demostración del Teorema 2.3).

También calculamos

$$\begin{aligned} & \sum_{i=1}^n \text{Var}(W_{ni} - EW_{ni}) \\ &= \sum_{i=1}^n E(W_{ni}^2) - \sum_{i=1}^n (EW_{ni})^2 \\ &= \sum_{i=1}^n \frac{(1 - F(y_j|x))^4}{m^{\alpha}(x)} \int K^4\left(\frac{x_i - u}{h}\right) V(u)^2 m^{\alpha}(u) du \end{aligned}$$

donde, utilizando Taylor, la última integral da

$$\begin{aligned} & \frac{1}{h} \int_{Z^*}^Z K^4 \left(\frac{x_i - u}{h} \right) (V(u))^2 m^{\alpha}(u) du \\ &= \int_{Z^*}^Z K^4(z) (V(x_i - hz))^2 m^{\alpha}(x_i - hz) dz \\ &= \int_{Z^*}^Z K^4(z) \left[V(x_i) - hzV'(x_i) + h^2 z^2 \frac{V''(x_i)}{2} \right]^2 m^{\alpha}(x_i - hz) dz \\ &= \int_{Z^*}^Z K^4(z) \left[m^{\alpha}(x_i) - hzm^{\alpha'}(x_i) + h^2 z^2 \frac{m^{\alpha''}(x_i)}{2} \right] dz \\ &= V(x_i) m^{\alpha}(x_i) \int_{Z^*}^Z K^4(z) dz + O(h^2). \end{aligned}$$

De esta forma se llega a

$$\begin{aligned} \sum_{i=1}^n \text{Var}(W_{ni}) &= \frac{\mu_{1i} F(y_j|x)}{m^{\alpha}(x)} \frac{1}{nh} \int_{Z^*}^Z K^4(z) dz + O(h^2) \\ &= O(nh)^{-1}. \end{aligned}$$

Por tanto, la desigualdad de Bernstein nos dice que

$$\begin{aligned} P \left(\sum_{i=1}^n (W_{ni} - EW_{ni}) > \epsilon \right) &\leq 2 \exp \left\{ - \frac{\epsilon^2}{2c_1 (nh)^{-1} + \frac{2}{3} c_2 \epsilon (nh)^{-1}} \right\} \\ &\leq 2 \exp \left\{ - \frac{2nh}{2c_1 \left(1 + \frac{2}{3} \epsilon\right)} \right\}. \end{aligned}$$

Tomando $\epsilon = c^{\alpha} \frac{\ln n}{nh}$ queda

$$2 \exp \left\{ - \frac{2nh}{2c_1 \left(1 + \frac{2}{3} c^{\alpha} \frac{\ln n}{nh}\right)} \right\} \leq 2 \exp \left\{ - \frac{c^{\alpha 2} \ln n}{c^0} \right\} = 2n^{-\frac{c^{\alpha 2}}{c^0}}$$

que es una serie sumable, sin más que elegir c^{α} de forma conveniente.

Aplicando de nuevo el lema de Borel Cantelli, se obtiene el resultado:

$$\sum_{i=1}^n (W_{ni} - EW_{ni}) = O \left(\frac{\ln n}{nh} \right) = O \left(\frac{(\ln n)^{1=2}}{n^{2=5}} \right) \quad \text{c.s.}$$

Finalmente tenemos

$$\begin{aligned} v_n(y_j | x) - s^2(y_j | x) &= [v_n(y_j | x) - \tilde{v}_n(y_j | x)] + \\ &+ [\tilde{v}_n(y_j | x) - E\tilde{v}_n(y_j | x)] + \\ &+ \frac{E\tilde{v}_n(y_j | x) - s^2(y_j | x)}{A} \\ &= O\left(\frac{(\ln n)^{1=2}}{n^{2=5}}\right) \quad \text{c:s:} \end{aligned}$$

sin más que utilizar (2.83), (2.84) y (2.85).

6. Demostración del Teorema 2.5.

Empezamos demostrando el apartado a) del teorema.

Aplicando la desigualdad triangular podemos acotar el supremo del enunciado por:

$$\begin{aligned} &\sup_{t \in \mathbb{R}} \left| \sum_{i=1}^n \frac{t_i b_n(y_j | x)}{(v_n(y_j | x))^{1=2}} - \sum_{i=1}^n \frac{t_i b(y_j | x)}{s(y_j | x)} \right| \\ &\leq \sup_{t \in \mathbb{R}} \sum_{i=1}^n \frac{t_i |b_n(y_j | x) - b(y_j | x)|}{(v_n(y_j | x))^{1=2}} + \sup_{t \in \mathbb{R}} \sum_{i=1}^n \frac{t_i |b(y_j | x)|}{s(y_j | x)} \end{aligned}$$

El orden del primer supremo lo da el Lema 2.7; es una $O(n^{1-2=5})$ c:s:

Para estudiar el orden del segundo supremo utilizaremos los Lemas 2.6 y 2.8.

De este modo, llegamos a

$$\begin{aligned} &\sup_{t \in \mathbb{R}} \sum_{i=1}^n \frac{t_i |b_n(y_j | x) - b(y_j | x)|}{(v_n(y_j | x))^{1=2}} + \sup_{t \in \mathbb{R}} \sum_{i=1}^n \frac{t_i |b(y_j | x)|}{s(y_j | x)} \\ &\leq O\left(\frac{(\ln n)^{1=2}}{n^{1=5}}\right) + O\left(\frac{(\ln n)^{1=2}}{n^{2=5}}\right) \quad \text{c:s:} \\ &= O\left(\frac{(\ln n)^{1=2}}{n^{1=5}}\right) \quad \text{c:s:} \end{aligned}$$

Como consecuencia de los órdenes de ambos supremos se obtiene el resultado deseado.

Para probar el apartado b), comenzamos demostrando el siguiente resultado:

$$\sup_{t \in \mathbb{R}} \mathbb{P} \left[\sum_{i=1}^n (nh)^{1/2} (1 - F(y_j | x)) B_{ni}(x) \sum_{j=1}^n t_j \frac{b(y_j | x)}{s(y_j | x)} \right] \leq \Phi_N \left(\frac{\sum_{j=1}^n t_j b(y_j | x)}{s(y_j | x)} \right) \rightarrow 0 \quad (2.86)$$

Aplicando la desigualdad triangular, podemos acotar el supremo anterior por los dos sumandos siguientes:

$$\sup_{t \in \mathbb{R}} \mathbb{P} \left[\sum_{i=1}^n (nh)^{1/2} (1 - F(y_j | x)) B_{ni}(x) \sum_{j=1}^n t_j \frac{b(y_j | x)}{(v_n(y_j | x))^{1/2}} \right] + \sup_{t \in \mathbb{R}} \mathbb{P} \left[\sum_{i=1}^n (nh)^{1/2} (1 - F(y_j | x)) B_{ni}(x) \sum_{j=1}^n t_j \frac{b(y_j | x)}{s(y_j | x)} \right] \quad (2.87)$$

$$\sup_{t \in \mathbb{R}} \mathbb{P} \left[\sum_{i=1}^n (nh)^{1/2} (1 - F(y_j | x)) B_{ni}(x) \sum_{j=1}^n t_j \frac{b(y_j | x)}{(v_n(y_j | x))^{1/2}} \right] \leq \Phi_N \left(\frac{\sum_{j=1}^n t_j b(y_j | x)}{s(y_j | x)} \right) \rightarrow 0 \quad (2.88)$$

Para analizar el primer sumando, tenemos en cuenta el Lema 2.7 y así se tiene

$$\sup_{t \in \mathbb{R}} \mathbb{P}^{Y|X} \left[\sum_{i=1}^n (nh)^{1/2} (1 - F(y_j | x)) B_{ni}(x) \sum_{j=1}^n t_j \frac{b(y_j | x)}{(v_n(y_j | x))^{1/2}} \right] \leq \Phi_N \left(\frac{\sum_{j=1}^n t_j b(y_j | x)}{(v_n(y_j | x))^{1/2}} \right) = o(1) \quad \text{c.s.}$$

Se trata de una sucesión de v.a. acotadas (funciones de distribución) y por tanto la sucesión de sus esperanzas converge también a 0, es decir:

$$\mathbb{E} \left[\sum_{i=1}^n (nh)^{1/2} (1 - F(y_j | x)) B_{ni}(x) \sum_{j=1}^n t_j \frac{b(y_j | x)}{(v_n(y_j | x))^{1/2}} \right] \leq \Phi_N \left(\frac{\sum_{j=1}^n t_j b(y_j | x)}{(v_n(y_j | x))^{1/2}} \right) \rightarrow 0$$

Por otra parte se cumple, para cualquier función aleatoria g que

$$\sup_{z \in \mathbb{R}} |E(g(z))| \cdot \sup_{z \in \mathbb{R}} E |g(z)| \cdot E \sup_{z \in \mathbb{R}} |g(z)| \quad (2.89)$$

y aplicando conjuntamente los dos últimos resultados se llega a

$$\begin{aligned} & \sup_{t \in \mathbb{R}} |E \Phi_N^{-1} \frac{t_i b_n(y_j x)}{(v_n(y_j x))^{1/2}}| \\ & \cdot E \sup_{t \in \mathbb{R}} |P^{Y_j X} (nh)^{1/2} (1 - F(y_j x)) \prod_{i=1}^X B_{ni}(x) \mathbb{1}_{\{Z_i \leq T_i; \pm_i \leq y; x\}} \cdot t_i| \\ & \leq O(b_n(y_j x) - b(y_j x)) + O((v_n(y_j x))^{1/2} - s(y_j x)) = o(1) \quad \text{c.s.} \end{aligned}$$

y esto implica que el primer supremo en estudio, dado por (2.87) tiende a cero, es decir

$$(2.87) \rightarrow 0 \quad (2.90)$$

Para estudiar el comportamiento del segundo sumando dado por (2.88) utilizaremos el Lema 2.6 y la convergencia c.s. de $b_n(y_j x)$ a $b(y_j x)$ y de $v_n(y_j x)$ a $s^2(y_j x)$ demostrada en el Lema 2.8. Así obtenemos que

$$\begin{aligned} & \sup_{t \in \mathbb{R}} |E \Phi_N^{-1} \frac{t_i b_n(y_j x)}{(v_n(y_j x))^{1/2}} - \mu \frac{t_i b(y_j x)}{s(y_j x)}| \\ & \leq O(b_n(y_j x) - b(y_j x)) + O((v_n(y_j x))^{1/2} - s(y_j x)) = o(1) \quad \text{c.s.} \end{aligned}$$

Tenemos nuevamente una sucesión de v.a. acotadas (funciones de distribución) que convergen a 0 y por tanto la sucesión de sus esperanzas converge también a 0. Esta propiedad conjuntamente con (2.89), permite obtener la convergencia a 0 de (2.88):

$$\begin{aligned} & \sup_{t \in \mathbb{R}} |E \Phi_N^{-1} \frac{t_i b_n(y_j x)}{(v_n(y_j x))^{1/2}} - \mu \frac{t_i b(y_j x)}{s(y_j x)}| \\ & \cdot E \sup_{t \in \mathbb{R}} |P^{Y_j X} \frac{t_i b_n(y_j x)}{(v_n(y_j x))^{1/2}} - \mu \frac{t_i b(y_j x)}{s(y_j x)}| = o(1) \quad (2.91) \end{aligned}$$

Finalmente, (2.90) y (2.91) prueban el resultado (2.86).

Estudiamos ahora el término:

$$\sup_{t \in \mathbb{R}} P \left[\frac{h}{(nh)^{1/2}} \sum_{i=1}^n \hat{F}_h(y_j | x) - F(y_j | x) \cdot t \right] \Phi_N \left[\frac{\mu_{t_i} b(y_j | x)}{s(y_j | x)} \right]$$

Teniendo en cuenta la representación casi segura dada en el Teorema 2.1c), se tiene

$$\begin{aligned} & P \left[\frac{h}{(nh)^{1/2}} \sum_{i=1}^n \hat{F}_h(y_j | x) - F(y_j | x) \cdot t \right] \\ &= P \left[\frac{h}{(nh)^{1/2}} (1 - F(y_j | x)) \sum_{i=1}^n B_{ni}(x) \mathbb{1}_{\{Z_i \leq T_i + t_i; y_i \leq x\}} \right. \\ & \quad \left. + (nh)^{1/2} R_n^0(y_j | x) \cdot t \right] \end{aligned}$$

y es fácil demostrar que para cualesquiera $U; V$ v.a. se veri...ca

$$P(U + V \cdot t) \cdot P(U \cdot t + \dots) + P(jVj, \dots) \cdot 8^n > 0:$$

Utilizando estos dos últimos resultados, y considerando $\{n\}$ como una sucesión de números reales convergente a 0, se tiene que

$$\begin{aligned} & P \left[\frac{h}{(nh)^{1/2}} \sum_{i=1}^n \hat{F}_h(y_j | x) - F(y_j | x) \cdot t \right] \Phi_N \left[\frac{\mu_{t_i} b(y_j | x)}{s(y_j | x)} \right] \\ & \cdot P \left[\frac{h}{(nh)^{1/2}} (1 - F(y_j | x)) \sum_{i=1}^n B_{ni}(x) \mathbb{1}_{\{Z_i \leq T_i + t_i; y_i \leq x\}} \cdot t + \dots \right. \\ & \quad \left. + P \left[\frac{h}{(nh)^{1/2}} R_n^0(y_j | x) \cdot t \right] \Phi_N \left[\frac{\mu_{t_i} b(y_j | x)}{s(y_j | x)} \right] \right] \end{aligned}$$

y por tanto

$$\begin{aligned} & \sup_{t \in \mathbb{R}} P \left[\frac{h}{(nh)^{1/2}} \sum_{i=1}^n \hat{F}_h(y_j | x) - F(y_j | x) \cdot t \right] \Phi_N \left[\frac{\mu_{t_i} b(y_j | x)}{s(y_j | x)} \right] \\ & \cdot \sup_{t \in \mathbb{R}} P \left[\frac{h}{(nh)^{1/2}} (1 - F(y_j | x)) \sum_{i=1}^n B_{ni}(x) \mathbb{1}_{\{Z_i \leq T_i + t_i; y_i \leq x\}} \cdot t + \dots \right. \\ & \quad \left. + \Phi_N \left[\frac{\mu_{t + \dots} b(y_j | x)}{s(y_j | x)} \right] \right] \\ & + \sup_{t \in \mathbb{R}} \Phi_N \left[\frac{\mu_{t + \dots} b(y_j | x)}{s(y_j | x)} \right] \Phi_N \left[\frac{\mu_{t_i} b(y_j | x)}{s(y_j | x)} \right] \\ & + P \left[\frac{h}{(nh)^{1/2}} R_n^0(y_j | x) \cdot t \right] \end{aligned}$$

El primer supremo del segundo miembro de la desigualdad anterior es una $o(1)$ por (2.86).

El segundo sumando se trata aplicando el Lema 2.6. De este modo, se obtiene:

$$\sup_{t \in \mathbb{R}} \left| \int_{\Phi_N} \frac{\mu_{t + \frac{b(y_j | x)}{s(y_j | x)}}}{s(y_j | x)} \right| \leq \int_{\Phi_N} \frac{\mu_{t + \frac{b(y_j | x)}{s(y_j | x)}}}{s(y_j | x)} \leq k \tilde{A}_N k_1 s(y_j | x)^{1-\alpha_n}$$

que tiende a 0 puesto que lo hace la sucesión α_n .

Para analizar el tercer sumando tendremos en cuenta que

$$(nh)^{1-2} \sup_{[a; b] \in I} \left| R_n^0(y_j | x) \right| = O \left(\frac{\tilde{A} (\ln n)^{3-4}}{(nh)^{1-4}} \right) = o(1) \quad \text{c.s.}$$

de donde se concluye que $\delta_n > 0$:

$$P \left((nh)^{1-2} \sup_{[a; b] \in I} \left| R_n^0(y_j | x) \right| \leq \delta_n \right) \rightarrow 1 \quad \text{c.s.}$$

Entonces, en $[a; b] \in I$ se verifica:

$$P \left((nh)^{1-2} \left| R_n^0(y_j | x) \right| \leq \delta_n \right) \cdot P \left((nh)^{1-2} \sup_{[a; b] \in I} \left| R_n^0(y_j | x) \right| \leq \delta_n \right) \rightarrow 1 \quad \text{c.s.}$$

Finalmente, la convergencia a 0 de cada uno de los tres sumandos prueba este segundo apartado del teorema.

Capítulo 3

Bootstrap para la función de distribución condicional en presencia de censura y truncamiento.

3.1 Antecedentes.

3.1.1 El método de remuestreo Bootstrap.

Este método de remuestreo, introducido por Efron en 1979, permite aproximar la distribución de un estadístico, función de una muestra aleatoria simple, mediante la distribución en el remuestreo (sobre nuevas muestras obtenidas a partir de algún estimador de la función de distribución poblacional desconocida) de un nuevo estimador, obtenido al reemplazar la muestra original por una nueva muestra denominada muestra bootstrap.

Concretamente, si denotamos por $R(\mathbf{X}; F)$ al estadístico dependiente de la muestra $\mathbf{X} = (X_1; X_2; \dots; X_n)$ y de la distribución poblacional desconocida F , el método bootstrap aproxima la distribución de dicho estadístico siguiendo el siguiente procedimiento:

1. Construir un estimador \hat{F}_n de la función de distribución F , mediante la distribución empírica.
2. Generar a partir de \hat{F}_n una nueva muestra $\mathbf{X}^a = (X_1^a; X_2^a; \dots; X_n^a)$ que se denominará muestra bootstrap (mecanismo que equivale a un muestreo aleatorio simple con reemplazamiento de la muestra original).

3. Aproximar la distribución de $R(\mathbf{X}; F)$ por la distribución de $R^n = R(\mathbf{X}^n; \hat{F}_n)$ en el remuestreo (que en la mayor parte de los casos prácticos se aproximará, a su vez, mediante el método de Montecarlo).

Este procedimiento bootstrap denominado standard o naive puede modificarse en sus dos primeros apartados para adaptarse de forma conveniente al contexto específico en el que uno se encuentre, utilizando de ese modo la información de que se disponga. Esto ocurre por ejemplo, en la estimación de la función de distribución cuando hay datos censurados, como se puede ver en Efron (1981), Akritas (1986) y Cao Abad (1988), o cuando hay datos truncados (consultar los trabajos de Wang (1991), Bilker y Wang (1997) y de Gross y Lai (1996)).

También se han definido distintas versiones del mecanismo bootstrap para la estimación de la función de regresión. Entre ellas destacan el wild bootstrap, estudiado en Härdle y Mammen (1993), Härdle y Marron (1991) y Mammen (1992), y el bootstrap suavizado presentado en Cao Abad (1990) y Cao Abad y González Manteiga (1993).

Es tarea del estadístico la elección o el diseño del plan de remuestreo que se adapte a cada situación, así como la comprobación de que la distribución de R^n aproxime convenientemente la distribución de R : En este sentido, un manual interesante donde se pueden consultar distintas formas de aproximación es Shao y Tu (1995).

3.1.2 El método Bootstrap para datos con censura.

Efron (1981) introduce dos métodos de remuestreo para datos censurados, conocidos en la literatura del bootstrap con censura como Bootstrap "simple" y Bootstrap "obvio".

Ambos parten de la muestra observada: $f(Z_i; \pm_i)g_{i=1}^n$ donde $Z_i = \min\{Y_i, C_i\}$ y $\pm_i = 1_{\{Y_i < C_i\}}$; $i = 1; 2; \dots; n$:

El método simple utilizado para aproximar la distribución de un estadístico $R = R(\mathbf{Z}; \pm)$ consiste en:

1. Generar la muestra bootstrap $f(\mathbf{Z}_i^n; \pm_i^n)g_{i=1}^n$ utilizando muestreo con reemplazamiento y masa $1/n$ en cada vector observado $(Z_i; \pm_i)$:
2. Obtener el estadístico de interés con la muestra bootstrap, esto es $R^n = R(\mathbf{Z}^n; \pm^n)$:
3. Aproximar la distribución de R por la de R^n en el remuestreo.

En cuanto al método obvio, construye la muestra bootstrap de acuerdo al siguiente procedimiento:

1. Calcular los estimadores límite producto de Y y de C , F_n^c y G_n^c , respectivamente (ver capítulo 1).
2. Generar muestras, de forma independiente: $Y_i^* \gg F_n^c$; $C_i^* \gg G_n^c$; para $i = 1; 2; \dots; n$:
 De...nir la muestra bootstrap $f(Z_i^*; \pm_i^*)g_{i=1}^n$ como $Z_i^* = \min(Y_i^*; C_i^*)$ y $\pm_i^* = 1_{fY_i^* \cdot C_i^*}g$:

A partir de este punto, los pasos para aproximar la distribución del estadístico $R = R(Z; \pm)$ son similares a los pasos 2 y 3 del bootstrap simple.

El mismo Efron (1981) prueba la equivalencia de estos dos métodos de remuestreo, y Akritas (1986) demuestra su validez, en el sentido de que $\hat{P}_n(F_n^{c*} | F_n^c)$ y $\hat{P}_n(F_n^c | F)$ convergen al mismo proceso límite en $[0; T]$, donde $T < b_H$ y F_n^{c*} denota el estimador límite producto de F construido con la muestra bootstrap (o equivalentemente, la versión bootstrap del estimador de Kaplan-Meier).

En el contexto de datos censurados con covariables, Van Keilegom y Veraverbeke (1997a) de...nen un procedimiento bootstrap para aproximar la distribución de $(nh)^{1-2} \hat{F}_n^c(y | j x) | F(y | j x)$, donde $\hat{F}_n^c(y | j x)$ es el estimador límite producto condicional de...nido en (2.13) con pesos de Gasser-Müller y bajo diseño ...jo.

El método, de tipo obvio condicional, consiste en:

1. Construir los estimadores límite producto condicionales de Y y de C , con una ventana piloto g (que tradicionalmente tiende a 0 de forma más lenta que h): $\hat{F}_g^c(t | j x_i)$ y $\hat{G}_g^c(t | j x_i)$, respectivamente.
2. Para cada x_i , con i desde 1 hasta n , obtener de forma independiente: Y_i^* de $\hat{F}_g^c(y | j x_i)$ y C_i^* de $\hat{G}_g^c(y | j x_i)$.
 De...nir $Z_i^* = \min(Y_i^*; C_i^*)$ y $\pm_i^* = 1_{fY_i^* \cdot C_i^*}g$ y así obtener la muestra bootstrap: $(x_1; Z_1^*; \pm_1^*); \dots; (x_n; Z_n^*; \pm_n^*)$.
3. Utilizando la muestra bootstrap de...nir el análogo bootstrap de $\hat{F}_n^c(y | j x)$, dado por:

$$\hat{F}_n^{c*}(y | j x) = \frac{1}{n} \sum_{i=1}^n \frac{1_{fZ_i^* \cdot y g \pm_i^*} B_{hi}(x)}{1_{fZ_i^* < Z_i^* g} B_{hj}(x)}$$

donde $B_{hi}(x)$ denota el peso de Gasser-Müller en x_i con ventana h .

4. Aproximar la distribución de $(nh)^{1-2}[\hat{F}_h^c(y|x) - F(y|x)]$ por la distribución en el remuestreo de $(nh)^{1-2}[\hat{F}_h^{c*}(y|x) - \hat{F}_g^c(y|x)]$.

Además, Van Keilegom y Veraverbeke (1997a) en su Teorema 5.1 prueban la consistencia uniforme fuerte de este método de remuestreo, es decir, que con probabilidad 1 se tiene la convergencia hacia 0 de

$$\sup_{t \in \mathbb{R}} \int_{\mathbb{R}} |(\hat{F}_h^{c*}(y|x) - \hat{F}_g^c(y|x)) - (F(y|x) - \hat{F}_h^c(y|x))| dt \rightarrow 0$$

En la demostración utilizan la herramienta de las representaciones casi seguras.

En la sección siguiente, nosotros abordaremos un estudio similar para el contexto de censura y truncamiento con covariables bajo diseño aleatorio.

3.1.3 El método Bootstrap para datos con censura y truncamiento.

En este apartado de la memoria presentamos los dos métodos de remuestreo bootstrap que trata la literatura para datos censurados y truncados (sin covariables), y que son una adaptación a este contexto de los métodos simple y obvio de Efron para datos censurados.

Ambos parten de una muestra observada: $f(T_i; Z_i; \delta_i)g_{i=1}^n$ donde $T_i \in [0, \infty)$; $Z_i \in \mathbb{R}^p$; $\delta_i = 1, 2, \dots, n$; $\delta_i = 1$ si $T_i \leq C_i$ y $\delta_i = 2$ si $T_i > C_i$.

El método simple genera la muestra bootstrap $(T_i^*, Z_i^*; \delta_i^*)$; $i = 1, \dots, n$; utilizando muestreo con reemplazamiento y masa $1/n$ en cada vector observado $(T_i; Z_i; \delta_i)$.

El método obvio que exponemos a continuación, está definido en Wang (1991) y Bilker y Wang (1997) para el siguiente modelo de censura y truncamiento:

Se supone que Y y $(T; C)$ son variables independientes y continuas, $C > T$; $D = C - T$ (con función de distribución Q) es independiente de T ; y también

$$b_L = \inf_{s \in \mathbb{R}} P(T \leq s \leq C) > 0 \text{ and } a_F = \sup_{s \in \mathbb{R}} P(T \leq s \leq C) < 1$$

Bajo estas hipótesis el estimador límite producto de L es igual a

$$L_n^{tc}(t) = \prod_{k=1}^{\infty} \frac{1 - F_n^{tc}(T_k)}{1 - F_n^{tc}(T_{k-1})} \prod_{i=1}^{\infty} \frac{1 - F_n^{tc}(T_i)}{1 - F_n^{tc}(T_{i-1})} 1_{F(T_i) \leq t}$$

y el de Q es el estimador límite producto para censura construido con Z_i y T_i como datos principales y $1_{\delta_i = 1}$ como variable indicadora, Q_n^c .

En este modelo, el método obvio para construir la muestra bootstrap, consiste en:

1. Obtener los estimadores de $F; L$ y Q : $F_n^{tc}, L_n^{tc}(t)$; y Q_n^c ; respectivamente.
2. De forma independiente, generar muestras $Y_i^a \gg F_n^{tc}; T_i^a \gg L_n^{tc}$; y $D_i^a \gg Q_n^c$; $i = 1; 2; \dots; n$:
 Definir $Z_i^a = \min(Y_i^a; C_i^a)$; donde $C_i^a = T_i^a + D_i^a$; y $\delta_i^a = 1_{F_{Y_i^a}, C_i^a}$; Si $T_i^a \cdot Y_i^a$; entonces mantener la observación, en caso contrario descartarla debido al truncamiento y generarla de nuevo. Continuar hasta obtener una muestra (bootstrap) $(T_1^a; Z_1^a; \delta_1^a); \dots; (T_n^a; Z_n^a; \delta_n^a)$ de tamaño n .

Respecto a estos dos métodos conviene destacar que:

- 2 Los dos métodos bootstrap presentados no son equivalentes para datos truncados y por tanto, no son equivalentes para datos bajo censura y truncamiento.
- 2 Bilker y Wang (1997) establecen la validez del método obvio probando que los procesos $n^{1/2} \{ S_n^{tc}(\cdot) - S_n^{tc}(\cdot) \}$ y $n^{1/2} \{ S_n^{tc}(\cdot) - S(\cdot) \}$ tienen la misma estructura asintótica (ellos trabajan con la función de supervivencia S ; en lugar de con la función de distribución $F = 1 - S$; y con sus correspondientes estimadores bootstrap y no bootstrap: S_n^{tc} y S_n^{tc} , respectivamente).
- 2 Gross y Lai (1996) desarrollan una teoría asintótica del método simple para datos censurados y truncados con covariables (que comentaremos a continuación) y, dicha teoría prueba, como caso particular, la convergencia en probabilidad hacia 0 de:

$$\frac{1}{\sqrt{n}} \left[\frac{S_n^a(y) - S_n(y)}{\hat{\sigma}_n} \right] \xrightarrow{d} N \left(0, \frac{S_n(y) - S(y)}{\sigma^2} \right)$$

donde $\hat{\sigma}_n$ es un estimador del error estandar de S verificando ciertas características.

Los métodos anteriores corresponden a un contexto de ausencia de covariables. Si consideramos la presencia de covariables, Gross y Lai (1996) definen un método bootstrap simple y no condicional del siguiente modo:

Sea $f(X_i; T_i; Z_i; \delta_i)_{i=1}^n$ la muestra observada donde $T_i \cdot Z_i; Z_i = \min f Y_i; C_i$ y $\delta_i = 1_{f Y_i, C_i}$; $i = 1; 2; \dots; n$:
 Sea Ψ_n la función de distribución empírica que asigna probabilidad $1/n$ a cada vector aleatorio $V_i = (X_i; T_i; Z_i; \delta_i)$; $i = 1; \dots; n$:

El método simple consiste en $V_1^a; \dots; V_n^a$ vectores i.i.d. con distribución común Ψ_n :

De este modo, si $R = R(V_1; \dots; V_n)$ es un estimador de $R(\Psi)$ y $\mathcal{R} = \mathcal{R}(V_1; \dots; V_n)$ es un estimador del error estándar de R ; el método de Gross y Lai estima la distribución $P[(R - R(\Psi)) / \mathcal{R} \leq z]$ por $P[(R^a - R(\Psi_n)) / \mathcal{R}^a \leq z | V_1; \dots; V_n]$ donde $R^a = R(V_1^a; \dots; V_n^a)$ y $\mathcal{R}^a = \mathcal{R}(V_1^a; \dots; V_n^a)$:

En muchas aplicaciones $(R - R(\Psi)) / \mathcal{R}$ puede ser expresada como un U-estadístico que asintóticamente tiene una distribución $N(0, 1)$: Lo mismo ocurre con el estadístico $(R^a - R(\Psi_n)) / \mathcal{R}^a$ y de ello se sigue la convergencia en probabilidad a cero de

$$P^a \left[\frac{R^a - R(\Psi_n)}{\mathcal{R}^a} \leq z \right] \rightarrow P \left[\frac{R - R(\Psi)}{\mathcal{R}} \leq z \right]$$

Además, bajo ciertas condiciones, el orden de convergencia es una $O_p(n^{-1/2})$:

3.2 Bootstrap condicional con censura y truncamiento.

3.2.1 Definición de una aproximación bootstrap.

En esta sección definimos un procedimiento de remuestreo bootstrap para aproximar la distribución de $(nh)^{1/2}[\hat{F}_n^{tc}(y | j, x) - F(y | j, x)]$; que llamaremos método bootstrap obvio condicional para datos truncados por la izquierda y censurados por la derecha. De esta forma tendremos una alternativa a la distribución asintótica normal derivada en el Teorema 2.3, cuyos parámetros dependen de cantidades teóricas de estimación complicada.

El proceso de remuestreo tiene en cuenta que, suponiendo T, Y, C condicionalmente independientes a $X = x$, y $a_{L(j|x)} < a_{H(j|x)}$ y $b_{L(j|x)} < b_{H(j|x)}$; se pueden calcular los estimadores de las funciones de distribución condicionales $F(y | j, x)$, $G(y | j, x)$ y $L(y | j, x)$ dados por $\hat{F}_n^{tc}(y | j, x)$, $\hat{G}_n^{tc}(y | j, x)$ y $\hat{L}_n^{tc}(y | j, x)$ respectivamente, de acuerdo con las definiciones (2.24), (2.29) y (2.28).

Observaciones. Antes de describir el método de muestreo, conviene aclarar que, por la naturaleza del mismo, utilizaremos dos ventanas distintas: h o g , según convenga. Por ello, a lo largo de este capítulo los estimadores límite producto condicionales: $\hat{F}_n^{tc}(y | j, x)$, $\hat{G}_n^{tc}(y | j, x)$ y $\hat{L}_n^{tc}(y | j, x)$ los denotaremos explicitando, en lugar de n , la ventana utilizada en cada caso, esto es: $\hat{F}_h^{tc}(y | j, x)$, $\hat{G}_h^{tc}(y | j, x)$ y $\hat{L}_h^{tc}(y | j, x)$ cuando se construyan utilizando pesos no paramétricos con ventana h ; y $\hat{F}_g^{tc}(y | j, x)$, $\hat{G}_g^{tc}(y | j, x)$ y $\hat{L}_g^{tc}(y | j, x)$ cuando se construyan utilizando pesos no paramétricos con ventana g : Lo mismo cabe decir para los estimadores no paramétricos de las

funciones $H_1^\alpha(y \mid x)$, $C(y \mid x)$; $\hat{H}_{1n}^\alpha(y \mid x)$; $\hat{C}_n(y \mid x)$ y para los pesos de Nadaraya-Watson $B_{ni}(x)$; sustituiremos el subíndice n por la ventana h o g que estemos utilizando.

En cuanto al capítulo anterior, todos los resultados han sido calculados para la ventana h . Es por esto que todos los estimadores y pesos no paramétricos del capítulo anterior (de subíndice n) se corresponden, con la notación de este capítulo, con ellos mismos denotados con subíndice h .

Otro punto que queremos aclarar es el relativo al uso del símbolo α como superíndice. Este símbolo tiene un significado propio para el bootstrap, significado que se mantiene a lo largo de este capítulo cuando acompaña a variables aleatorias (las réplicas bootstrap) y a probabilidades, esperanzas y varianzas de ciertas variables dependientes condicionadas a la muestra observada inicial. Sin embargo, el símbolo α como superíndice de una función, ya utilizado en capítulos anteriores en la definición de diversas funciones (H_1^α , H^α , L^α , m^α , ...) asociadas a los modelos con presencia de truncamiento, así como en la de sus respectivos estimadores (\hat{H}_{1n}^α , \hat{H}_n^α , \hat{L}_n^α , \hat{m}_n^α, \dots) cuyos análogos también aparecen en este capítulo, conserva el significado de condicionado a la ausencia de truncamiento. Ante la necesidad de denotar en el presente capítulo los análogos bootstrap de algunas de dichas funciones hemos optado por denotarlas mediante el superíndice Boot. De este modo, el símbolo α como superíndice de una variable signifi...ca que es una réplica bootstrap, y como superíndice de una función indica cierta probabilidad condicionada a la ausencia de truncamiento. En cuanto a las funciones que se definen en el universo bootstrap se denotan con el superíndice Boot.

Teniendo en cuenta las consideraciones efectuadas, y partiendo de una muestra observada $f(X_i; T_i; Z_i; \pm_i)g_{i=1}^n$; llevaremos a cabo el siguiente plan de remuestreo:

1. Para cada X_i , independientemente, obtener:

$$Y_i^\alpha \text{ de } \hat{F}_g^{\text{tc}}(y \mid X_i), C_i^\alpha \text{ de } \hat{G}_g^{\text{tc}}(y \mid X_i) \text{ y } T_i^\alpha \text{ de } \hat{L}_g^{\text{tc}}(y \mid X_i);$$

donde $\hat{F}_g^{\text{tc}}(y \mid X_i)$, $\hat{G}_g^{\text{tc}}(y \mid X_i)$ y $\hat{L}_g^{\text{tc}}(y \mid X_i)$ son los estimadores de las respectivas funciones de distribución condicionales con ventana $g = g_n$ (posiblemente diferente de $h = h_n$).

Definir $Z_i^\alpha = \min(Y_i^\alpha; C_i^\alpha)$ y $\pm_i^\alpha = 1_{F_{Y_i^\alpha} \cdot C_i^\alpha g}$. Si $T_i^\alpha \cdot Z_i^\alpha$ mantener la observación $(X_i; T_i^\alpha; Z_i^\alpha; \pm_i^\alpha)$; en otro caso repetir el proceso. Continuar hasta obtener la muestra bootstrap $(X_1; T_1^\alpha; Z_1^\alpha; \pm_1^\alpha); \dots; (X_n; T_n^\alpha; Z_n^\alpha; \pm_n^\alpha)$.

2. Utilizando la muestra bootstrap definir el análogo bootstrap de $\hat{F}_h^{\text{tc}}(y \mid x)$,

dado por:

$$\hat{F}_h^{\text{Boot}}(y | x) = \frac{\sum_{i=1}^n \mathbb{1}_{\{Z_i \leq y\}} \prod_{j=1}^n \mathbb{1}_{\{T_j \leq x\}} B_{hi}(x)}{\sum_{j=1}^n \mathbb{1}_{\{T_j \leq x\}} \prod_{i=1}^n B_{hj}(x)}$$

donde $B_{hi}(x)$ denota el peso de la estimación no paramétrica con ventana h .

3. Aproximar la distribución de $(nh)^{1/2}[\hat{F}_h^{\text{tc}}(y | x) - F(y | x)]$ por la distribución en el remuestreo de $(nh)^{1/2}[\hat{F}_h^{\text{Boot}}(y | x) - \hat{F}_g^{\text{tc}}(y | x)]$.

3.2.2 Propiedades de la aproximación bootstrap.

A continuación derivamos una representación casi segura para la aproximación bootstrap, y apoyándonos en ella, demostramos la consistencia del método de remuestreo propuesto. En la demostración de ambas propiedades, se utilizan las hipótesis H1-H10 enunciadas en el apartado 2.3.3 del capítulo anterior, junto con las siguientes:

(H11) $a < a_{H(\cdot|x)}$ y $b_{L(\cdot|x)} < b_{H(\cdot|x)}$; $\delta x \in I_{\pm}$:

(H12) K es dos veces continuamente diferenciable y con derivada primera acotada.

(H13) a) $g \neq 0$ y $\ln n = ng \neq 0$.

b) Además $g = h \neq 1$; $\frac{ng^5}{\ln n} \neq 1$ y $\frac{ng^5 h}{\ln n} = O(1)$:

(H14) Las funciones $m(x)$ y $\mathbb{R}(x)$ tienen tercera derivada respecto a x continua en I_{\pm} y las funciones $L(y | x)$; $H(y | x)$ y $H_1(y | x)$ tienen tercera derivada respecto a x continua y acotada en $(y; x) \in [0; 1) \in I_{\pm}$:

(H15) Las funciones $m(x)$ y $\mathbb{R}(x)$ tienen cuarta derivada respecto a x continua en I_{\pm} y las funciones $L(y | x)$; $H(y | x)$ y $H_1(y | x)$ tienen cuarta derivada respecto a x continua y acotada en $(y; x) \in [0; 1) \in I_{\pm}$:

Teorema 3.6 (representación casi segura). Bajo las hipótesis (H1)-(H13), para $x \in I$ e $y \in [a; b]$, se verifica que:

a)

$$\hat{\Lambda}_n^{\text{Boot}}(y | x) - \hat{\Lambda}_g^{\text{tc}}(y | x) = \sum_{i=1}^n B_{hi}(x) \mathbb{R}(Z_i; T_i; \pm_i; y; x) - \sum_{i=1}^n B_{gi}(x) \mathbb{R}(Z_i; T_i; \pm_i; y; x) + R_n^{\text{Boot}}(y | x) \quad (3.1)$$

donde

$$\hat{\Lambda}_h^{\text{Boot}}(y | x) = \frac{\sum_{i=1}^n 1_{fZ_i^\pm \cdot y g_i^\pm} B_{hi}(x)}{\sum_{j=1}^n 1_{fT_j^\pm \cdot Z_j^\pm g} B_{hj}(x)}$$

$$g(Z; T; \pm; y; x) = \frac{1_{fZ \cdot y; \pm=1g}}{C(Z | x)} \int_0^Z y \frac{1_{fT \cdot u \cdot Zg}}{(C(u | x))^2} dH_1^\pm(u | x)$$

y

$$\sup_{[a;b] \in I} |R_n^{\text{Boot}}(y | x)| = O_{P^n} \left(\frac{\tilde{A} \mu_{\ln n} \mathbb{1}_{3=4}}{nh} \right) \quad c:S(P)$$

b)

$$\hat{F}_h^{\text{Boot}}(y | x) - \hat{F}_g^{\text{tc}}(y | x) =$$

$$= (1 - F(y | x)) \sum_{i=1}^n B_{hi}(x) g(Z_i^\pm; T_i^\pm; \pm; y; x) - \sum_{i=1}^n B_{gi}(x) g(Z_i; T_i; \pm; y; x) +$$

$$+ R_n^{\text{Boot}}(y | x)$$

donde

$$\sup_{[a;b] \in I} |R_n^{\text{Boot}}(y | x)| = O_{P^n} \left(\frac{\tilde{A} \mu_{\ln n} \mathbb{1}_{3=4}}{nh} \right) \quad c:S(P):$$

Teorema 3.7 (consistencia). Bajo las hipótesis del teorema anterior junto con (H14), (H15) y $h = Cn^{-1-5}$; para $x \in I$ e $y \in [a; b]$; se tiene la convergencia hacia 0 en probabilidad de

$$\sup_{t \in \mathbb{R}^+} P^n \left(\frac{h}{(nh)^{1-2}} \left| \hat{F}_h^{\text{Boot}}(y | x) - \hat{F}_g^{\text{tc}}(y | x) \right| \cdot t \right) \rightarrow 0$$

$$\text{y } P^n \left(\frac{h}{(nh)^{1-2}} \left| \hat{F}_h^{\text{tc}}(y | x) - F(y | x) \right| \cdot t \right) \rightarrow 0$$

Conviene observar que las expresiones P^n ; E^n y Var^n denotan, respectivamente, las probabilidades, esperanzas y varianzas de ciertas variables condicionadas a la muestra observada inicial $f(X_i; T_i; Z_i; \pm_i)g_{i=1}^n$:

3.2.3 Lemas y demostraciones.

Hemos organizado este apartado de acuerdo con el siguiente esquema:

1. Lemas utilizados en la demostración del Teorema 3.6.
2. Demostración del Teorema 3.6.
3. Lemas utilizados en la demostración del Teorema 3.7.
4. Demostración del Teorema 3.7.

1. Lemas utilizados en la demostración del Teorema 3.6.

Lema 3.9 Bajo las hipótesis H1, H2a, H3, H4, H9, H10 y H13a, se verifica:

$$\sup_{[a:b] \in I} |\hat{H}_{1g}^{\alpha}(y|x) - H_1^{\alpha}(y|x)| = O\left(\frac{\tilde{\mu}_{\ln n}^{\alpha} \mathbb{1}_{1=2}}{ng}\right) + O(g^2) \quad \text{c.s.}$$

$$\sup_{[a:b] \in I} |\hat{C}_g(y|x) - C(y|x)| = O\left(\frac{\tilde{\mu}_{\ln n}^{\alpha} \mathbb{1}_{1=2}}{ng}\right) + O(g^2) \quad \text{c.s.}$$

Si además se verifica H13b, el orden anterior es una $O\left(\frac{\ln n}{nh}\right)^{1=2}$ c.s.:

Demostración. Como $g \rightarrow 0$ y $\ln n = (ng) \rightarrow 0$; podemos razonar como en la Propiedad 2.42 y en el Lema 2.1 y de esta forma obtener el orden $O\left(\frac{\ln n}{ng}\right)^{1=2} + O(g^2)$ c.s. para los dos supremos anteriores.

Por otra parte, teniendo en cuenta

$$\frac{\left(\frac{\ln n}{ng}\right)^{1=2}}{\left(\frac{\ln n}{nh}\right)^{1=2}} = \frac{\mu_h \mathbb{1}_{1=2}}{g} \rightarrow 0$$

y

$$\frac{g^2}{\left(\frac{\ln n}{nh}\right)^{1=2}} = \frac{\mu_{g^4 nh} \mathbb{1}_{1=2}}{\ln n} = \frac{\mu_{ng^5 h} \mathbb{1}_{1=2}}{\ln n g} = O(1)$$

se llega a que

$$O\left(\frac{\tilde{\mu}_{\ln n}^{\alpha} \mathbb{1}_{1=2}}{ng}\right) + O(g^2) = O\left(\frac{\tilde{\mu}_{\ln n}^{\alpha} \mathbb{1}_{1=2}}{nh}\right) :$$

Lema 3.10 Bajo las hipótesis H1-H4 y H9-H13, se veri...ca:

$$\sup_{[a;b] \in I} \|\hat{H}_{1h}^{\alpha \text{Boot}}(y | j, x) - \hat{H}_{1g}^{\alpha}(y | j, x)\| = O_{P^{\alpha}} \left(\frac{\tilde{A} \mu_{\ln n} \mathbb{1}_{1=2}}{nh} \right) \quad c:s:(P)$$

y

$$\sup_{[a;b] \in I} \|\hat{C}_h^{\alpha \text{Boot}}(y | j, x) - \hat{C}_g^{\alpha}(y | j, x)\| = O_{P^{\alpha}} \left(\frac{\tilde{A} \mu_{\ln n} \mathbb{1}_{1=2}}{nh} \right) \quad c:s:(P)$$

donde $\hat{H}_{1h}^{\alpha \text{Boot}}(y | j, x)$ y $\hat{C}_h^{\alpha \text{Boot}}(y | j, x)$ son los análogos bootstrap de $\hat{H}_{1h}^{\alpha}(y | j, x)$ y $\hat{C}_h^{\alpha}(y | j, x)$ respectivamente, es decir

$$\begin{aligned} \hat{H}_{1h}^{\alpha \text{Boot}}(y | j, x) &= \sum_{i=1}^X B_{hi}(x) 1_{F_{Z_i^{\alpha}} \cdot y; \pm_i^{\alpha} = 1g} \\ \hat{C}_h^{\alpha \text{Boot}}(y | j, x) &= \sum_{i=1}^X B_{hi}(x) 1_{F_{T_i^{\alpha}} \cdot y; Z_i^{\alpha} g} \end{aligned}$$

Demostración.

Comenzamos estudiando $\|\hat{H}_{1h}^{\alpha \text{Boot}}(y | j, x) - \hat{H}_{1g}^{\alpha}(y | j, x)\|$; expresión que acotamos por tres sumandos:

$$\begin{aligned} &\|\hat{H}_{1h}^{\alpha \text{Boot}}(y | j, x) - \hat{H}_{1g}^{\alpha}(y | j, x)\| \\ &= \|\hat{H}_{1h}^{\alpha \text{Boot}}(y | j, x) - E^{\alpha} \hat{H}_{1h}^{\alpha \text{Boot}}(y | j, x)\| + \|E^{\alpha} \hat{H}_{1h}^{\alpha \text{Boot}}(y | j, x) - \hat{H}_{1g}^{\alpha}(y | j, x)\| + \\ &\quad + \|\hat{H}_{1g}^{\alpha}(y | j, x) - \hat{H}_{1g}^{\alpha}(y | j, x)\| \\ &= S1(y | j, x) + S2(y | j, x) + S3(y | j, x) \end{aligned} \quad (3.2)$$

que hemos denotado por $S1(y | j, x)$, $S2(y | j, x)$ y $S3(y | j, x)$, respectivamente.

Para el sumando $S3(y | j, x)$ de la última expresión, en virtud del Lema 3.9, se veri...ca que:

$$\sup_{[a;b] \in I} S3(y | j, x) = \sup_{[a;b] \in I} \|\hat{H}_{1g}^{\alpha}(y | j, x) - \hat{H}_{1g}^{\alpha}(y | j, x)\| = O \left(\frac{\tilde{A} \mu_{\ln n} \mathbb{1}_{1=2}}{nh} \right) \quad c:s: \quad (3.3)$$

En el estudio de $\sup_{[a;b] \in I} S1(y | j, x)$ utilizaremos que

$$E^{\alpha} 1_{F_{Z_i^{\alpha}} \cdot y; \pm_i^{\alpha} = 1g} = \hat{H}_{1g}^{\alpha}(y | j, X_i) \quad (3.4)$$

cuestión que probamos a continuación.

Por comodidad y sin pérdida de generalidad estudiamos $E^{\pi} 1_{f_{Z_1^{\pi}, y; \pm_1^{\pi}}=1g}$:
Para ello se genera X_1 y se generan:

$$\begin{aligned} T_1^{\pi} &= T_j \text{ con probabilidad } d\hat{L}_g^{tc}(T_j | X_1) \\ (Z_1^{\pi}; \pm_1^{\pi}) &= (Z_i; \pm_i) \text{ con probabilidad } d\hat{H}_g^{tc}(Z_i | X_1) \end{aligned}$$

por tanto

$$\begin{aligned} P((T_1^{\pi}; Z_1^{\pi}; \pm_1^{\pi}) = (T_j; Z_i; \pm_i = 1)) &= P(T_1^{\pi} = T_j; Y_1^{\pi} = Z_i; C_1^{\pi} > Z_i) 1_{f_{T_j, Z_i}} \\ &= \frac{1_{f_{T_j, Z_i}} d\hat{L}_g^{tc}(T_j | X_1) d\hat{F}_g^{tc}(Z_i | X_1) 1_i \hat{G}_g^{tc}(Z_i | X_1)}{\prod_{k=1}^n \prod_{l=1}^n 1_{f_{T_k, Z_l}} d\hat{L}_g^{tc}(T_k | X_1) d\hat{H}_g^{tc}(Z_l | X_1)} \end{aligned}$$

Entonces $E^{\pi} 1_{f_{Z_1^{\pi}, y; \pm_1^{\pi}}=1g}$ es igual a

$$\begin{aligned} & \int_{\mathcal{X}} 1_{f_{Z_i, y}} P(Z_1^{\pi} = Z_i; \pm_1^{\pi} = 1) \\ &= \frac{\prod_{i=1}^n \int_{\mathcal{Y}} 1_{f_{Z_i, y}} 1_{f_{T_j, Z_i}} d\hat{L}_g^{tc}(T_j | X_1) d\hat{F}_g^{tc}(Z_i | X_1) 1_i \hat{G}_g^{tc}(Z_i | X_1)}{\prod_{k=1}^n \prod_{l=1}^n \int_{\mathcal{R}} \int_{\mathcal{R}} 1_{f_{z, y}} 1_{f_{t, z}} 1_i \hat{G}_g^{tc}(z | X_1) d\hat{L}_g^{tc}(t | X_1) d\hat{F}_g^{tc}(z | X_1)} \\ &= \frac{\int_{\mathcal{R}} \int_{\mathcal{R}} 1_{f_{z, y}} 1_{f_{t, z}} d\hat{L}_g^{tc}(t | X_1) d\hat{H}_g^{tc}(z | X_1)}{\int_{\mathcal{R}} 1_{f_{z, y}} d\hat{L}_g^{tc}(z | X_1) d\hat{H}_g^{tc}(z | X_1)} = \\ &= \frac{\int_{\mathcal{R}} 1_{f_{z, y}} \hat{L}_g^{tc}(z | X_1) 1_i \hat{G}_g^{tc}(z | X_1) d\hat{F}_g^{tc}(z | X_1)}{\int_{\mathcal{R}} \hat{L}_g^{tc}(z | X_1) d\hat{H}_g^{tc}(z | X_1)} \end{aligned} \quad (3.5)$$

Utilizando ahora el valor de $d\hat{F}_g^{tc}(z | X_1)$ (Propiedad 2.54) el numerador de la expresión anterior se escribe de la forma

$$\begin{aligned} & \int_{\mathcal{Z}} 1_{f_{z, y}} \hat{L}_g^{tc}(z | X_1) 1_i \hat{G}_g^{tc}(z | X_1) d\hat{F}_g^{tc}(z | X_1) \\ &= \int_{i=1}^n 1_{f_{Z_i, y}} \hat{L}_g^{tc}(Z_i | X_1) 1_i \hat{G}_g^{tc}(Z_i | X_1) \frac{\pm_i B_{gi}(X_1)}{\hat{C}_g(Z_i | X_1)} 1_i \hat{F}_g^{tc}(Z_i | X_1) \\ &= \int_{i=1}^n 1_{f_{Z_i, y}} \hat{L}_g^{tc}(Z_i | X_1) \pm_i 1_i \hat{H}_g^{tc}(Z_i | X_1) \frac{B_{gi}(X_1)}{\hat{C}_g(Z_i | X_1)} \\ &= \int_{i=1}^n 1_{f_{Z_i, y}} \hat{L}_g^{tc}(Z_i | X_1) \pm_i d\hat{H}_g^{tc}(Z_i | X_1) \end{aligned} \quad (3.6)$$

donde en los dos últimos pasos se han utilizado las propiedades 2.60 y 2.57, respectivamente.

Finalmente, utilizando (3.5) y (3.6) y la Propiedad 2.61, se llega al resultado deseado de:

$$\begin{aligned} & E^n \frac{1}{n} \sum_{i=1}^n 1_{fZ_i^{\pm}, y; \pm_i^{\pm}=1g} \\ &= \frac{\frac{1}{n} \sum_{i=1}^n 1_{fZ_i^{\pm}, yg} \hat{L}_g^{tc}(Z_i | X_1) \pm_i \frac{B_{gi}(X_1) \hat{L}_g^{tc}(Z_i | X_1)}{\frac{1}{n} \sum_{i=1}^n B_{gi}(X_1) \hat{L}_g^{tc}(Z_i | X_1)}}{\frac{1}{n} \sum_{i=1}^n \hat{L}_g^{tc}(Z_i | X_1) \frac{B_{gi}(X_1) \hat{L}_g^{tc}(Z_i | X_1)}{\frac{1}{n} \sum_{i=1}^n B_{gi}(X_1) \hat{L}_g^{tc}(Z_i | X_1)}}} \\ &= \frac{\frac{1}{n} \sum_{i=1}^n 1_{fZ_i^{\pm}, yg} \pm_i B_{gi}(X_1)}{\frac{1}{n} \sum_{i=1}^n B_{gi}(X_1)} = \hat{H}_{1g}^{\pm}(y | X_1): \end{aligned}$$

Haciendo uso entonces de $E^n \frac{1}{n} \sum_{i=1}^n 1_{fZ_i^{\pm}, y; \pm_i^{\pm}=1g} = \hat{H}_{1g}^{\pm}(y | X_i)$; el sumando denotado por $S1(y | x)$ se puede escribir como

$$S1(y | x) = \sum_{i=1}^n B_{hi}(x) 1_{fZ_i^{\pm}, y; \pm_i^{\pm}=1g} \pm_i \hat{H}_{1g}^{\pm}(y | X_i) = \sum_{i=1}^n W_i$$

donde

$$W_i = B_{hi}(x) 1_{fZ_i^{\pm}, y; \pm_i^{\pm}=1g} \pm_i \hat{H}_{1g}^{\pm}(y | X_i)$$

En el universo bootstrap (condicionando a la muestra observada inicial) las variables W_i son i.i.d. y $E^n W_i = 0$: Además

$$\begin{aligned} Var^n W_i &= E^n W_i^2 = E^n (B_{hi}(x))^2 1_{fZ_i^{\pm}, y; \pm_i^{\pm}=1g} \pm_i \hat{H}_{1g}^{\pm}(y | X_i)^2 \\ &= (B_{hi}(x))^2 E \\ &= (B_{hi}(x))^2 E \frac{1}{n} \sum_{i=1}^n 1_{fZ_i^{\pm}, y; \pm_i^{\pm}=1g} \pm_i \hat{H}_{1g}^{\pm}(y | X_i) + \frac{1}{n} \sum_{i=1}^n \hat{H}_{1g}^{\pm}(y | X_i)^2 \\ &= (B_{hi}(x))^2 \hat{H}_{1g}^{\pm}(y | X_i) \pm_i \hat{H}_{1g}^{\pm}(y | X_i) \end{aligned}$$

y por tanto

$$\begin{aligned} \sum_{i=1}^n Var^n W_i &= \sum_{i=1}^n (B_{hi}(x))^2 \hat{H}_{1g}^{\pm}(y | X_i) \pm_i \hat{H}_{1g}^{\pm}(y | X_i) \\ &= O\left(\frac{1}{nh}\right) \text{ c.s.} \end{aligned}$$

También se verifica que

$$|W_{ij}| = |B_{hi}(x) - \mathbb{E}[\hat{H}_{1g}^{\mu}(y_j | X_i)]| \leq \max B_{hi}(x) = O\left(\frac{1}{nh}\right) \quad \text{c.s.}$$

y aplicando la desigualdad de Bernstein en este contexto obtenemos, de forma casi segura, que:

$$P^{\mu} \left\{ \sum_{i=1}^n |W_{ij}| > \frac{\tilde{A}}{2} \right\} \leq 2 \exp \left\{ - \frac{\tilde{A}^2}{2c_1 \frac{1}{nh} + \frac{2}{3} c_2 \frac{1}{nh}} \right\} + 2 \exp \left\{ - \frac{\tilde{A}^2 nh}{2c_1 (1 + \frac{1}{3})} \right\} + 2 \exp \left\{ - \frac{\tilde{A}^2 nh}{4c} \right\}$$

Tomando $\tilde{A} = c^{\mu} \frac{\ln n}{nh} \mathbb{1}_{1=2}$ se tiene

$$P^{\mu} \left\{ \sum_{i=1}^n |W_{ij}| > \frac{\tilde{A}}{2} \right\} \leq 2 \exp \left\{ - \frac{c^{\mu 2} \ln n}{4c} \right\} = 2n^{-\frac{c^{\mu 2}}{4c}} \quad (3.7)$$

sucesión que converge a cero tomando c^{μ} de forma apropiada. Por tanto

$$S_1(y_j | x) = \sum_{i=1}^n |W_{ij}| = O_P^{\mu} \left(\frac{\ln n}{nh} \right) \mathbb{1}_{1=2} \quad \text{c.s.}(P) \quad (3.8)$$

Nosotros queremos estudiar el $\sup_{x \in \mathcal{X}} \sup_{y \in \mathcal{Y}_{2[a;b]}} S_1(y_j | x)$. Para ello, razonamos de la forma siguiente:

En un primer paso sea x fijo, y estudiemos

$$S_1(x) = \sup_{y \in \mathcal{Y}_{2[a;b]}} |\hat{H}_{1h}^{\mu \text{Boot}}(y_j | x) - \mathbb{E}^{\mu} \hat{H}_{1h}^{\mu \text{Boot}}(y_j | x)|$$

Dividimos el intervalo $[a; b]$ en $c_n \gg \frac{\ln n}{nh} \mathbb{1}_{1=2}$ subintervalos $[y_j; y_{j+1}]$ con $j = 1; \dots; c_n$, donde $y_1 = a$, $y_{c_n+1} = b$. Debido a la monotonía de $\hat{H}_{1h}^{\mu \text{Boot}}(y_j | x)$ y de $\mathbb{E}^{\mu} \hat{H}_{1h}^{\mu \text{Boot}}(y_j | x)$ en y ; se tiene que:

$$S_1(x) = \sup_{y \in \mathcal{Y}_{2[a;b]}} |\hat{H}_{1h}^{\mu \text{Boot}}(y_j | x) - \mathbb{E}^{\mu} \hat{H}_{1h}^{\mu \text{Boot}}(y_j | x)| \\ \leq \max_j |\hat{H}_{1h}^{\mu \text{Boot}}(y_j | x) - \mathbb{E}^{\mu} \hat{H}_{1h}^{\mu \text{Boot}}(y_j | x)| \\ + \max_j |\mathbb{E}^{\mu} \hat{H}_{1h}^{\mu \text{Boot}}(y_{j+1} | x) - \mathbb{E}^{\mu} \hat{H}_{1h}^{\mu \text{Boot}}(y_j | x)| \quad (3.9)$$

Además, el segundo sumando de la expresión anterior se acota por:

$$\begin{aligned} & \max_j \left| E^n \hat{H}_{1h}^{n, \text{Boot}}(y_{j+1} | j, X) - E^n \hat{H}_{1h}^{n, \text{Boot}}(y_j | j, X) \right| \\ &= \max_j \left| \sum_{i=1}^3 B_{hi}(X) \left(\hat{H}_{1g}^n(y_{j+1} | j, X_i) - \hat{H}_{1g}^n(y_j | j, X_i) \right) \right| \\ & \cdot \max_j \left| \sum_{i=1}^3 B_{hi}(X) \left(\hat{H}_{1g}^n(y_{j+1} | j, X_i) - \hat{H}_{1g}^n(y_j | j, X_i) \right) \right| \\ & \quad + \left(\hat{H}_1^n(y_{j+1} | j, X) - \hat{H}_1^n(y_j | j, X) \right) j \\ & \quad + \max_j j \left| \hat{H}_1^n(y_{j+1} | j, X) - \hat{H}_1^n(y_j | j, X) \right| \end{aligned}$$

que, a su vez, se puede mayorar por

$$\begin{aligned} & \max_j \left| \sum_{i=1}^3 B_{hi}(X) \left(\hat{H}_{1g}^n(y_{j+1} | j, X_i) - \hat{H}_{1g}^n(y_{j+1} | j, X) \right) \right| + \\ & \quad + \max_j \left| \sum_{i=1}^3 B_{hi}(X) \left(\hat{H}_{1g}^n(y_j | j, X_i) - \hat{H}_{1g}^n(y_j | j, X) \right) \right| + \\ & \quad + \max_j j \left| \hat{H}_1^n(y_{j+1} | j, X) - \hat{H}_1^n(y_j | j, X) \right| \\ & \cdot 2 \sup_{y \in [a; b]} \left| \sum_{i=1}^3 B_{hi}(X) \left(\hat{H}_{1g}^n(y | j, X_i) - \hat{H}_{1g}^n(y | j, X) \right) \right| + \end{aligned} \tag{3.10}$$

$$+ \max_j j \left| \hat{H}_1^n(y_{j+1} | j, X) - \hat{H}_1^n(y_j | j, X) \right|: \tag{3.11}$$

Utilizando (3.9), (3.10) y (3.11) se llega a la siguiente acotación de $S_1(x)$:

$$\begin{aligned} S_1(x) & \cdot \max_j \left| \hat{H}_{1h}^{n, \text{Boot}}(y_j | j, X) - E^n \hat{H}_{1h}^{n, \text{Boot}}(y_j | j, X) \right| + \\ & \quad + \max_j j \left| \hat{H}_1^n(y_{j+1} | j, X) - \hat{H}_1^n(y_j | j, X) \right| j + \\ & \quad + 2 \sup_{y \in [a; b]} \left| \sum_{i=1}^3 B_{hi}(X) \left(\hat{H}_{1g}^n(y | j, X_i) - \hat{H}_{1g}^n(y | j, X) \right) \right| \\ & = S_{11}(x) + S_{12}(x) + S_{13}(x) \end{aligned}$$

y por tanto

$$\sup_{x \in \mathcal{X}} S_1(x) \leq \sup_{x \in \mathcal{X}} S_{11}(x) + \sup_{x \in \mathcal{X}} S_{12}(x) + \sup_{x \in \mathcal{X}} S_{13}(x): \tag{3.12}$$

Estudiamos estos tres sumandos separadamente.

En virtud de la hipótesis H5 y de las condiciones sobre la partición $f[y_j; y_{j+1}]g$ se tiene que

$$\sup_{x \in I} S_{12}(x) = \sup_{x \in I} \max_j |H_{1h}^{\alpha}(y_{j+1} | x) - H_{1h}^{\alpha}(y_j | x)| = O\left(\frac{\mu_{1=2}}{nh}\right) \quad (3.13)$$

En cuanto al primer sumando del segundo miembro de (3.12) probaremos ahora que

$$\sup_{x \in I} S_{11}(x) = O_{P_{\alpha}}\left(\frac{\mu_{1=2}}{nh}\right) \quad c.s.(P) \quad (3.14)$$

Para ello (siguiendo a Cheng y Cheng, 1987) consideramos en el intervalo I un conjunto E_n de, a lo sumo $n^2 + 1$ puntos, $E_n = \{x_n\}$, de forma que $x_n \in I$; $x_n - x_{n-1} \leq C_1/n^2$ (donde C_1 denota cierta constante). Entonces, se veri...ca:

$$\begin{aligned} \sup_{x \in I} S_{11}(x) &= \sup_{x \in I} \max_j | \hat{H}_{1h}^{\alpha \text{Boot}}(y_j | x) - E^{\alpha} \hat{H}_{1h}^{\alpha \text{Boot}}(y_j | x) | \quad (3.15) \\ &\leq \max_{x_n \in E_n} \max_j | \hat{H}_{1h}^{\alpha \text{Boot}}(y_j | x_n) - E^{\alpha} \hat{H}_{1h}^{\alpha \text{Boot}}(y_j | x_n) | + \\ &\quad + \sup_{x \in I} \max_j | \hat{H}_{1h}^{\alpha \text{Boot}}(y_j | x) - \hat{H}_{1h}^{\alpha \text{Boot}}(y_j | x_n) | + \\ &\quad + | E^{\alpha} \hat{H}_{1h}^{\alpha \text{Boot}}(y_j | x) - E^{\alpha} \hat{H}_{1h}^{\alpha \text{Boot}}(y_j | x_n) | \end{aligned}$$

En primer lugar, para todo $x \in I$; se tiene que

$$\begin{aligned} P^{\alpha} \left[\max_j | \hat{H}_{1h}^{\alpha \text{Boot}}(y_j | x) - E^{\alpha} \hat{H}_{1h}^{\alpha \text{Boot}}(y_j | x) | > \frac{\mu_{1=2}}{3nh} \right] &> 0 \\ &= P^{\alpha} \left[\max_j | \hat{H}_{1h}^{\alpha \text{Boot}}(y_j | x) - \hat{H}_{1h}^{\alpha \text{Boot}}(y_j | x_n) | > \frac{\mu_{1=2}}{3nh} \right] \\ &\leq P^{\alpha} \left[\max_j | \hat{H}_{1h}^{\alpha \text{Boot}}(y_j | x) - \hat{H}_{1h}^{\alpha \text{Boot}}(y_j | x_n) | > \frac{\mu_{1=2}}{3nh} \right] \quad (3.16) \end{aligned}$$

Tomando $\mu_{1=2} = c^{\alpha} \frac{\ln n}{nh} \mu_{1=2}$ y utilizando (3.7), resultado que se cumple uniformemente en $y \in [a; b]$; se llega a que

$$P^{\alpha} \left[\max_j | \hat{H}_{1h}^{\alpha \text{Boot}}(y_j | x) - E^{\alpha} \hat{H}_{1h}^{\alpha \text{Boot}}(y_j | x) | > c^{\alpha} \frac{\mu_{1=2}}{nh} \right] \leq (C_n + 1) 2n^{-\frac{c^{\alpha} n^2}{4c}}$$

resultado que se tiene para todo $x \in \mathbb{R}^1$ y por tanto para cualquier \hat{x}_n : De este modo y razonando como en (3.16), llegamos a que:

$$P^{\alpha} \max_{\hat{x}_n \in E_n} \max_j \left| \hat{H}_{1h}^{\alpha \text{Boot}}(y_j | \hat{x}_n) - E^{\alpha} \hat{H}_{1h}^{\alpha \text{Boot}}(y_j | \hat{x}_n) \right| > c^{\alpha} \frac{\mu_{1=2}}{nh} \\ \times P^{\alpha} \max_{\hat{x}_n \in E_n} \max_j \left| \hat{H}_{1h}^{\alpha \text{Boot}}(y_j | \hat{x}_n) - E^{\alpha} \hat{H}_{1h}^{\alpha \text{Boot}}(y_j | \hat{x}_n) \right| > c^{\alpha} \frac{\mu_{1=2}}{nh} \\ \cdot \text{card}(E_n) (c_n + 1) 2n^{\frac{c^{\alpha 2}}{4c}}$$

que para $\text{card}(E_n) = n^2 + 1$ y $c_n \gg \frac{\mu_{1=2}}{nh}$ es sumable tomando un c^{α} conveniente. Esto prueba, para el primer sumando que acota (3.15) que

$$\max_{\hat{x}_n \in E_n} \max_j \left| \hat{H}_{1h}^{\alpha \text{Boot}}(y_j | \hat{x}_n) - E^{\alpha} \hat{H}_{1h}^{\alpha \text{Boot}}(y_j | \hat{x}_n) \right| = O_{P^{\alpha}} \left(\frac{\mu_{1=2}}{nh} \right) \quad c:s:(P): \quad (3.17)$$

En cuanto al segundo sumando que acota (3.15) se puede escribir como

$$\sup_{x \in \mathbb{R}^1} \max_{j=1}^3 \left| (B_{hi}(x) - B_{hi}(\hat{x}_n)) \cdot 1_{f_{Z_i^{\alpha}}(y_j | x) = 1} \cdot \hat{H}_{1g}^{\alpha}(y_j | X_i) \right| \\ \cdot 2 \sup_{x \in \mathbb{R}^1} |B_{hi}(x) - B_{hi}(\hat{x}_n)| \quad (3.18)$$

e inmediatamente comprobaremos que

$$\sup_{x \in \mathbb{R}^1} |B_{hi}(x) - B_{hi}(\hat{x}_n)| = O \left(\frac{\mu_{1=2}}{nh} \right) \quad c:s: \quad (3.19)$$

De este modo, (3.15), (3.17), (3.18) y (3.19) prueban (3.14).

Probemos ahora (3.19).

En primer lugar, escribimos $B_{hi}(x) - B_{hi}(\hat{x}_n)$ de la forma:

$$|B_{hi}(x) - B_{hi}(\hat{x}_n)| = \left| \frac{\frac{1}{nh} \sum_{i=1}^n \frac{x_i X_i}{h}}{\hat{m}_h^{\alpha}(x)} - \frac{\frac{1}{nh} \sum_{i=1}^n \frac{\hat{x}_{ni} X_i}{h}}{\hat{m}_h^{\alpha}(\hat{x}_n)} \right| \\ = \frac{1}{\hat{m}_h^{\alpha}(x) \hat{m}_h^{\alpha}(\hat{x}_n)} \left[\hat{m}_h^{\alpha}(\hat{x}_n) \frac{1}{nh} \sum_{i=1}^n \frac{x_i X_i}{h} - \hat{m}_h^{\alpha}(x) \frac{1}{nh} \sum_{i=1}^n \frac{\hat{x}_{ni} X_i}{h} \right]$$

Sumando y restando $\hat{m}_h^a(\hat{x}_n) \frac{1}{nh} \sum_{i=1}^n K\left(\frac{\hat{x}_n - X_i}{h}\right)$; agrupando términos y utilizando la desigualdad triangular acotamos la expresión anterior por

$$\begin{aligned} & \frac{1}{\hat{m}_h^a(x)\hat{m}_h^a(\hat{x}_n)} \left(\hat{m}_h^a(\hat{x}_n) - \hat{m}_h^a(x) \right) \frac{1}{nh} \sum_{i=1}^n K\left(\frac{\hat{x}_n - X_i}{h}\right) + \\ & + \frac{1}{\hat{m}_h^a(x)\hat{m}_h^a(\hat{x}_n)} \hat{m}_h^a(\hat{x}_n) \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x - X_i}{h}\right) - \frac{1}{nh} \sum_{i=1}^n K\left(\frac{\hat{x}_n - X_i}{h}\right) \end{aligned} \quad (3.20)$$

Y entonces:

$$\begin{aligned} & \sup_{x \in I} \left| \sum_{i=1}^n B_{hi}(x) - B_{hi}(\hat{x}_n) \right| \\ & \cdot \sup_{x \in I} \frac{1}{\hat{m}_h^a(x)} \left| \hat{m}_h^a(\hat{x}_n) - \hat{m}_h^a(x) \right| + \end{aligned} \quad (3.21)$$

$$+ \sup_{x \in I} \frac{1}{\hat{m}_h^a(x)} \left| \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x - X_i}{h}\right) - \frac{1}{nh} \sum_{i=1}^n K\left(\frac{\hat{x}_n - X_i}{h}\right) \right| \quad (3.22)$$

El término (3.21) se acota por

$$\begin{aligned} (3.21) \cdot & \sup_{x \in I} \frac{1}{\hat{m}_h^a(x)} \left| \hat{m}_h^a(\hat{x}_n) - \hat{m}_h^a(x) \right| + \\ & + \sup_{x \in I} \frac{1}{\hat{m}_h^a(x)} \left| \hat{m}_h^a(\hat{x}_n) - \hat{m}_h^a(x) \right| + \\ & + \sup_{x \in I} \frac{1}{\hat{m}_h^a(x)} \left| \hat{m}_h^a(x) - \hat{m}_h^a(\hat{x}_n) \right| \end{aligned}$$

expresiones que tienen orden $O\left(\frac{\ln n}{nh}\right)^{c_1=2}$ c:s: debido a:

i) $\hat{m}_h^a(x) \geq m^a(x) > 0$; $\forall x \in I$; de forma que

$$\sup_{x \in I} \left| \hat{m}_h^a(x) - m^a(x) \right| = O\left(\frac{\ln n}{nh}\right)^{c_1=2} \quad \text{c:s:}$$

ii) En virtud de H3 se tiene

$$\sup_x \left| \hat{m}_h^a(\hat{x}_n) - \hat{m}_h^a(x) \right| \cdot \sup_x \left| m^a(\hat{x}_n) - m^a(x) \right| \leq C \frac{C_1}{n^2}$$

y

$$\frac{\left(\frac{\ln n}{nh}\right)^{c_1=2}}{\frac{1}{n^2}} = \frac{\ln^3 n}{h} \ll 1$$

Por tanto

$$(3:21) = O\left(\frac{\tilde{A}\mu_{\ln n}^{(1)}!}{nh}\right) \quad \text{c:s:} \quad (3:23)$$

En cuanto al término (3.22), es fácil ver que

$$\begin{aligned} (3:22) &= \sup_x \frac{1}{\hat{m}_h^a(x)} \sum_{i=1}^n \frac{1}{nh} K\left(\frac{x_i - X_i}{h}\right) \frac{1}{nh} K\left(\frac{\hat{x}_{ni} - X_i}{h}\right) \\ &\cdot \sup_x \frac{1}{\hat{m}_h^a(x)} \sum_{i=1}^n \frac{1}{nh} K^0(\cdot) \left(\frac{x_i - X_i}{h}\right) \frac{\hat{x}_{ni} - X_i}{h} \\ &= \sup_x \frac{1}{\hat{m}_h^a(x)} \frac{1}{h^2} K^0(\cdot) \sum_{i=1}^n (x_i - \hat{x}_{ni}) = O\left(\frac{1}{n^2 h^2}\right) \quad \text{c:s:} \end{aligned} \quad (3:24)$$

sin más que aplicar la hipótesis H12. Este último resultado junto con (3.23), (3.22) y (3.21) prueban (3.19).

El siguiente paso en esta demostración consiste en acotar el tercer sumando del segundo miembro de la expresión (3.12), dado por

$$\sup_{x \in I} S_{13}(x) = 2 \sup_{x \in I} \sum_{i=1}^n B_{hi}(x) \hat{H}_{1g}^a(y_j | X_i) - H_1^a(y_j | x)$$

Utilizando la desigualdad triangular se obtiene fácilmente que

$$\sum_{i=1}^n B_{hi}(x) \hat{H}_{1g}^a(y_j | X_i) - H_1^a(y_j | x) + \sum_{i=1}^n B_{hi}(x) \hat{H}_{1g}^a(y_j | X_i) - H_1^a(y_j | X_i) \quad (3:25)$$

$$+ \sum_{i=1}^n B_{hi}(x) (H_1^a(y_j | X_i) - H_1^a(y_j | x)) \quad (3:26)$$

Analizamos el $\sup_{[a;b] \in I}$ de (3:25).

En primer lugar, gracias a la hipótesis H9, dicho supremo se puede acotar de la siguiente forma:

$$\begin{aligned} &\sup_{[a;b] \in I} \sum_{i=1}^n B_{hi}(x) \hat{H}_{1g}^a(y_j | X_i) - H_1^a(y_j | X_i) \\ &\cdot \sup_{x \in I} \sup_{f \times 2(x_i - h; x+h); y \in [a;b]g} \hat{H}_{1g}^a(y_j | \tilde{x}) - H_1^a(y_j | \tilde{x}) \end{aligned} \quad (3:27)$$

Ahora, recordando la definición de los intervalos I e I_{\pm} en la hipótesis H2 y haciendo tender $h \rightarrow 0$ podemos encontrar un $\pm_1 = 0 < \pm_1 < \pm$ de forma que

$$(3.27) = \sup_{[a;b] \in I_{\pm_1}} |\hat{H}_{1g}^n(y | \tilde{x}) - H_1^n(y | \tilde{x})|$$

donde $I_{\pm_1} = [x_1 - \pm_1; x_2 + \pm_1]$: Revisando las hipótesis de los Lemas 2.1 y 3.9 respecto al intervalo I , es sencillo comprobar que se siguen verificando para I_{\pm_1} ; y de este modo

$$\sup_{[a;b] \in I_{\pm_1}} |\hat{H}_{1g}^n(y | \tilde{x}) - H_1^n(y | \tilde{x})| = O\left(\frac{\tilde{A}_\mu \ln n}{nh} \right)_{1=2} \quad \text{c.s.}$$

lo cual, finalmente prueba que para (3.25) se tiene:

$$\sup_{[a;b] \in I} \sum_{i=1}^n B_{hi}(x) |\hat{H}_{1g}^n(y | X_i) - H_1^n(y | X_i)| = O\left(\frac{\tilde{A}_\mu \ln n}{nh} \right)_{1=2} \quad \text{c.s.} \quad (3.28)$$

El $\sup_{[a;b] \in I}$ de (3.26), esto es

$$\sup_{[a;b] \in I} \sum_{i=1}^n B_{hi}(x) (H_1^n(y | X_i) - H_1^n(y | x))$$

se puede ver como el $\sup_{[a;b] \in I} |\hat{z}(y | x) - z(y | x)|$; donde

$$\hat{z}(y | x) = \sum_{i=1}^n B_{hi}(x) H_1^n(y | X_i)$$

es un estimador tipo núcleo con pesos de Nadaraya-Watson de la función $z(y | x) = H_1^n(y | x)$:

Teniendo en cuenta que las variables Z_i son continuas y positivas, que $H_1^n(y | X_i)$ toma valores entre 0 y 1, y que además las funciones $m^n(x)$ y $E(H_1^n(y | X_i) | X_i = x) = H_1^n(y | x)$ verifican las hipótesis de la Propiedad 2.42, se puede reproducir la demostración de dicha propiedad para el estimador $\hat{z}(y | x)$; y así obtener que

$$\sup_{[a;b] \in I} \sum_{i=1}^n B_{hi}(x) |H_1^n(y | X_i) - H_1^n(y | x)| = O\left(\frac{\tilde{A}_\mu \ln n}{nh} \right)_{1=2} \quad \text{c.s.}$$

Como consecuencia de este último resultado, (3.28), (3.25) y (3.26), se prueba que

$$\sup_{x \in I} S1_3(x) = O\left(\frac{\tilde{A}_\mu \ln n}{nh} \right)_{1=2} \quad \text{c.s.} \quad (3.29)$$

Finalmente, (3.29), (3.14), (3.13) y (3.12) llevan a

$$\sup_{x \in I} S1(x) = \sup_{[a;b] \in I} S1(y | x) = O_{P^n} \left(\frac{\tilde{A} \mu_{\ln n} \mathbb{1}_{1=2}}{nh} \right) \quad c:s:(P) \quad (3.30)$$

El siguiente paso en esta demostración es acotar $\sup_{[a;b] \in I} S2(y | x)$ (ver (3.2)), lo cual resulta inmediato puesto que

$$\sup_{[a;b] \in I} S2(y | x) = \frac{1}{2} \sup_{x \in I} S1_3(x) = O \left(\frac{\tilde{A} \mu_{\ln n} \mathbb{1}_{1=2}}{nh} \right) \quad c:s: \quad (3.31)$$

como se probó en (3.29).

De este último resultado junto con (3.30), (3.3) y (3.2) se sigue la primera parte del presente lema:

$$\sup_{x \in I; y \in [a;b]} \left| \hat{H}_{1h}^{n,Boot}(y | x) - \hat{H}_{1g}^n(y | x) \right| = O_{P^n} \left(\frac{\tilde{A} \mu_{\ln n} \mathbb{1}_{1=2}}{nh} \right) \quad c:s:(P) \quad (3.32)$$

Para la demostración de la segunda parte, tendremos en cuenta que:

$$\begin{aligned} & \sup_{[a;b] \in I} \left| \hat{C}_h^{n,Boot}(y | x) - \hat{C}_g(y | x) \right| \\ &= \sup_{[a;b] \in I} \left| \hat{L}_h^{n,Boot}(y | x) - \hat{L}_g^n(y | x) \right| + \sup_{[a;b] \in I} \left| \hat{H}_h^{n,Boot}(y | x) - \hat{H}_g^n(y | x) \right| \\ & \cdot \sup_{[a;b] \in I} \left| \hat{L}_h^{n,Boot}(y | x) - \hat{L}_g^n(y | x) \right| + \sup_{[a;b] \in I} \left| \hat{H}_h^{n,Boot}(y | x) - \hat{H}_g^n(y | x) \right| \end{aligned}$$

donde

$$\hat{L}_h^{n,Boot}(y | x) = \sum_{i=1}^X B_{hi}(x) 1_{F_{T_i}^n \leq y}$$

y

$$\hat{H}_h^{n,Boot}(y | x) = \sum_{i=1}^X B_{hi}(x) 1_{F_{Z_i}^n \leq y}$$

Ahora, basta seguir paso a paso la demostración del apartado a), pero para $\sup_{[a;b] \in I} \left| \hat{L}_h^{n,Boot}(y | x) - \hat{L}_g^n(y | x) \right|$ y $\sup_{[a;b] \in I} \left| \hat{H}_h^{n,Boot}(y | x) - \hat{H}_g^n(y | x) \right|$; respectivamente.

Conviene solamente destacar que:

$$E^n \hat{L}_h^{a, \text{Boot}}(y | x) = \sum_{i=1}^n B_{hi}(x) E^n \frac{1}{f_{T_i}^a} y g = \sum_{i=1}^n B_{hi}(x) \hat{L}_g^a(y | X_i)$$

y

$$E^n \hat{H}_h^{a, \text{Boot}}(y | x) = \sum_{i=1}^n B_{hi}(x) E^n \frac{1}{f_{Z_i}^a} y g = \sum_{i=1}^n B_{hi}(x) \hat{H}_g^a(y | X_i);$$

resultados que se obtienen de forma análoga a (3.4).

Lema 3.11 Bajo las hipótesis H1, H2a), H3-H10 y H13, se verifica, con probabilidad 1, que:

$$\sup_{x \in \mathcal{X}} \sup_{f_s; t_2[a;b]=s_i} | \hat{H}_{1g}^a(t | x) - \hat{H}_{1g}^a(s | x) - H_1^a(t | x) + H_1^a(s | x) |$$

y

$$\sup_{x \in \mathcal{X}} \sup_{f_s; t_2[a;b]=s_i} | \hat{C}_g(t | x) - \hat{C}_g(s | x) - C(t | x) + C(s | x) |$$

tienen orden $O \left(\frac{\ln n}{nh} \right)^{3-4}$, donde $b_n = O \left(\frac{\ln n}{nh} \right)^{1-2}$:

Demostración. Como $g \neq 0$ y $\ln n = (ng) \neq 0$; podemos utilizar los Lemas 2.2 y 2.3 con ventana g para las funciones del presente lema, y de esta forma obtener el orden de los supremos anteriores, que será:

$$O \left(\frac{\ln n}{ng} \right)^{1-2} + O \left(\frac{\ln n}{ng} \right)^2 + O(b_n g) \quad \text{c:s:} \quad (3.33)$$

donde $(b_n)_{n \geq 2N}$ denota cualquier sucesión convergente a 0 y tal que $b_n \left(\frac{\ln n}{ng} \right)^{1-2} > \Delta > 0$:

La sucesión $O \left(\frac{\ln n}{nh} \right)^{1-2}$ es una tal sucesión b_n , puesto que:

$$\frac{\ln n}{nh} \left(\frac{\ln n}{ng} \right)^{1-2} = O \left(\frac{\ln n}{nh} \right)^{1-2} = \frac{ng^2}{h \ln n} \left(\frac{\ln n}{ng} \right)^{1-2} = \frac{ng}{\ln n h} \left(\frac{\ln n}{ng} \right)^{1-2} \quad \text{c:s:} \quad (> \Delta > 0):$$

Además, para dicha elección de b_n se tiene que

$$(3.33) = O \left(\frac{\ln n}{nh} \right)^{3-4} \quad \text{c:s:} \quad (3.34)$$

sin más que utilizar que $g=h \{ 1 ; \frac{ng^5}{\ln n} \} \{ 1 \}$ y $\frac{ng^5 h}{\ln n g} = O(1)$; y teniendo en cuenta las relaciones:

$$\frac{\frac{3}{\ln n} b_n \{ 1=2 \}}{\frac{1}{\ln n} \{ 3=4 \}} = \frac{\frac{3}{\ln n} \{ 1=2 \} \frac{1}{\ln n} \{ 1=4 \}}{\frac{1}{\ln n} \{ 3=4 \}} = \frac{\mu h \{ 1=2 \}}{g} \{ 0 \}$$

y

$$\frac{b_n g}{\frac{1}{\ln n} \{ 3=4 \}} = \frac{\frac{1}{\ln n} \{ 1=2 \} g}{\frac{1}{\ln n} \{ 3=4 \}} = \frac{g}{\frac{1}{\ln n} \{ 1=4 \}} = \frac{\mu ng^5 h}{\ln n g} \{ 1=4 \} = O(1):$$

De este modo, (3.33) y (3.34) prueban el resultado deseado.

Lema 3.12 Bajo las hipótesis H1-H13, se veri...ca que

$$\sup_{x \geq 1} \sup_{fs: t2[a;b]=js_i} \{ \hat{H}_{1h}^{aBoot}(t j x) \} \{ \hat{H}_{1h}^{aBoot}(s j x) \} \{ \hat{H}_{1g}^a(t j x) + \hat{H}_{1g}^a(s j x) \}$$

y

$$\sup_{x \geq 1} \sup_{fs: t2[a;b]=js_i} \{ \hat{C}_h^{Boot}(t j x) \} \{ \hat{C}_h^{Boot}(s j x) \} \{ \hat{C}_g(t j x) + \hat{C}_g(s j x) \}$$

tienen orden

$$O_P \{ \frac{\bar{A} \mu \ln n \{ 1=3=4 \}}{nh} \} \{ c:s(P) \}$$

donde $b_n = O \{ \frac{1}{\ln n} \{ 1=2 \} \}$:

Demostración.

Analizamos

$$\sup_{x \geq 1} \sup_{fs: t2[a;b]=js_i} \{ \hat{H}_{1h}^{aBoot}(t j x) \} \{ \hat{H}_{1h}^{aBoot}(s j x) \} \{ \hat{H}_{1g}^a(t j x) + \hat{H}_{1g}^a(s j x) \} \tag{3.35}$$

que se puede acotar por la suma de los respectivos $\sup_{x \geq 1} \sup_{fs: t2[a;b]=js_i} \{ t j \cdot b_n g \}$ de los dos términos siguientes:

$$\{ \hat{H}_{1h}^{aBoot}(t j x) \} \{ \hat{H}_{1h}^{aBoot}(s j x) \} \{ E^a \hat{H}_{1h}^{aBoot}(t j x) + E^a \hat{H}_{1h}^{aBoot}(s j x) \} \tag{3.36}$$

y

$$E^{\alpha} \hat{H}_{1g}^{\alpha \text{Boot}}(t_j | x) - E^{\alpha} \hat{H}_{1g}^{\alpha \text{Boot}}(s_j | x) - \hat{H}_{1g}^{\alpha}(t_j | x) + \hat{H}_{1g}^{\alpha}(s_j | x) \quad (3.37)$$

Comenzamos estudiando $\sup_{x \in I} \sup_{f_{S; t_2[a; b] = s_j} | t_j \in b_{ng}} (3.36)$ y en un primer momento ...jemos $x \in I$:

Razonando como en la demostración del Lema A.5 de Van Keilegom y Veraverbeke (1997a) dividimos el intervalo $[a; b]$ en subintervalos I_i de centro t_i y radio b_n ; de forma que $t_0 = a$, $t_m = b$ y, si $s_j \in I_i$, existe un I_i con $i = 1; 2; \dots; m$, tal que $s_j \in I_i$. Cada intervalo I_i se divide a su vez en subintervalos cuyos extremos denotamos por t_{ij} , de forma que $t_{ij} = t_i + j \frac{b_n}{a_n}$, con $j = 1; 2; \dots; a_n$ y $a_n \gg \frac{\ln n}{nh} \rightarrow 0$. Utilizando estas particiones y la monotonía de $\hat{H}_{1h}^{\alpha \text{Boot}}(\cdot | x)$ y de $E^{\alpha} \hat{H}_{1h}^{\alpha \text{Boot}}(\cdot | x)$, se tiene:

$$\begin{aligned} & \sup_{f_{S; t_2[a; b] = s_j} | t_j \in b_{ng}} \quad (3.36) \\ & \max_i \max_{j; k} \left[\hat{H}_{1h}^{\alpha \text{Boot}}(t_{ik} | x) - \hat{H}_{1h}^{\alpha \text{Boot}}(t_{ij} | x) \right] \\ & \leq E^{\alpha} \hat{H}_{1h}^{\alpha \text{Boot}}(t_{ik} | x) + E^{\alpha} \hat{H}_{1h}^{\alpha \text{Boot}}(t_{ij} | x) + \\ & + 2 \max_i \max_j \left[E^{\alpha} \hat{H}_{1h}^{\alpha \text{Boot}}(t_{i, j+1} | x) - E^{\alpha} \hat{H}_{1h}^{\alpha \text{Boot}}(t_{ij} | x) \right] \quad (3.38) \end{aligned}$$

El valor absoluto que aparece en el primer sumando de la expresión (3.38), se puede escribir como la siguiente suma de variables:

$$\begin{aligned} & \left| \hat{H}_{1h}^{\alpha \text{Boot}}(t_{ik} | x) - \hat{H}_{1h}^{\alpha \text{Boot}}(t_{ij} | x) - E^{\alpha} \hat{H}_{1h}^{\alpha \text{Boot}}(t_{ik} | x) + E^{\alpha} \hat{H}_{1h}^{\alpha \text{Boot}}(t_{ij} | x) \right| \\ & = \sum_{r=1}^h B_{hr}(x) \mathbb{1}_{F_{Z_r}^{\alpha}(t_{ik}; \pm_r = 1g)} - \mathbb{1}_{F_{Z_r}^{\alpha}(t_{ij}; \pm_r = 1g)} \left[\hat{H}_{1g}^{\alpha}(t_{ik} | X_r) + \hat{H}_{1g}^{\alpha}(t_{ij} | X_r) \right] \\ & = \sum_{r=1}^h W_{r(ijk)}(x) \quad (3.39) \end{aligned}$$

En el universo bootstrap, las variables $W_{r(ijk)}(x)$ son i.i.d. y $E^{\alpha} W_{r(ijk)}(x) = 0$. Además

$$\begin{aligned} & \text{Var}^{\alpha} W_{r(ijk)}(x) \\ & = E^{\alpha} \left[\sum_{i=1}^h W_{r(ijk)}(x) \right]^2 = (B_{hr}(x))^2 E \\ & \leq E^{\alpha} \left[\mathbb{1}_{F_{Z_r}^{\alpha}(t_{ik}; \pm_r = 1g)} \hat{H}_{1g}^{\alpha}(t_{ik} | X_r) + \mathbb{1}_{F_{Z_r}^{\alpha}(t_{ij}; \pm_r = 1g)} \hat{H}_{1g}^{\alpha}(t_{ij} | X_r) \right]^2 \end{aligned}$$

$$\begin{aligned}
 &= (B_{hr}(x))^2 \mathbb{E} \left[\sum_i \left(\hat{H}_{1g}^a(t_{ik} | X_r) - \hat{H}_{1g}^a(t_{ij} | X_r) + \hat{H}_{1g}^a(t_{ij} | X_r) - \hat{H}_{1g}^a(t_{ik} | X_r) \right)^2 \right. \\
 &\quad \left. + 2(B_{hr}(x))^2 \sum_i \left(\hat{H}_{1g}^a(\min\{t_{ik}, t_{ij}\} | X_r) - \hat{H}_{1g}^a(t_{ik} | X_r) \hat{H}_{1g}^a(t_{ij} | X_r) \right) \right] \\
 &= (B_{hr}(x))^2 \sum_i \left(\hat{H}_{1g}^a(t_{ik} | X_r) + \hat{H}_{1g}^a(t_{ij} | X_r) - 2\hat{H}_{1g}^a(\min\{t_{ik}, t_{ij}\} | X_r) \right) \\
 &\quad + (B_{hr}(x))^2 \sum_i \left(\hat{H}_{1g}^a(t_{ik} | X_r) + \hat{H}_{1g}^a(t_{ij} | X_r) - 2\hat{H}_{1g}^a(t_{ik} | X_r) \hat{H}_{1g}^a(t_{ij} | X_r) \right)
 \end{aligned}$$

expresión que se acota por

$$\begin{aligned}
 &\cdot (B_{hr}(x))^2 \mathbb{E} \left[\sum_i \left(\hat{H}_{1g}^a(t_{ik} | X_r) - \hat{H}_{1g}^a(\min\{t_{ik}, t_{ij}\} | X_r) + \right. \right. \\
 &\quad \left. \left. + \hat{H}_{1g}^a(t_{ij} | X_r) - \hat{H}_{1g}^a(\min\{t_{ik}, t_{ij}\} | X_r) \right)^2 \right. \\
 &\quad \left. + \sum_i \left(\hat{H}_{1g}^a(t_{ik} | X_r) + \hat{H}_{1g}^a(t_{ij} | X_r) - 2\hat{H}_{1g}^a(t_{ik} | X_r) \hat{H}_{1g}^a(t_{ij} | X_r) \right) \right]
 \end{aligned}$$

y por tanto

$$\begin{aligned}
 \text{Var}^a W_{r(ij)k}(x) &\cdot 2(B_{hr}(x))^2 \max_i \left(\hat{H}_{1g}^a(t_{i+1} | X_r) - \hat{H}_{1g}^a(t_i | X_r) \right) \\
 &\cdot 2(B_{hr}(x))^2 \max_i \left(\hat{H}_{1g}^a(t_{i+1} | X_r) - \hat{H}_{1g}^a(t_i | X_r) \right) \\
 &\quad + H_1^a(t_{i+1} | X_r) + H_1^a(t_i | X_r) + \\
 &\quad + 2(B_{hr}(x))^2 \max_i \left(H_1^a(t_{i+1} | X_r) - H_1^a(t_i | X_r) \right)
 \end{aligned}$$

Analizamos por separado los dos últimos términos.

Para el primero, utilizando un razonamiento análogo al de la acotación de la expresión (3.27) (puesto que $x \geq 1$ y $h \neq 0$), junto con el lema anterior, se tiene

$$2(B_{hr}(x))^2 \max_i \left(\hat{H}_{1g}^a(t_i | X_r) - \hat{H}_{1g}^a(t_{i+1} | X_r) - H_1^a(t_i | X_r) + H_1^a(t_{i+1} | X_r) \right) \tag{3.40}$$

$$\begin{aligned}
 &\cdot 2(B_{hr}(x))^2 \mathbb{E} \left[\sup_{x \geq 1} \sup_{h: x+h \leq 1} \sup_{f \in \mathcal{F}_n} \sup_{t \in [a, b]} \left| \hat{H}_{1g}^a(s | \tilde{X}) - \hat{H}_{1g}^a(t | \tilde{X}) - H_1^a(s | \tilde{X}) + H_1^a(t | \tilde{X}) \right| \right] \\
 &= (B_{hr}(x))^2 O \left(\frac{\ln n}{nh} \right) \quad \text{c.s.}
 \end{aligned}$$

Para el segundo:

$$\begin{aligned}
 & 2(B_{hr}(x))^2 \max_i |H_1^a(t_i | X_r) - H_1^a(t_{i+1} | X_r)| \\
 & \cdot 2(B_{hr}(x))^2 \sup_{x \in (x_i - h, x_i + h)} \max_i |H_1^a(t_i | \tilde{x}) - H_1^a(t_{i+1} | \tilde{x})| \\
 & \cdot 2(B_{hr}(x))^2 C(H_1^{a0}) \max_i |t_{i+1} - t_i| = (B_{hr}(x))^2 O(b_n) \\
 & = (B_{hr}(x))^2 O\left(\frac{\tilde{A} \mu_{\ln n}^{11=2}}{nh}\right)
 \end{aligned}$$

donde $C(H_1^{a0})$ denota una cota de dicha función en $[a; b] \in I_{\pm}$. Por tanto:

$$\begin{aligned}
 \sum_{r=1}^{\infty} \text{Var}^a W_{r(ijk)}(x) & \cdot O\left(\frac{\tilde{A} \mu_{\ln n}^{13=4}}{nh}\right) + O\left(\frac{\tilde{A} \mu_{\ln n}^{11=2}}{nh}\right) \sum_{r=1}^{\infty} (B_{hr}(x))^2 \\
 & \cdot O\left(\frac{\tilde{A} \mu_{\ln n}^{11=2}}{nh} \frac{1}{nh}\right) \quad \text{c.s.}
 \end{aligned}$$

También se verifica que $W_{r(ijk)}(x)$ se acota de la forma:

$$\begin{aligned}
 & |B_{hr}(x) \mathbb{1}_{F_r^a(t_{ik}; \pm_r=1g)} - \mathbb{1}_{F_r^a(t_{ij}; \pm_r=1g)}| \leq |H_{1g}^a(t_{ik} | X_r) - H_{1g}^a(t_{ij} | X_r)| \\
 & \cdot 2 \max_r B_{hr}(x) = O\left(\frac{1}{nh}\right) \quad \text{c.s.}
 \end{aligned}$$

La desigualdad de Bernstein en este contexto nos lleva a que:

$$P^a \left(\sum_{r=1}^{\infty} W_{r(ijk)}(x) > \frac{\tilde{A}}{2} \right) \leq 2 \exp \left(- \frac{\tilde{A}^2}{2C_1 \frac{1}{nh} \frac{\ln n}{nh}^{1=2} + \frac{2}{3} C_2 \frac{1}{nh}} \right)$$

Tomando $\frac{\tilde{A}}{2} = c^a \frac{\ln n}{nh}^{3=4}$ obtenemos:

$$\begin{aligned}
 P^a \left(\sum_{r=1}^{\infty} W_{r(ijk)}(x) > \frac{\tilde{A}}{2} \right) & \leq 2 \exp \left(- \frac{\tilde{A}^2}{2C_1 \frac{\ln n}{nh}^{1=2} + \frac{2C_2 c^a}{3} \frac{\ln n}{nh}^{3=4}} \right) \\
 & \leq 2 \exp \left(- \frac{c^{a2} \frac{\ln n}{nh}^{3=2}}{2(C_1 + C_2 c^a) \frac{\ln n}{nh}^{1=2}} \right) \\
 & = 2n^i \frac{c^{a2}}{2(c_1 + c_2 c^a)}
 \end{aligned}$$

sucesión que converge a cero, tomando c^a de forma apropiada. Por tanto, se tiene:

$$\sum_{r=1}^{\infty} W_{r(ijk)}(x) = O_P^a \left(\frac{\tilde{A} \mu_{\ln n} \mathbb{1}_{3=4}}{nh} \right) \quad c:s:(P)$$

Es más, también se verifica que:

$$\begin{aligned} P^a \max_i \max_{j;k} \sum_{r=1}^{\infty} W_{r(ijk)}(x) &> \tilde{A} \\ \cdot (m_i - 1)(2a_n + 1)^2 P^a \sum_{r=1}^{\infty} W_{r(ijk)}(x) &> \tilde{A} \\ &= 2(m_i - 1)(2a_n + 1)^2 n^i \frac{c^{a2}}{2(c_1 + c_2 c^a)} = \\ &= 2C \frac{\mu_{\ln n} \mathbb{1}_{i=1}}{nh} n^i \frac{c^{a2}}{2(c_1 + c_2 c^a)} \end{aligned} \quad (3.41)$$

sucesión que sigue tendiendo a cero tomando c^a de forma apropiada. Por tanto, se llega al resultado:

$$\max_i \max_{j;k} \sum_{r=1}^{\infty} W_{r(ijk)}(x) = O_P^a \left(\frac{\tilde{A} \mu_{\ln n} \mathbb{1}_{3=4}}{nh} \right) \quad c:s:(P) \quad (3.42)$$

De esta forma hemos acotado el primer sumando de (3.38). A continuación, estudiamos el segundo sumando, dado por:

$$\begin{aligned} &2 \max_i \max_j \left[E^a \hat{H}_{1h}^{aBoot}(t_{ij+1} | j, X) \mid E^a \hat{H}_{1h}^{aBoot}(t_{ij} | j, X) \right] \\ &= 2 \max_i \max_j \sum_{r=1}^{\infty} B_{hr}(x) \left[\hat{H}_{1g}^a(t_{ij+1} | j, X_r) \mid \hat{H}_{1g}^a(t_{ij} | j, X_r) \right] \\ &\cdot 2 \max_i \max_j \sum_{r=1}^{\infty} B_{hr}(x) E \\ &E \left[\hat{H}_{1g}^a(t_{ij+1} | j, X_r) \mid \hat{H}_{1g}^a(t_{ij} | j, X_r) \mid H_1^a(t_{ij+1} | j, X_r) + H_1^a(t_{ij} | j, X_r) \right] + \\ &+ 2 \max_i \max_j \sum_{r=1}^{\infty} B_{hr}(x) \left(H_1^a(t_{ij+1} | j, X_r) \mid H_1^a(t_{ij} | j, X_r) \right) \end{aligned}$$

El primer sumando del segundo miembro de la desigualdad anterior se acota (de forma análoga a (3.40)) por

$$\begin{aligned}
 & \frac{\tilde{A}_n}{2} \sum_{r=1}^{\infty} E \sup_{x \in \mathcal{X}} |B_{hr}(x)| \\
 & \leq \frac{\tilde{A}_n}{2} \sup_{x \in \mathcal{X}} \sup_{f \in \mathcal{F}_{[a,b]}(x)} \sup_{t_j \in \mathcal{T}_n} |\hat{H}_{1g}^{\alpha}(s_j \tilde{x}) - \hat{H}_{1g}^{\alpha}(t_j \tilde{x}) - H_1^{\alpha}(s_j \tilde{x}) + H_1^{\alpha}(t_j \tilde{x})| \\
 & = O\left(\frac{\ln n}{nh}\right) \quad \text{c.s.}
 \end{aligned}$$

El segundo por

$$\begin{aligned}
 & \frac{\tilde{A}_n}{2} \sum_{r=1}^{\infty} E \sup_{x \in \mathcal{X}} |B_{hr}(x)| \\
 & \leq \frac{\tilde{A}_n}{2} \sup_{x \in \mathcal{X}} \max_i \max_j |H_1^{\alpha}(t_{i+1} \tilde{x}) - H_1^{\alpha}(t_{ij} \tilde{x})| \\
 & \cdot C(H_1^{\alpha}) \max_i \max_j |t_{i+1} - t_{ij}| = O\left(\frac{\ln n}{nh}\right)
 \end{aligned}$$

y ambos resultados conducen al orden del segundo sumando en (3.38):

$$2 \max_i \max_j |E^{\alpha} \hat{H}_{1h}^{\alpha \text{Boot}}(t_{i+1} \tilde{x}) - E^{\alpha} \hat{H}_{1h}^{\alpha \text{Boot}}(t_{ij} \tilde{x})| = O\left(\frac{\ln n}{nh}\right) \quad \text{c.s.} \tag{3.43}$$

Entonces en virtud de (3.38), (3.39), (3.42) y (3.43) se tiene que:

$$\sup_{f \in \mathcal{F}_{[a,b]}(x)} (3.36) = O_P\left(\frac{\ln n}{nh}\right) \quad \text{c.s.}(P)$$

Ahora, necesitamos calcular el $\sup_{x \in \mathcal{X}} S_1(x)$ y que claramente coincide con:

$$\begin{aligned}
 & \sup_{x \in \mathcal{X}} \max_i \max_{j,k} |\hat{H}_{1h}^{\alpha \text{Boot}}(t_{ik} \tilde{x}) - \hat{H}_{1h}^{\alpha \text{Boot}}(t_{ij} \tilde{x})| \\
 & \leq E^{\alpha} \hat{H}_{1h}^{\alpha \text{Boot}}(t_{ik} \tilde{x}) + E^{\alpha} \hat{H}_{1h}^{\alpha \text{Boot}}(t_{ij} \tilde{x}) \\
 & + 2 \sup_{x \in \mathcal{X}} \max_i \max_j |E^{\alpha} \hat{H}_{1h}^{\alpha \text{Boot}}(t_{i+1} \tilde{x}) - E^{\alpha} \hat{H}_{1h}^{\alpha \text{Boot}}(t_{ij} \tilde{x})| \\
 & = \sup_{x \in \mathcal{X}} S_1(x) + \sup_{x \in \mathcal{X}} S_2(x):
 \end{aligned}$$

Es inmediato comprobar que

$$\sup_{x \in I} S_{12}(x) = O_{P^*} \left(\frac{\mu_{\ln n}^{(3)}!}{nh} \right) \quad c:s: \quad (3.44)$$

Para estudiar $\sup_{x \in I} S_{11}(x)$ utilizaremos un razonamiento análogo al de la demostración de (3.14).

Por tanto, consideramos en el intervalo I un conjunto E_n de, a lo sumo $n^2 + 1$ puntos, $E_n = \{x_{n1}, \dots, x_{nn}\}$, de forma que $x_{ni} \in I$; $x_{ni} \in E_n = \{x_{nj} : x_{nj} \in I\}$. $C_n = n^2$: Se verifica:

$$\begin{aligned} \sup_{x \in I} S_{11}(x) &= \max_{x_{ni} \in E_n} S_{11}(x_{ni}) + \\ &+ \sup_{x \in I} \max_i \max_{j:k} \left[\hat{H}_{1h}^{*Boot}(t_{ik} | j, x) - \hat{H}_{1h}^{*Boot}(t_{ij} | j, x) \right] \quad (3.45) \\ &= E^* \hat{H}_{1h}^{*Boot}(t_{ik} | j, x) + E^* \hat{H}_{1h}^{*Boot}(t_{ij} | j, x) \\ &+ \hat{H}_{1h}^{*Boot}(t_{ik} | j, x_n) + \hat{H}_{1h}^{*Boot}(t_{ij} | j, x_n) + \\ &+ E^* \hat{H}_{1h}^{*Boot}(t_{ik} | j, x_n) - E^* \hat{H}_{1h}^{*Boot}(t_{ij} | j, x_n) \end{aligned}$$

Utilizando (3.41) obtenemos:

$$\begin{aligned} P^* \left(\max_{x_{ni} \in E_n} S_{11}(x_{ni}) > c^* \frac{\mu_{\ln n}^{(3)}!}{nh} \right) &\leq P^* \left(\max_{x_{ni} \in E_n} S_{11}(x_{ni}) > c^* \frac{\mu_{\ln n}^{(3)}!}{nh} \right) \\ &\leq \text{card}(E_n) 2C \frac{\mu_{\ln n}^{(3)}!}{nh} n^{\frac{c^* 2}{2(c_1 + c_2 c^*)}} \end{aligned}$$

que para $\text{card}(E_n) = n^2 + 1$ es sumable tomando un c^* conveniente. Así llegamos a

$$\max_{x_{ni} \in E_n} S_{11}(x_{ni}) = O_{P^*} \left(\frac{\mu_{\ln n}^{(3)}!}{nh} \right) \quad c:s:(P)$$

Por otra parte,

$$(3.45) = \sup_{x \in I} \max_i \max_{j:k} \sum_{r=1}^{\infty} [B_{hr}(x) - B_{hr}(x_n)] A_{r(ij,k)}$$

donde

$$A_{r(ij,k)} = 1_{F_{Z_r}^*} \cdot 1_{t_{ik} \pm \frac{1}{r} = 1g} - 1_{F_{Z_r}^*} \cdot 1_{t_{ij} \pm \frac{1}{r} = 1g} + \hat{H}_{1g}^*(t_{ik} | j, X_r) + \hat{H}_{1g}^*(t_{ij} | j, X_r):$$

Mediante cálculos análogos a los efectuados en (3.20) acotamos (3.45) de la forma

$$\sup_{x \in I} \max_i \max_{j:k} \frac{1}{\hat{m}_h^a(x) \hat{m}_h^a(\hat{x}_n)} \left| \hat{m}_h^a(\hat{x}_n) - \hat{m}_h^a(x) \right| \sum_{r=1}^{\infty} \frac{1}{nh} K \left(\frac{\hat{x}_n - x}{h} \right) A_{r(ijk)} +$$

$$+ \sup_{x \in I} \max_i \max_{j:k} \frac{2}{\hat{m}_h^a(x)} \sum_{r=1}^{\infty} \frac{1}{nh} K \left(\frac{x_i - x_r}{h} \right) \left| \frac{1}{nh} K \left(\frac{\hat{x}_n - x_r}{h} \right) \right|$$

El segundo sumando es una $O\left(\frac{1}{n^2 h^2}\right)$ c.s: como ya se vió en (3.24). En cuanto al primer sumando se tiene

$$\sup_{x \in I} \max_i \max_{j:k} \frac{1}{\hat{m}_h^a(x) \hat{m}_h^a(\hat{x}_n)} \left| \hat{m}_h^a(\hat{x}_n) - \hat{m}_h^a(x) \right| \sum_{r=1}^{\infty} \frac{1}{nh} K \left(\frac{\hat{x}_n - x}{h} \right) A_{r(ijk)} =$$

$$\cdot \sup_{x \in I} \frac{|\hat{m}_h^a(\hat{x}_n) - \hat{m}_h^a(x)|}{\hat{m}_h^a(x) \hat{m}_h^a(\hat{x}_n)} \max_{\hat{x}_n \in E_n} \max_i \max_{j:k} \sum_{r=1}^{\infty} B_{hr}(\hat{x}_n) A_{r(ijk)}$$

donde fácilmente se ve que

$$\sup_{x \in I} \frac{|\hat{m}_h^a(\hat{x}_n) - \hat{m}_h^a(x)|}{\hat{m}_h^a(x) \hat{m}_h^a(\hat{x}_n)} = O\left(\frac{\tilde{A}_\mu \ln n}{nh}\right) \quad \text{c.s:}$$

y

$$\max_{\hat{x}_n \in E_n} \max_i \max_{j:k} \sum_{r=1}^{\infty} B_{hr}(\hat{x}_n) A_{r(ijk)} = O_{P^*} \left(\frac{\tilde{A}_\mu \ln n}{nh}\right) \quad \text{c.s:(P)}$$

sin más que tener en cuenta que $\sum_{r=1}^{\infty} B_{hr}(\hat{x}_n) A_{r(ijk)} = \sum_{r=1}^{\infty} W_{r(ijk)}(\hat{x}_n)$ (ver (3.39)) y que (3.42) se tiene de modo uniforme para todo $x \in I$:

De este modo hemos ...nalizado la prueba de que

$$\sup_{x \in I} S1_1(x) = O_{P^*} \left(\frac{\tilde{A}_\mu \ln n}{nh}\right) \quad \text{c.s:(P)}$$

y éste resultado junto con (3.44) lleva a que

$$\sup_{x \in I} S1(x) = \sup_{x \in I} \sup_{f_s; t_2[a;b]=j_s \cdot t_j \cdot b_n} (3:36) = O_{P^*} \left(\frac{\tilde{A}_\mu \ln n}{nh}\right) \quad \text{c.s:(P):}$$

El siguiente paso en esta demostración es acotar el término (3.37).

Para ello, efectuamos la descomposición siguiente:

$$\begin{aligned} & E^n \hat{H}_{1g}^{n, \text{Boot}}(t | j, x) - E^n \hat{H}_{1g}^{n, \text{Boot}}(s | j, x) - \hat{H}_{1g}^n(t | j, x) + \hat{H}_{1g}^n(s | j, x) \\ &= \sum_{i=1}^n B_{hi}(x) \left(\hat{H}_{1g}^n(t | j, X_i) - \hat{H}_{1g}^n(s | j, X_i) - H_{1g}^n(t | j, X_i) + H_{1g}^n(s | j, X_i) \right) \\ &+ \sum_{i=1}^n B_{hi}(x) \left(H_{1g}^n(t | j, X_i) - H_{1g}^n(s | j, X_i) \right) - H_{1g}^n(t | j, x) + H_{1g}^n(s | j, x) \\ &+ H_{1g}^n(t | j, x) - H_{1g}^n(s | j, x) - \hat{H}_{1g}^n(t | j, x) + \hat{H}_{1g}^n(s | j, x) : \end{aligned}$$

Es fácil ver que $\sup_{x \in \mathcal{X}} \sup_{f \in \mathcal{F}_s; t \in [a, b]; j \in \mathcal{J}_s; t_j \in \mathcal{T}_j; b_{ng}}$ del primer sumando se puede acotar por

$$\sup_{x \in \mathcal{X}} \sup_{x_2 \in \mathcal{X}_2(x; h; x+h)} \sup_{f \in \mathcal{F}_s; t \in [a, b]; j \in \mathcal{J}_s; t_j \in \mathcal{T}_j; b_{ng}} \left| \hat{H}_{1g}^n(s | j, \tilde{x}) - \hat{H}_{1g}^n(t | j, \tilde{x}) - H_{1g}^n(s | j, \tilde{x}) + H_{1g}^n(t | j, \tilde{x}) \right|$$

y aplicando el Lema 3.11 en un cierto intervalo $I_{\pm} = I_{-1/2} \cup I_{\pm 1} \cup I_{\pm}$ (definido como en el estudio de (3.27)), se llega a un orden $O\left(\frac{\ln n}{nh}\right)^{3-4}$ c.s.:

El mismo orden se obtiene para el $\sup_{x \in \mathcal{X}} \sup_{f \in \mathcal{F}_s; t \in [a, b]; j \in \mathcal{J}_s; t_j \in \mathcal{T}_j; b_{ng}}$ del tercer sumando, aplicando directamente el Lema 3.11.

En cuanto al análisis del $\sup_{x \in \mathcal{X}} \sup_{f \in \mathcal{F}_s; t \in [a, b]; j \in \mathcal{J}_s; t_j \in \mathcal{T}_j; b_{ng}}$ para el segundo sumando, basta considerar que $\hat{z}_h(y | j, x) = \sum_{i=1}^n B_{hi}(x) H_{1g}^n(y | j, X_i)$ es un estimador de tipo núcleo con pesos Nadaraya-Watson de la función $z(y | j, x) = H_{1g}^n(y | j, x)$:

Teniendo en cuenta entonces, que las variables Z_i son continuas y positivas, que $H_{1g}^n(y | j, X_i)$ toma valores entre 0 y 1, y que además, las funciones $m^n(x)$ y $E(H_{1g}^n(y | j, X_i) | X_i = x) = H_{1g}^n(y | j, x)$ verifican las hipótesis del Lema 2.2, se puede aplicar dicho lema al estimador $\hat{z}_h(y | j, x)$ y así obtener el orden del sumando en estudio, que resulta ser una $O\left(\frac{\ln n}{nh}\right)^{3-4}$ c.s.:

Finalmente, al haber acotado los respectivos $\sup_{x \in \mathcal{X}} \sup_{f \in \mathcal{F}_s; t \in [a, b]; j \in \mathcal{J}_s; t_j \in \mathcal{T}_j; b_{ng}}$ de las expresiones (3.36) y (3.37) obteniendo, de forma casi segura una $O_{P^n}\left(\frac{\ln n}{nh}\right)^{3-4}$ y una $O\left(\frac{\ln n}{nh}\right)^{3-4}$ respectivamente, hemos demostrado que:

$$(3.35) = O_{P^n}\left(\frac{\ln n}{nh}\right)^{3-4} \quad \text{c.s.} (P)$$

resultado que constituye la primera parte de este lema.

La segunda parte del lema se demuestra de forma totalmente análoga a la primera parte, una vez tenida en cuenta la relación entre las funciones C , H y L , y las de sus respectivos estimadores no paramétricos y bootstrap.

2. Demostración del Teorema 3.6.

a) Tengamos en cuenta que

$$\hat{\Lambda}_h^{\text{Boot}}(y | x) = \frac{\prod_{i=1}^n 1_{F_{Z_i^*} \cdot y g_i^*} B_{hi}(x)}{\prod_{j=1}^n 1_{F_{T_j^*} \cdot Z_i^* \cdot Z_{jg}^*} B_{hj}(x)} = \frac{\int_0^Z y d\hat{H}_{1h}^{\text{Boot}}(t | x)}{\hat{C}_h^{\text{Boot}}(t | x)}$$

$$\hat{\Lambda}_g^{\text{tc}}(y | x) = \frac{\prod_{i=1}^n 1_{F_{Z_i} \cdot y g_i} B_{gi}(x)}{\prod_{j=1}^n 1_{F_{T_j} \cdot Z_i \cdot Z_{jg}} B_{gj}(x)} = \frac{\int_0^Z y d\hat{H}_{1g}^{\text{tc}}(t | x)}{\hat{C}_g(t | x)}$$

y por tanto

$$\begin{aligned} & \frac{\hat{\Lambda}_h^{\text{Boot}}(y | x) \hat{\Lambda}_g^{\text{tc}}(y | x)}{\int_0^Z y d\hat{H}_{1h}^{\text{Boot}}(t | x) \int_0^Z y d\hat{H}_{1g}^{\text{tc}}(t | x)} = \frac{\int_0^Z y d\hat{H}_{1h}^{\text{Boot}}(t | x)}{\hat{C}_h^{\text{Boot}}(t | x)} \int_0^Z y d\hat{H}_{1g}^{\text{tc}}(t | x) + \frac{\int_0^Z y d\hat{H}_{1h}^{\text{tc}}(t | x)}{\hat{C}_h(t | x)} \int_0^Z y d\hat{H}_{1g}^{\text{Boot}}(t | x) \\ & = \int_0^Z y \frac{d\hat{H}_{1h}^{\text{Boot}}(t | x)}{\hat{C}_h^{\text{Boot}}(t | x)} \int_0^Z y \frac{d\hat{H}_{1g}^{\text{tc}}(t | x)}{\hat{C}_g(t | x)} + \int_0^Z y \frac{d\hat{H}_{1h}^{\text{tc}}(t | x)}{\hat{C}_h(t | x)} \int_0^Z y \frac{d\hat{H}_{1g}^{\text{Boot}}(t | x)}{\hat{C}_g(t | x)} = \\ & = \int_0^Z y \frac{d\hat{H}_{1h}^{\text{Boot}}(t | x) \int_0^Z y \frac{d\hat{H}_{1g}^{\text{tc}}(t | x)}{\hat{C}_g(t | x)} + \int_0^Z y \frac{d\hat{H}_{1h}^{\text{tc}}(t | x) \int_0^Z y \frac{d\hat{H}_{1g}^{\text{Boot}}(t | x)}{\hat{C}_g(t | x)}}{\hat{C}_h^{\text{Boot}}(t | x) \hat{C}_g(t | x)} + \\ & + \int_0^Z y \frac{d\hat{H}_{1h}^{\text{tc}}(t | x) \int_0^Z y \frac{d\hat{H}_{1g}^{\text{Boot}}(t | x)}{\hat{C}_g(t | x)}}{\hat{C}_h(t | x) \hat{C}_g(t | x)} = \end{aligned}$$

$$\begin{aligned} & = \int_0^Z y \frac{d\hat{H}_{1h}^{\text{Boot}}(t | x) \int_0^Z y \frac{d\hat{H}_{1g}^{\text{tc}}(t | x)}{\hat{C}_g(t | x)} + \int_0^Z y \frac{d\hat{H}_{1h}^{\text{tc}}(t | x) \int_0^Z y \frac{d\hat{H}_{1g}^{\text{Boot}}(t | x)}{\hat{C}_g(t | x)}}{\hat{C}_h^{\text{Boot}}(t | x) \hat{C}_g(t | x)} + \\ & + \int_0^Z y \frac{d\hat{H}_{1h}^{\text{tc}}(t | x) \int_0^Z y \frac{d\hat{H}_{1g}^{\text{Boot}}(t | x)}{\hat{C}_g(t | x)}}{\hat{C}_h(t | x) \hat{C}_g(t | x)} = \end{aligned}$$

$$\begin{aligned}
&= \int_0^Z \frac{d\hat{H}_{1h}^{\text{Boot}}(t|x) - d\hat{H}_{1g}^{\text{a}}(t|x)}{C(t|x)} \\
&+ \int_0^Z \frac{\tilde{A}}{\hat{C}_h^{\text{Boot}}(t|x)} \left(\frac{1}{\hat{C}_g(t|x)} - \frac{1}{C(t|x)} \right) d\hat{H}_{1h}^{\text{Boot}}(t|x) + \\
&+ \int_0^Z \frac{\tilde{A}}{\hat{C}_g(t|x)} \left(\frac{1}{C(t|x)} - \frac{1}{\hat{C}_g(t|x)} \right) d\hat{H}_{1h}^{\text{Boot}}(t|x) + \\
&+ \int_0^Z \frac{\tilde{A}}{\hat{C}_h^{\text{Boot}}(t|x)} \left(\frac{1}{\hat{C}_g(t|x)} - \frac{1}{C(t|x)} \right) d\hat{H}_{1g}^{\text{a}}(t|x) \\
&+ \int_0^Z \frac{\hat{C}_h^{\text{Boot}}(t|x) - \hat{C}_g(t|x)}{(C(t|x))^2} d\hat{H}_{1g}^{\text{a}}(t|x) + \\
&+ \int_0^Z \frac{\hat{C}_h^{\text{Boot}}(t|x) - \hat{C}_g(t|x)}{(C(t|x))^2} d\hat{H}_{1h}^{\text{a}}(t|x) :
\end{aligned}$$

Es decir,

$$\begin{aligned}
&\frac{\hat{\Lambda}_h^{\text{Boot}}(y|x) - \hat{\Lambda}_g^{\text{tc}}(y|x)}{\int_0^Z \frac{d\hat{H}_{1h}^{\text{Boot}}(t|x) - d\hat{H}_{1g}^{\text{a}}(t|x)}{C(t|x)}} + \int_0^Z \frac{\tilde{A}}{(C(t|x))^2} \left(\hat{C}_h^{\text{Boot}}(t|x) - \hat{C}_g(t|x) \right) d\hat{H}_{1g}^{\text{a}}(t|x) \\
&+ R1 + R2 + R3
\end{aligned} \tag{3.46}$$

donde

$$\begin{aligned}
\int_0^Z \frac{d\hat{H}_{1h}^{\text{Boot}}(t|x) - d\hat{H}_{1g}^{\text{a}}(t|x)}{C(t|x)} &= \sum_{i=1}^n B_{hi}(x) \frac{1_{F_{Z_i^{\text{a}}}(y; z_i^{\text{a}}=1g)}}{C(Z_i^{\text{a}}|x)} - \sum_{i=1}^n B_{gi}(x) \frac{1_{F_{Z_i^{\text{a}}}(y; z_i^{\text{a}}=1g)}}{C(Z_i^{\text{a}}|x)} \\
&= \int_0^Z \frac{\sum_{i=1}^n B_{hi}(x) 1_{F_{Z_i^{\text{a}}}(t; z_i^{\text{a}}=1g)} - \sum_{i=1}^n B_{gi}(x) 1_{F_{Z_i^{\text{a}}}(t; z_i^{\text{a}}=1g)}}{(C(t|x))^2} d\hat{H}_{1g}^{\text{a}}(t|x) \\
R1 &= \int_0^Z \frac{\tilde{A}}{\hat{C}_h^{\text{Boot}}(t|x)} \left(\frac{1}{\hat{C}_g(t|x)} - \frac{1}{C(t|x)} \right) d\hat{H}_{1h}^{\text{Boot}}(t|x) \\
R2 &= \int_0^Z \frac{\tilde{A}}{\hat{C}_g(t|x)} \left(\frac{1}{C(t|x)} - \frac{1}{\hat{C}_g(t|x)} \right) d\hat{H}_{1h}^{\text{Boot}}(t|x)
\end{aligned}$$

$$\begin{aligned}
 R_3 &= \int_0^{\tilde{A}} \frac{1}{\hat{C}_h^{\text{Boot}}(t_j x)} \left(\frac{1}{\hat{C}_g(t_j x)} - \frac{1}{C(t_j x)} \right) d\hat{H}_{1g}^{\alpha}(t_j x) + \\
 &+ \int_0^{\tilde{A}} \frac{\hat{C}_h^{\text{Boot}}(t_j x) - C(t_j x)}{(C(t_j x))^2} \frac{1}{\hat{C}_g(t_j x)} d\hat{H}_1^{\alpha}(t_j x) \\
 &= \int_0^{\tilde{A}} \frac{\hat{C}_h^{\text{Boot}}(t_j x) - C(t_j x)}{(C(t_j x))^2} d\hat{H}_1^{\alpha}(t_j x) + \int_0^{\tilde{A}} \frac{1}{\hat{C}_g(t_j x)} \frac{1}{C(t_j x)} d\hat{H}_{1g}^{\alpha}(t_j x) \\
 &+ \int_0^{\tilde{A}} \frac{\hat{C}_h^{\text{Boot}}(t_j x) - C(t_j x)}{(C(t_j x))^2} \frac{1}{\hat{C}_g(t_j x)} d\hat{H}_{1g}^{\alpha}(t_j x) :
 \end{aligned}$$

Los dos primeros sumandos de (3.46) constituyen el término dominante de la expresión (3.1), dado por:

$$\sum_{i=1}^{\infty} B_{hi}(x) \mathbb{E} \left[\frac{1}{\hat{C}_h^{\text{Boot}}(t_j x)} \left(\frac{1}{\hat{C}_g(t_j x)} - \frac{1}{C(t_j x)} \right) \right] + \sum_{i=1}^{\infty} B_{gi}(x) \mathbb{E} \left[\frac{1}{\hat{C}_g(t_j x)} \frac{1}{C(t_j x)} \right]$$

y los tres últimos su término de error: $R_n^{\text{Boot}}(y | x) = R_1 + R_2 + R_3$:

Para acotar el supremo en $[a; b] \in I$ del término de error, analizaremos cada uno de sus sumandos por separado.

Los términos R_1 , R_2 y el primer sumando de la expresión ...nal de R_3 que llamaremos R_{31} , presentan una forma similar al término $R_{2na}(y | x)$ de...nido en la demostración del Teorema 2.1 y, en virtud de los resultados de los Lemas 3.9, 3.10, 3.11 y 3.12 previamente demostrados, los respectivos supremos en $[a; b] \in I$ se pueden tratar de forma totalmente semejante a $\sup_{[a; b] \in I} |R_{2na}(y | x)|$ y así ser acotados por una $O_{P^*} \left(\frac{1}{nh} \right)^{3-4} c:S(P)$:

En cuanto al $\sup_{[a; b] \in I}$ del segundo sumando de la expresión ...nal de R_3 , supremo que denotaremos por $\sup_{[a; b] \in I} |R_{32}|$; no es difícil ver que se acota por una $O_{P^*} \left(\frac{1}{nh} \right) c:S(P)$; utilizando los Lemas 3.9 y 3.10.

Para ello, procedemos como sigue:

$$\begin{aligned}
 &\sup_{[a; b] \in I} |R_{32}| \cdot \\
 &\cdot \frac{1}{\mu^4} \sup_{[a; b] \in I} \left| \frac{\hat{C}_h^{\text{Boot}}(t_j x) - C(t_j x)}{(C(t_j x))^2} \frac{1}{\hat{C}_g(t_j x)} \right| \sup_{[a; b] \in I} \left| \frac{\hat{C}_h^{\text{Boot}}(t_j x) - C(t_j x)}{(C(t_j x))^2} \right|.
 \end{aligned}$$

$$\begin{aligned}
 & \frac{1}{\mu^4} \sup_{[a;b] \in L} \left| \hat{C}_h^{\text{Boot}}(t_j | x) - \hat{C}_g(t_j | x) \right| \leq \\
 & \leq \sup_{[a;b] \in L} \left| \hat{C}_h^{\text{Boot}}(t_j | x) \hat{C}_g(t_j | x) + \hat{C}_g(t_j | x) C(t_j | x) \right| \\
 & \quad + \left| \hat{C}_g(t_j | x) C(t_j | x) - (C(t_j | x))^2 \right| \\
 & \frac{1}{\mu^4} \sup_{[a;b] \in L} \left| \hat{C}_h^{\text{Boot}}(t_j | x) - \hat{C}_g(t_j | x) \right| \leq \\
 & \leq \sup_{[a;b] \in L} \left| \hat{C}_g(t_j | x) - \hat{C}_h^{\text{Boot}}(t_j | x) \right| C(t_j | x) + C(t_j | x) \left| \hat{C}_g(t_j | x) - C(t_j | x) \right| \\
 & \frac{1}{\mu^4} \sup_{[a;b] \in L} \left| \hat{C}_h^{\text{Boot}}(t_j | x) - \hat{C}_g(t_j | x) \right| \leq \\
 & \leq \sup_{[a;b] \in L} \left| \hat{C}_g(t_j | x) - \hat{C}_h^{\text{Boot}}(t_j | x) + \hat{C}_g(t_j | x) - \hat{C}_g(t_j | x) \right| C(t_j | x) + \\
 & \quad + C(t_j | x) \left| \hat{C}_g(t_j | x) - C(t_j | x) \right| \\
 & = O_{P^n} \left(\frac{\ln n}{nh} \right) \quad c:s:(P)
 \end{aligned}$$

De esta forma se tiene que

$$\sup_{[a;b] \in L} \left| \hat{R}_n^{\text{Boot}}(y_j | x) \right| = O_{P^n} \left(\frac{\tilde{\mu} \ln n}{nh} \right) \quad c:s:(P)$$

con lo que ...naliza la demostración del apartado a) de este teorema.

b) Escribamos $\hat{F}_h^{\text{Boot}}(y_j | x) - \hat{F}_g^{\text{tc}}(y_j | x)$ de la forma siguiente:

$$\begin{aligned}
 & \frac{1}{\hat{F}_g^{\text{tc}}(y_j | x)} - \frac{1}{\hat{F}_h^{\text{Boot}}(y_j | x)} = \\
 & = \frac{1}{\hat{F}_g^{\text{tc}}(y_j | x)} - \frac{1}{\hat{F}_g^{\text{tc}}(y_j | x)} \frac{\hat{F}_h^{\text{Boot}}(y_j | x)}{\hat{F}_g^{\text{tc}}(y_j | x)} = \\
 & = \frac{1}{\hat{F}_g^{\text{tc}}(y_j | x)} \left(1 - \frac{\hat{F}_h^{\text{Boot}}(y_j | x)}{\hat{F}_g^{\text{tc}}(y_j | x)} \right) = \\
 & = \frac{1}{\hat{F}_g^{\text{tc}}(y_j | x)} \left(1 - \frac{e^{\ln(1 - \hat{F}_h^{\text{Boot}}(y_j | x))}}{e^{\ln(1 - \hat{F}_g^{\text{tc}}(y_j | x))}} \right) = \\
 & = \frac{1}{\hat{F}_g^{\text{tc}}(y_j | x)} \left(1 - e^{\ln(1 - \hat{F}_h^{\text{Boot}}(y_j | x)) - \ln(1 - \hat{F}_g^{\text{tc}}(y_j | x))} \right)
 \end{aligned}$$

Aplicando Taylor a $1 - F(y|x)$ en $x = 0$, se tiene

$$\begin{aligned} \hat{F}_h^{\text{Boot}}(y|x) - \hat{F}_g^{\text{tc}}(y|x) &= (1 - \hat{F}_g^{\text{tc}}(y|x)) - \left[(1 - \hat{F}_g^{\text{tc}}(y|x)) \right. \\ &\quad \left. + \frac{1}{2} e^{\#_n} \ln(1 - \hat{F}_g^{\text{tc}}(y|x)) \right] \\ &\quad + \frac{1}{2} e^{\#_n} \ln(1 - \hat{F}_g^{\text{tc}}(y|x)) - \frac{1}{2} e^{\#_n} \ln(1 - \hat{F}_g^{\text{tc}}(y|x)) : \end{aligned}$$

Llamando $A = \ln(1 - \hat{F}_h^{\text{Boot}}(y|x)) - \ln(1 - \hat{F}_g^{\text{tc}}(y|x))$ podemos escribir la expresión anterior como

$$\begin{aligned} & (1 - \hat{F}_g^{\text{tc}}(y|x)) - (1 - \hat{F}_g^{\text{tc}}(y|x)) + A + \frac{1}{2} e^{\#_n} A^2 \\ &= (1 - \hat{F}_g^{\text{tc}}(y|x)) - (1 - \hat{F}_g^{\text{tc}}(y|x)) + A + \frac{1}{2} e^{\#_n} A^2 + \\ & \quad + (1 - \hat{F}_g^{\text{tc}}(y|x)) - (1 - \hat{F}_g^{\text{tc}}(y|x)) - A + \frac{1}{2} e^{\#_n} A^2 \end{aligned}$$

donde $\#_n$ está comprendido entre 0 y $\ln(1 - \hat{F}_h^{\text{Boot}}(y|x)) - \ln(1 - \hat{F}_g^{\text{tc}}(y|x))$:

Analizamos el término A ; que dividimos en tres sumandos

$$\begin{aligned} A &= \ln(1 - \hat{F}_h^{\text{Boot}}(y|x)) - \ln(1 - \hat{F}_g^{\text{tc}}(y|x)) = \quad (3.47) \\ &= \ln(1 - \hat{F}_h^{\text{Boot}}(y|x)) - \ln(1 - \hat{\Lambda}_h^{\text{Boot}}(y|x)) + \\ & \quad + \hat{\Lambda}_h^{\text{Boot}}(y|x) - \ln(1 - \hat{\Lambda}_g^{\text{tc}}(y|x)) + \\ & \quad + \hat{\Lambda}_g^{\text{tc}}(y|x) - \ln(1 - \hat{F}_g^{\text{tc}}(y|x)) : \end{aligned}$$

Y como consecuencia de esta descomposición y de que el segundo sumando es el principal y está desarrollado en el apartado a) de este teorema, obtenemos

$$\begin{aligned} & \hat{F}_h^{\text{Boot}}(y|x) - \hat{F}_g^{\text{tc}}(y|x) = \\ &= (1 - F(y|x)) \sum_{i=1}^n B_{hi}(x) \mathcal{X}(Z_i^{\#}; T_i^{\#}; \pm_i^{\#}; y; x) - \sum_{i=1}^n B_{gi}(x) \mathcal{X}(Z_i; T_i; \pm_i; y; x) \\ & \quad + R_n^{\text{Boot}}(y|x) \end{aligned}$$

donde

$$\begin{aligned}
 & R_n^{0\text{Boot}}(y|x) = \frac{1}{h} \int_{\mathcal{I}_3} \left[(1 - F(y|x)) \ln \frac{1}{1 - \hat{F}_h^{\text{Boot}}(y|x)} - \hat{\Lambda}_h^{\text{Boot}}(y|x) \right] + \\
 & + (1 - F(y|x)) \frac{1}{h} R_n^{\text{Boot}}(y|x) + \frac{1}{h} \int_{\mathcal{I}_3} \left[(1 - F(y|x)) \hat{\Lambda}_g^{\text{tc}}(y|x) + \ln \frac{1}{1 - \hat{F}_g^{\text{tc}}(y|x)} \right] + \\
 & + (1 - F(y|x)) \frac{1}{2} e^{\#n} A^2 + \\
 & + \frac{1}{h} \int_{\mathcal{I}_3} \left[\hat{F}_g^{\text{tc}}(y|x) - (1 - F(y|x)) \right] A \frac{1}{2} e^{\#n} A^2 :
 \end{aligned} \tag{3.48}$$

A continuación acotaremos los tres primeros sumandos que forman el error, prescindiendo de $(1 - F(y|x))$ por ser siempre ≤ 1 :

En primer lugar, tendremos en cuenta que

$$\sup_{[a;b] \in \mathcal{I}} \left| \hat{\Lambda}_g^{\text{tc}}(y|x) + \ln \frac{1}{1 - \hat{F}_g^{\text{tc}}(y|x)} \right| = O\left(\frac{1}{ng}\right) \quad \text{c:s} \tag{3.49}$$

resultado inmediato sin más que reproducir el estudio de la expresión (2.58) del apartado c) del Teorema 2.1 para la ventana g . También es fácil ver, siguiendo paso a paso el desarrollo citado, y utilizando

$$\sup_{[a;b] \in \mathcal{I}} \left| \hat{C}_h^{\text{Boot}}(t|x) - \hat{C}_g(t|x) \right| = O_{P^n} \left(\frac{\tilde{A} \mu_{\ln n} \mathfrak{I}_{1=2}!}{nh} \right) \quad \text{c:s:(P)}$$

que se tiene

$$\sup_{[a;b] \in \mathcal{I}} \left| \ln \frac{1}{1 - \hat{F}_h^{\text{Boot}}(y|x)} - \hat{\Lambda}_h^{\text{Boot}}(y|x) \right| = O_{P^n} \left(\frac{1}{nh} \right) \quad \text{c:s:(P)} \tag{3.50}$$

Por otro lado en el apartado a) de este teorema se probó que

$$\sup_{[a;b] \in \mathcal{I}} \left| R_n^{\text{Boot}}(y|x) \right| = O_{P^n} \left(\frac{\tilde{A} \mu_{\ln n} \mathfrak{I}_{3=4}!}{nh} \right) \quad \text{c:s:(P)} \tag{3.51}$$

Y entonces, el $\sup_{[a;b] \in \mathcal{I}}$ para los tres primeros sumandos de $R_n^{0\text{Boot}}(y|x)$ tiene un orden dominante de $O_{P^n} \left(\frac{\ln n}{nh} \right) \mathfrak{C}_{3=4} \quad \text{c:s:(P)}$:

En cuanto al cuarto y quinto sumandos, observemos que:

i) El valor de $e^{\#n}$ está comprendido entre 1 y $e^{\ln(1 - \hat{F}_h^{\text{Boot}}(y|x)) - \ln(1 - \hat{F}_g^{\text{tc}}(y|x))}$; esto es

$$e^{\#n} \cdot \frac{1 - \hat{F}_h^{\text{Boot}}(y|x)}{1 - \hat{F}_g^{\text{tc}}(y|x)} \cdot \frac{1}{1 - \hat{F}_g^{\text{tc}}(y|x)} \quad \text{c.s.}$$

cantidad que está acotada en $[a; b] \in I$ de forma c.s.; al estarlo $\frac{1}{1 - \hat{F}_g^{\text{tc}}(y|x)}$ y

$$\sup_{[a;b] \in I} \frac{|\hat{F}_g^{\text{tc}}(y|x) - F(y|x)|}{F(y|x)} = O\left(\frac{\tilde{A} \mu_{1=2}^{\ln n}}{nh}\right) \quad \text{c.s.}$$

(el Teorema 2.2 se sigue verificando para $\hat{F}_g^{\text{tc}}(y|x)$ con las condiciones sobre h y g).

ii) Respecto a A (expresión (3.47)), teniendo en cuenta la representación del apartado a) de este teorema y los resultados (3.49), (3.50) y (3.51), se tiene:

$$\begin{aligned} \sup_{[a;b] \in I} |A_j| &= \sup_{[a;b] \in I} \left| \sum_{i=1}^n B_{hi}(x) \mathbb{1}_{\{Z_i^{\#} \leq T_i^{\#} \pm \epsilon_i^{\#}; y; x\}} - \sum_{i=1}^n B_{gi}(x) \mathbb{1}_{\{Z_i \leq T_i \pm \epsilon_i; y; x\}} \right| + \\ &+ O_{P^{\#}}\left(\frac{\tilde{A} \mu_{3=4}^{\ln n}}{nh}\right) + O\left(\frac{\mu_{1=1}}{ng}\right) + O_{P^{\#}}\left(\frac{\mu_{i=1}^{\ln n}}{nh}\right) \quad \text{c.s.}(P); \end{aligned}$$

Inmediatamente veremos que

$$\begin{aligned} &\sup_{[a;b] \in I} \left| \sum_{i=1}^n B_{hi}(x) \mathbb{1}_{\{Z_i^{\#} \leq T_i^{\#} \pm \epsilon_i^{\#}; y; x\}} - \sum_{i=1}^n B_{gi}(x) \mathbb{1}_{\{Z_i \leq T_i \pm \epsilon_i; y; x\}} \right| \quad (3.52) \\ &= O_{P^{\#}}\left(\frac{\tilde{A} \mu_{1=2}^{\ln n}}{nh}\right) \quad \text{c.s.}(P) \end{aligned}$$

y por tanto, se verificará que

$$\sup_{[a;b] \in I} |A_j| = O_{P^{\#}}\left(\frac{\tilde{A} \mu_{1=2}^{\ln n}}{nh}\right) \quad \text{c.s.}(P);$$

Como consecuencia, los $\sup_{[a;b] \in I}$ de los sumandos cuarto y quinto en (3.48) se acotan, respectivamente, por

$$\sup_{[a;b] \in I} \left| \frac{1}{1 - \hat{F}_g^{\text{tc}}(y|x)} - \frac{1}{1 - F(y|x)} \right| = O_{P^{\#}}\left(\frac{\mu_{1=2}^{\ln n}}{nh}\right) \quad \text{c.s.}(P)$$

y

$$\begin{aligned}
 & \sup_{[a;b] \in I} \left| \frac{1}{n} \sum_{i=1}^n \left(\hat{F}_g^{tc}(y_j | x_i) - (1 - F(y_j | x_i)) \right) \right| A_i \frac{1}{2} e^{\#n} A^2 \\
 & \cdot \sup_{[a;b] \in I} \left| \frac{\hat{F}_g^{tc}(y_j | x_i) - F(y_j | x_i)}{A \mu_{\ln n}^{11=2}} - \frac{\tilde{A} \mu_{\ln n}^{11=2}}{A \mu_{\ln n}^{11=2}} \right| \sup_{[a;b] \in I} A_i \frac{1}{2} e^{\#n} A^2 = \\
 & = O_{P^n} \frac{\ln n}{nh} = O_{P^n} \frac{\ln n}{nh} \quad c.s.(P):
 \end{aligned}$$

Este orden junto con los obtenidos para los $\sup_{[a;b] \in I}$ de los sumandos anteriores de (3.48) prueban, ...nalmente, que

$$\sup_{[a;b] \in I} \left| R_n^{0Boot}(y_j | x) \right| = O_{P^n} \frac{\tilde{A} \mu_{\ln n}^{13=4}}{nh} \quad c.s.(P):$$

Analizamos, pues (3.52). Para ello escribimos:

$$\begin{aligned}
 & \int_0^{\infty} B_{hi}(x) \mathbb{E} \left(Z_i^{\#}; T_i^{\#}; \pm_i^{\#}; y; x \right) \int_0^{\infty} B_{gi}(x) \mathbb{E} \left(Z_i; T_i; \pm_i; y; x \right) \\
 & = \int_0^{\infty} \frac{d\hat{H}_{1h}^{\#Boot}(t_j | x) \int_0^{\infty} d\hat{H}_{1g}^{\#}(t_j | x)}{C(t_j | x)} \int_0^{\infty} \frac{\tilde{A} \hat{C}_h^{Boot}(t_j | x) \int_0^{\infty} \hat{C}_g(t_j | x)}{(C(t_j | x))^2} dH_1^{\#}(t_j | x) \\
 & = \int_0^{\infty} \frac{d\hat{H}_{1h}^{\#Boot}(t_j | x) \int_0^{\infty} d\hat{H}_{1g}^{\#}(t_j | x)}{C(t_j | x)} \int_0^{\infty} \frac{\tilde{A} \hat{C}_h^{Boot}(t_j | x) \int_0^{\infty} \hat{C}_g(t_j | x)}{(C(t_j | x))^2} dH_1^{\#}(t_j | x) \\
 & = \frac{\hat{H}_{1h}^{\#Boot}(t_j | x) \int_0^{\infty} \hat{H}_{1g}^{\#}(t_j | x) dC(t_j | x)}{C(t_j | x)} + \int_0^{\infty} \frac{\hat{H}_{1h}^{\#Boot}(t_j | x) \int_0^{\infty} \hat{H}_{1g}^{\#}(t_j | x) dC(t_j | x)}{(C(t_j | x))^2} \\
 & \quad + \int_0^{\infty} \frac{\tilde{A} \hat{C}_h^{Boot}(t_j | x) \int_0^{\infty} \hat{C}_g(t_j | x)}{(C(t_j | x))^2} dH_1^{\#}(t_j | x)
 \end{aligned}$$

y entonces

$$\sup_{[a;b] \in I} \left| \int_0^{\infty} B_{hi}(x) \mathbb{E} \left(Z_i^{\#}; T_i^{\#}; \pm_i^{\#}; y; x \right) \int_0^{\infty} B_{gi}(x) \mathbb{E} \left(Z_i; T_i; \pm_i; y; x \right) \right|$$

$$\begin{aligned}
 & \frac{1}{\mu} + \frac{2}{\mu^2} \sup_{[a,b] \in I} |\hat{H}_{1h}^{\text{Boot}}(y|x) - \hat{H}_{1g}(y|x)| + \\
 & + \frac{1}{\mu^2} \sup_{[a,b] \in I} |\hat{C}_h^{\text{Boot}}(t|x) - \hat{C}_g(t|x)| \\
 = & O_{P^n} \left(\frac{\ln n}{nh} \right) \quad c:s:(P)
 \end{aligned}$$

...nalizando, de esta forma, la presente demostración.

3. Lemas utilizados en la demostración del Teorema 3.7.

Lema 3.13 Bajo las hipótesis H1-H13, para $x \in I$ e $y \in [a,b]$, se veri...ca:

$$\begin{aligned}
 & \sup_{t \in R} P^n \left(\frac{h^3}{(nh)^{1=2}} (1 - F(y|x)) \leq \right. \\
 & \left. \sum_{i=1}^n (B_{hi}(x)) \sum_{j=1}^n (Z_i^n; T_i^n; \pm_i^n; y; x) - B_{gi}(x) \sum_{j=1}^n (Z_i; T_i; \pm_i; y; x) \right) \leq t \\
 & \leq \Phi_N \left(\frac{t}{(v_n^n(y|x))^{1=2}} \right) = O_3 \left((nh)^{i-1=2} \right) \quad c:s:
 \end{aligned}$$

donde $b_n^n(y|x)$ es

$$\begin{aligned}
 & E^n \left((nh)^{1=2} (1 - F(y|x)) \sum_{i=1}^n (B_{hi}(x)) \sum_{j=1}^n (Z_i^n; T_i^n; \pm_i^n; y; x) - B_{gi}(x) \sum_{j=1}^n (Z_i; T_i; \pm_i; y; x) \right) \\
 & = (nh)^{1=2} (1 - F(y|x)) \sum_{i=1}^n (B_{hi}(x)) E^n \sum_{j=1}^n (Z_i^n; T_i^n; \pm_i^n; y; x) - B_{gi}(x) \sum_{j=1}^n (Z_i; T_i; \pm_i; y; x)
 \end{aligned}$$

y $v_n^n(y|x)$ es

$$\begin{aligned}
 & \text{Var}^n \left((nh)^{1=2} (1 - F(y|x)) \sum_{i=1}^n (B_{hi}(x)) \sum_{j=1}^n (Z_i^n; T_i^n; \pm_i^n; y; x) - B_{gi}(x) \sum_{j=1}^n (Z_i; T_i; \pm_i; y; x) \right) \\
 & = nh (1 - F(y|x))^2 \sum_{i=1}^n (B_{hi}(x))^2 \text{Var}^n \sum_{j=1}^n (Z_i^n; T_i^n; \pm_i^n; y; x):
 \end{aligned}$$

Demostración.

Sigue los mismos pasos que la demostración del Lema 2.7. Por ello, aplicaremos la desigualdad de Berry-Esséen a las variables

$$V_i^n = (nh)^{1-2} (1 - F(y_j | x)) B_{hi}(x) [j_{\gg_i}^n(Z_i^n; T_i^n; \pm_i^n; y; x) - E^n j_{\gg_i}^n(Z_i^n; T_i^n; \pm_i^n; y; x)]$$

que abreviadamente escribiremos como

$$V_i^n = (nh)^{1-2} (1 - F(y_j | x)) B_{hi}(x) [j_{\gg_i}^n(y; x) - E^n j_{\gg_i}^n(y; x)]:$$

Así, obtenemos que el supremo del enunciado se acota por

$$\begin{aligned} & \frac{(nh)^{3-2} (1 - F(y_j | x))^3 \sum_{i=1}^n (B_{hi}(x))^3 E^n [j_{\gg_i}^n(y; x) - E^n j_{\gg_i}^n(y; x)]^3}{nh (1 - F(y_j | x))^2 \sum_{i=1}^n (B_{hi}(x))^2 \text{Var}^n j_{\gg_i}^n(y; x)} \\ &= A \frac{\sum_{i=1}^n (B_{hi}(x))^3 E^n [j_{\gg_i}^n(y; x) - E^n j_{\gg_i}^n(y; x)]^3}{\sum_{i=1}^n (B_{hi}(x))^2 \text{Var}^n j_{\gg_i}^n(y; x)}: \end{aligned} \tag{3.53}$$

A continuación veremos que $E^n [j_{\gg_i}^n(y; x) - E^n j_{\gg_i}^n(y; x)]^3$ está acotada de modo uniforme en i y como consecuencia, el numerador de la cota (3.53) tendrá el orden de

$$\sum_{i=1}^n (B_{hi}(x))^3 \cdot \max_i B_{hi}(x) \sum_{i=1}^n B_{hi}(x) = O\left(\frac{1}{nh}\right) \quad \text{c.s.}$$

Estudiamos entonces:

$$\begin{aligned} & E^n [j_{\gg_i}^n(y; x) - E^n j_{\gg_i}^n(y; x)]^3 \\ &= E^n [j_{\gg_i}^n(y; x)]^3 + 3E^n [j_{\gg_i}^n(y; x)]^2 E^n [j_{\gg_i}^n(y; x) - E^n j_{\gg_i}^n(y; x)] + \\ &+ 3E^n [j_{\gg_i}^n(y; x) - E^n j_{\gg_i}^n(y; x)] (E^n [j_{\gg_i}^n(y; x)]^2 + (E^n [j_{\gg_i}^n(y; x) - E^n j_{\gg_i}^n(y; x)])^2) + \\ &+ 4E^n [j_{\gg_i}^n(y; x) - E^n j_{\gg_i}^n(y; x)]^3 + (E^n [j_{\gg_i}^n(y; x) - E^n j_{\gg_i}^n(y; x)])^3 \end{aligned}$$

(en la última desigualdad se ha utilizado que $j_{\gg_i}^n(y; x)$ y $j_{\gg_i}^n(y; x) - E^n j_{\gg_i}^n(y; x)$ tienen covarianza positiva y por tanto $E^n [j_{\gg_i}^n(y; x) - E^n j_{\gg_i}^n(y; x)]^2 \leq E^n [j_{\gg_i}^n(y; x) - E^n j_{\gg_i}^n(y; x)] E^n [j_{\gg_i}^n(y; x)]^2$:

Como por otra parte, se tiene de modo uniforme en i ; que:

$$\begin{aligned} j_{\gg_i}^n(y; x) &= \frac{\int_{-1}^1 1_{F(Z_i^n; T_i^n; \pm_i^n) \leq g} dH_1^n(u | j | x)}{C(Z_i^n | j | x)} \\ &= \frac{\int_0^1 1_{F(\tau_i^n; u; Z_i^n) \leq g} dH_1^n(u | j | x)}{(C(u | j | x))^2} \\ &= \sup_{s_2 \in [0; y]} \frac{1}{C(s | j | x)} + \sup_{s_2 \in [0; y]} \frac{1}{(C(s | j | x))^2} \cdot 2 \sup_{s_2 \in [0; y]} \frac{1}{(C(s | j | x))^2} \cdot \frac{2}{\mu^2} \end{aligned}$$

se llega al resultado deseado de

$$\begin{aligned}
 & E_h^{\mu} \left[\frac{1}{h^3} \sum_{i=1}^3 j_{i-1}^{\mu}(y; x) - E_{\mu}^{\mu} \left(\frac{1}{h^3} \sum_{i=1}^3 j_{i-1}^{\mu}(y; x) \right) \right]^3 \\
 & \leq \frac{4}{h^3} E_{\mu}^{\mu} \left[\sum_{i=1}^3 j_{i-1}^{\mu}(y; x) \right]^3 + \left(E_{\mu}^{\mu} \left[\sum_{i=1}^3 j_{i-1}^{\mu}(y; x) \right] \right)^3 \\
 & \leq \frac{4}{h^3} \frac{\tilde{A} \mu^2 \Gamma_3}{\mu^2} + \frac{\mu^2 \Gamma_3}{\mu^2} = 8 \frac{\mu^2 \Gamma_3}{\mu^2} \quad c:s:
 \end{aligned}$$

Continuamos estudiando el denominador de (3.53) dado por

$$\sum_{i=1}^3 (B_{hi}(x))^2 \text{Var}_{\mu}^{\mu} \left(\frac{1}{h^3} \sum_{i=1}^3 j_{i-1}^{\mu}(y; x) \right)$$

Como consecuencia de los Lemas 3.14 (probado a continuación) y 2.8, se verifica que:

$$\sum_{i=1}^3 (B_{hi}(x))^2 \text{Var}_{\mu}^{\mu} \left(\frac{1}{h^3} \sum_{i=1}^3 j_{i-1}^{\mu}(y; x) \right) = \frac{s^2(y|x)}{(1 - F(y|x))^2 nh} + O_{\tilde{A}} \left(\frac{(\ln n)^{1=2}}{nh (nh)^{1=2}} \right) \quad c:s:$$

lo que, de forma inmediata, demuestra:

$$\sum_{i=1}^3 (B_{hi}(x))^2 \text{Var}_{\mu}^{\mu} \left(\frac{1}{h^3} \sum_{i=1}^3 j_{i-1}^{\mu}(y; x) \right) = O \left(\frac{1}{(nh)^{i-3=2}} \right) \quad c:s:$$

y por tanto la cota en (3.53) tiene orden

$$O \left(\frac{(\ln n)^{1=2}}{(nh)^{i-3=2}} \right) = O \left(\frac{1}{(nh)^{i-1=2}} \right) \quad c:s:$$

Lema 3.14 Bajo las hipótesis H1-H13, para $x \in I$ e $y \in [a; b]$, se tiene que:

$$v_n^{\mu}(y|x) - v_n(y|x) = O \left(\frac{\tilde{A} \mu \Gamma_{1=2}}{nh} \right) \quad c:s:$$

Si además se verifican las hipótesis H14 y H15, entonces

$$b_n^{\mu}(y|x) - b_n(y|x) = O_P \left(\frac{1}{nh^5 g^4} \right)^{1=2} :$$

Demostración.

Para analizar la diferencia entre sesgos utilizaremos el Teorema 3 de Härdle y Marron (1991).

Para aplicar dicho resultado veremos, en primer lugar, que denotando por $\Phi(u) = E(\mathbb{1}_{\{Z; T; \pm; y; x\}} \mid T \cdot Z; X = u)$ y, a su estimador tipo núcleo, por $\hat{\Phi}_n(x) = \frac{1}{n} \sum_{i=1}^n B_{hi}(x) \mathbb{1}_{\{Z_i; T_i; \pm_i; y; x\}}$, las expresiones $b_n(y \mid x)$ y $b_n^a(y \mid x)$ se pueden escribir de la forma

$$\begin{aligned} b_n(y \mid x) &= (nh)^{1-2} (1 - F(y \mid x)) E^{Y \mid X} \sum_{i=1}^n B_{hi}(x) \mathbb{1}_{\{Z_i; T_i; \pm_i; y; x\}} \\ &= (nh)^{1-2} (1 - F(y \mid x)) E^{Y \mid X} \hat{\Phi}_n(x) - \Phi(x) \end{aligned} \tag{3.54}$$

y

$$\begin{aligned} b_n^a(y \mid x) &= (nh)^{1-2} (1 - F(y \mid x)) E \sum_{i=1}^n (B_{hi}(x) E^a \mathbb{1}_{\{Z_i^a; T_i^a; \pm_i^a; y; x\}} - B_{gi}(x) \mathbb{1}_{\{Z_i; T_i; \pm_i; y; x\}}) \\ &= (nh)^{1-2} (1 - F(y \mid x)) \sum_{i=1}^n B_{hi}(x) \hat{\Phi}_g(X_i) - \hat{\Phi}_g(x) \end{aligned} \tag{3.55}$$

respectivamente: (3.54) es inmediato, ya que $\Phi(x) = 0$ (ver (2.60)); (3.55) se obtiene teniendo en cuenta que

$$\begin{aligned} &E^a \mathbb{1}_{\{Z_i^a; T_i^a; \pm_i^a; y; x\}} \\ &= E^a \int_0^y \frac{1_{f_{Z_i^a, y; \pm_i^a=1g}}}{C(Z_i^a \mid x)} \int_0^Z \frac{1_{f_{T_i^a, u, Z_i^a g}}}{(C(u \mid x))^2} dH_1^a(u \mid x) = \\ &= \int_0^Z \frac{d\hat{H}_{1g}^a(u \mid X_i)}{C(u \mid x)} \int_0^Z \frac{\hat{C}_g(u \mid X_i)}{(C(u \mid x))^2} dH_1^a(u \mid x) = \\ &= \sum_{r=1}^X B_{gr}(X_i) \frac{1_{f_{Z_r, y; \pm_r=1g}}}{C(Z_r \mid x)} \int_0^Z \frac{1_{f_{T_r, u, Z_r g}}}{(C(u \mid x))^2} dH_1^a(u \mid x) \\ &= \sum_{r=1}^X B_{gr}(X_i) \mathbb{1}_{\{Z_r; T_r; \pm_r; y; x\}} = \hat{\Phi}_g(X_i) \end{aligned}$$

Por otro lado, se comprueba que las hipótesis H1-H15 implican las del teorema citado de Härdle y Marron, y por ello podemos aplicarlo en este contexto. Así, se establece la convergencia en probabilidad hacia 1 de la expresión

$$\frac{E^{Y \mid X} (nh)^{i-1} (1 - F(y \mid x))^{i-2} (b_n^a(y \mid x) - b_n(y \mid x))^2}{[h^4 (C_1 n^{i-1} g^{i-5} + C_2 g^4)]}$$

donde

$$C_1 = \frac{1}{4m^{\alpha}(x)} d_K^2 \int_{-\infty}^{\infty} V(u) K^{(0)}(u)^2 du$$

y

$$C_2 = \frac{1}{16m^{\alpha}(x)^2} d_K^4 \int_{-\infty}^{\infty} (m^{\alpha}(x)\Phi(x))^{1V} \int_{-\infty}^{\infty} m^{\alpha(0)}(x)\Phi(x) \int_{-\infty}^{\infty} \mathbb{1}_2$$

siendo

$$V(u) = \text{Var}(\mathfrak{N}(Z; T; \pm; y; x) | T \cdot Z; X = u)$$

y

$$d_K = \int_{-\infty}^{\infty} u^2 K(u) du$$

Razonando de modo análogo al de la demostración del Teorema 3.18 de la tesis de Cao Abad (1990), se prueba que

$$b_n^{\alpha}(y | x) - b_n(y | x) = O_P(n^{-1} h^5) C_1 n^{-1} g^{i-5} + C_2 g^{4} \mathbb{C}_{1=2}^{\mathbb{C}};$$

Y teniendo en cuenta las hipótesis sobre las ventanas h y g , es sencillo ver que dicho orden es una $O_P(n^{-1} h^5 g^{4} \mathbb{C}_{1=2}^{\mathbb{C}})$:

Estudiamos ahora la diferencia entre las varianzas.

Con el fin de hacer más manejables las expresiones para dicho estudio, escribiremos de forma abreviada las variables $\mathfrak{N}(Z_i^{\alpha}; T_i^{\alpha}; \pm_i^{\alpha}; y; x)$ y $\mathfrak{N}(Z_i; T_i; \pm_i; y; x)$ denotándolas por $\mathfrak{N}_i^{\alpha}(y; x)$ y $\mathfrak{N}_i(y; x)$, respectivamente. De este modo, se tiene que:

$$\begin{aligned} & v_n^{\alpha}(y | x) - v_n(y | x) \\ &= nh(1 - F(y | x))^2 \sum_{i=1}^{\infty} (B_{hi}(x))^2 h \text{Var} \mathfrak{N}_i^{\alpha}(y; x) - \text{Var} \mathfrak{N}_i(y; x) = \\ &= nh(1 - F(y | x))^2 \sum_{i=1}^{\infty} (B_{hi}(x))^2 h E^{\alpha} \mathfrak{N}_i^{\alpha}(y; x)^2 - E \mathfrak{N}_i(y; x)^2 \end{aligned} \quad (3.56)$$

$$+ nh(1 - F(y | x))^2 \sum_{i=1}^{\infty} B_{hi}(x)^2 E \mathfrak{N}_i(y; x)^2 - E^{\alpha} \mathfrak{N}_i^{\alpha}(y; x)^2 \quad (3.57)$$

Analizamos el primer sumando.

Comenzamos con

$$\begin{aligned}
 & E^n \left(\frac{\tilde{A}_i^n(y; X_i)}{C(Z_i^n | X_i)} \right)^2 \\
 = & E^n \frac{\tilde{A}_i^n(y; X_i)^2}{(C(Z_i^n | X_i))^2} + E^n \int_0^Z \int_y^Z \frac{1_{F_{T_i^n} \cdot u \cdot Z_i^n g}}{(C(u | X_i))^2} dH_1^n(u | X_i) \\
 & + 2 E^n \int_0^Z \int_y^Z \frac{1_{F_{T_i^n} \cdot y; \pm_i^n = 1g}}{C(Z_i^n | X_i)} \int_0^Z \int_y^Z \frac{1_{F_{T_i^n} \cdot u \cdot Z_i^n g}}{(C(u | X_i))^2} dH_1^n(u | X_i) \\
 = & A_i^n + B_i^n + 2C_i^n
 \end{aligned}$$

donde (utilizando integración por partes) se llega a

$$\begin{aligned}
 A_i^n &= E^n \frac{\tilde{A}_i^n(y; X_i)^2}{(C(Z_i^n | X_i))^2} = \int_0^Z \int_y^Z \frac{d\hat{H}_{1g}^n(u | X_i)}{(C(u | X_i))^2} = \\
 &= \frac{\hat{H}_{1g}^n(y | X_i)}{(C(y | X_i))^2} + 2 \int_0^Z \int_y^Z \frac{\hat{H}_{1g}^n(u | X_i)}{(C(u | X_i))^3} dC(u | X_i)
 \end{aligned}$$

Para B_i^n se tiene:

$$\begin{aligned}
 & E^n \int_0^Z \int_y^Z \frac{1_{F_{T_i^n} \cdot u \cdot Z_i^n g}}{(C(u | X_i))^2} dH_1^n(u | X_i) \\
 = & \int_0^Z \int_y^Z \int_0^Z \int_y^Z \frac{1_{F_{T_i^n} \cdot u \cdot Z_i^n g} 1_{F_{T_i^n} \cdot v \cdot Z_i^n g}}{(C(u | X_i))^2 (C(v | X_i))^2} dH_1^n(u | X_i) dH_1^n(v | X_i) \\
 = & \int_0^Z \int_y^Z \int_0^Z \int_y^Z \frac{1_{\hat{H}_g^{tc}(u \wedge v | X_i)} \hat{L}_g^{tc}(u \wedge v | X_i)}{(C(u | X_i))^2 (C(v | X_i))^2} dH_1^n(u | X_i) dH_1^n(v | X_i)
 \end{aligned}$$

donde $\int_0^Z \int_y^Z \int_0^Z \int_y^Z$ se define en (3.58).

Descomponiendo la integral anterior en la suma de dos: una para el recinto donde $u \leq v$ y otra para el correspondiente a $u > v$; se puede escribir B_i^n de la

forma siguiente:

$$\begin{aligned}
 B_i^a &= \int_0^1 \int_0^1 \frac{\hat{L}_g^{tc}(u_j X_i)}{(C(u_j x))^2} \int_0^1 \int_0^1 \frac{\hat{H}_g^{tc}(v_j X_i)}{(C(v_j x))^2} dH_1^a(v_j x) A dH_1^a(u_j x) + \\
 &+ \int_0^1 \int_0^1 \frac{\hat{L}_g^{tc}(v_j X_i)}{(C(v_j x))^2} \int_0^1 \int_0^1 \frac{\hat{H}_g^{tc}(u_j X_i)}{(C(u_j x))^2} dH_1^a(u_j x) A dH_1^a(v_j x) \\
 &= 2 \int_0^1 \int_0^1 \frac{\hat{L}_g^{tc}(u_j X_i)}{(C(u_j x))^2} \int_0^1 \int_0^1 \frac{\hat{H}_g^{tc}(v_j X_i)}{(C(v_j x))^2} dH_1^a(v_j x) A dH_1^a(u_j x):
 \end{aligned}$$

Y, en cuanto a C_i^a ; es

$$\begin{aligned}
 C_i^a &= E^a \int_0^1 \int_0^1 \frac{1_{f_{Z_i^a \cdot y; \pm_i^a=1g}}}{C(Z_i^a j x)} \int_0^1 \int_0^1 \frac{1_{f_{T_i^a \cdot u \cdot Z_i^a g}}}{(C(u_j x))^2} dH_1^a(u_j x) \\
 &= \int_0^1 \int_0^1 \frac{1_{f_{u \cdot Z_i^a \cdot y; \pm_i^a=1g}}}{C(Z_i^a j x)} \int_0^1 \int_0^1 \frac{dH_1^a(u_j x)}{(C(u_j x))^2} \\
 &= \int_0^1 \int_0^1 \frac{\hat{L}_g^{tc}(u_j X_i)}{(C(u_j x))^2} \int_0^1 \int_0^1 \frac{d\hat{H}_{1g}^{tc}(v_j X_i)}{C(v_j x)} \frac{dH_1^a(u_j x)}{(C(u_j x))^2}:
 \end{aligned}$$

Es importante resaltar que el cálculo de las distintas esperanzas bootstrap efectuadas al analizar los términos A_i^a ; B_i^a y C_i^a se realiza de forma análoga al de $E^a 1_{f_{Z_i^a \cdot y; \pm_i^a=1g}}$ (ver (3.5)) y que por $\int_0^1 \int_0^1 (X_i)$ se entiende

$$\int_0^1 \int_0^1 (X_i) = \int_0^1 \int_0^1 \hat{L}_g^{tc}(z_j X_i) d\hat{H}_g^{tc}(z_j X_i): \tag{3.58}$$

De forma paralela al estudio de $E^a (\hat{y}_i^a(y; x))^2$ analizamos $E^{YjX} (\hat{y}_i^a(y; x))^2$, obteniendo que $E^{YjX} (\hat{y}_i^a(y; x))^2$ es igual a

$$\begin{aligned}
 &E^{YjX} \int_0^1 \int_0^1 \frac{1_{f_{Z_i \cdot y; \pm_i=1g}}}{(C(Z_i j x))^2} + E^{YjX} \int_0^1 \int_0^1 \frac{1_{f_{T_i \cdot u \cdot Z_i g}}}{(C(u_j x))^2} dH_1^a(u_j x) \\
 &+ 2 E^{YjX} \int_0^1 \int_0^1 \frac{1_{f_{Z_i \cdot y; \pm_i=1g}}}{C(Z_i j x)} \int_0^1 \int_0^1 \frac{1_{f_{T_i \cdot u \cdot Z_i g}}}{(C(u_j x))^2} dH_1^a(u_j x) \\
 &= A_i + B_i + 2C_i
 \end{aligned}$$

$I_{\pm 1} \approx \frac{1}{2} I_{\pm}$ tal que

$$\begin{aligned}
 & \sup_{[a;b] \in I} \sum_{i=1}^n B_{hi}(x) (A_i^{\alpha} - A_i) \\
 & \cdot \sup_{[a;b] \in I} \sum_{i=1}^n B_{hi}(x) \frac{\hat{H}_{1g}^{\alpha}(y_j | X_i) - H_1^{\alpha}(y_j | X_i)}{(C(y_j | x))^2} + \\
 & + \sup_{[a;b] \in I} \sum_{i=1}^n 2 \int_0^y \frac{\hat{H}_{1g}^{\alpha}(u_j | X_i) - H_1^{\alpha}(u_j | X_i)}{(C(u_j | x))^3} dC(u_j | x) \\
 & \cdot \frac{1}{\mu^2} \sup_{[a;b] \in I_{\pm 1}} \frac{\hat{H}_{1g}^{\alpha}(y_j | \tilde{x}) - H^{\alpha}(y_j | \tilde{x})}{\mu} + \frac{4}{\mu^3} \sup_{[a;b] \in I_{\pm 1}} \frac{\hat{H}_{1g}^{\alpha}(y_j | \tilde{x}) - H^{\alpha}(y_j | \tilde{x})}{\mu^3} \\
 & = O\left(\frac{\ln n}{nh}\right) \quad \text{c.s.} \quad (3.60)
 \end{aligned}$$

En cuanto a B_i^{α} y B_i , utilizando que $\int_0^1 (X_i) L(u_j | X_i) (1 - H(u_j | X_i)) = C(u_j | X_i)$ y $\int_0^1 (X_i) (1 - H_g^{\alpha}(u_j | X_i)) \hat{L}_g^{\alpha}(u_j | X_i) = \hat{C}_g(u_j | X_i)$; se tiene:

$$B_i^{\alpha} - B_i =$$

$$\int_0^y \frac{\hat{C}_g(u_j | X_i)}{(1 - H_g^{\alpha}(u_j | X_i)) (C(u_j | x))^2} - \int_0^y \frac{(1 - H_g^{\alpha}(v_j | X_i))}{(C(v_j | x))^2} dH_1^{\alpha}(v_j | x) - \int_0^y \frac{1}{(C(v_j | x))^2} dH_1^{\alpha}(u_j | x)$$

$$\int_0^y \frac{C(u_j | X_i)}{(1 - H(u_j | X_i)) (C(u_j | x))^2} - \int_0^y \frac{(1 - H(v_j | X_i))}{(C(v_j | x))^2} dH_1^{\alpha}(v_j | x) - \int_0^y \frac{1}{(C(v_j | x))^2} dH_1^{\alpha}(u_j | x)$$

Si de...nimos

$$\begin{aligned}
 \hat{r}_{1g}^{\alpha}(u_j | X_i) &= \int_0^y \frac{(1 - H_g^{\alpha}(v_j | X_i))}{(C(v_j | x))^2} dH_1^{\alpha}(v_j | x) \\
 \hat{r}_1(u_j | X_i) &= \int_0^y \frac{(1 - H(v_j | X_i))}{(C(v_j | x))^2} dH_1^{\alpha}(v_j | x)
 \end{aligned}$$

para $(y; x) \in [a; b] \times I$ y $X_i \in (x_i - h; x_i + h)$ y hacemos tender $n \rightarrow \infty$, podemos razonar como en la acotación de (3.60), de tal modo que

$$\begin{aligned} & \sup_{u \in [a; b]} \sup_{x \in I} \left| \hat{r}_{1g}(u; X_i) - r_{1g}(u; X_i) \right| \\ &= \sup_{u \in [a; b]} \sup_{x \in I} \int_u^y \frac{1 - \hat{H}_g^{\text{tc}}(v; X_i) - (1 - H(v; X_i))}{(C(v; x))^2} dH_1^{\alpha}(v; x) \\ & \leq \frac{1}{\mu^2} \sup_{[a; b] \times I_{\pm 1}} \left| \hat{H}_g^{\text{tc}}(y; \tilde{x}) - H(y; \tilde{x}) \right| = O\left(\frac{\ln n}{nh}\right) \quad \text{c.s.} \end{aligned}$$

donde el orden de convergencia de $\sup_{[a; b] \times I_{\pm 1}} \left| \hat{H}_g^{\text{tc}}(y; \tilde{x}) - H(y; \tilde{x}) \right|$ se obtiene como consecuencia del Teorema 2.2 y de los Lemas 3.9 y 3.11, aplicados al caso particular de sólo truncamiento.

Entonces:

$$\begin{aligned} & \int_{Z_y} \frac{B_i^{\alpha} - B_i}{A_i} \hat{C}_g(u; X_i) \hat{r}_{1g}(u; X_i) \frac{C(u; X_i) - r_{1g}(u; X_i)}{1 - \hat{H}_g^{\text{tc}}(u; X_i)} \frac{dH_1^{\alpha}(u; x)}{(C(u; x))^2} \\ &= \int_{Z_y} \frac{\hat{C}_g(u; X_i) \hat{r}_{1g}(u; X_i) (1 - H(u; X_i))}{C(u; X_i) - r_{1g}(u; X_i)} \frac{dH_1^{\alpha}(u; x)}{1 - \hat{H}_g^{\text{tc}}(u; X_i) (1 - H(u; X_i)) (C(u; x))^2} \\ & \leq \frac{3}{1 - \hat{H}_g^{\text{tc}}(u; X_i) (1 - H(u; X_i)) (C(u; x))^2} \end{aligned}$$

A partir de esta expresión, y teniendo en cuenta que en $(y; x) \in [a; b] \times I$ el denominador está acotado por encima de 0 y que

$$\begin{aligned} & \sup_{[a; b] \times I_{\pm 1}} \left| \hat{r}_{1g}(u; \tilde{x}) - r_{1g}(u; \tilde{x}) \right|; \sup_{[a; b] \times I_{\pm 1}} \left| \hat{H}_g^{\text{tc}}(y; \tilde{x}) - H(y; \tilde{x}) \right| \\ & \text{y} \sup_{[a; b] \times I_{\pm 1}} \left| \hat{C}_g(u; \tilde{x}) - C(u; \tilde{x}) \right| \end{aligned}$$

tienen orden $O\left(\frac{\ln n}{nh}\right)$ c.s., sin más que sumar y restar los términos convenientes, llegamos a que

$$\sup_{[a; b] \times I} \left| B_{hi}(x) (B_i^{\alpha} - B_i) \right| = O\left(\frac{\ln n}{nh}\right) \quad \text{c.s.} \quad (3.61)$$

Analizamos ahora C_{ij} C_i^a :

Nos interesa escribir C_i como

$$C_i = \int_0^Z \int_{\mu}^y L(u|j, X_i) \int_u^Z \frac{dH_1^a(v|j, X_i)}{L(v|j, X_i)C(v|j, X_i)} \frac{dH_1^a(u|j, X_i)}{(C(u|j, X_i))^2}$$

lo cual se prueba utilizando que $dH_1^a(v|j, X_i) = \mathbb{1}(X_i) \int_0^1 L(v|j, X_i) dH_1(v|j, X_i)$:

Por otro lado, teniendo en cuenta la expresión (3.58) junto con la Propiedad 2.61, llegamos a

$$\begin{aligned} & \frac{1}{\mathbb{1}_g(X_i)} \int_u^Z \frac{d\hat{H}_{1g}(v|j, X_i)}{C(v|j, X_i)} \\ &= \frac{\prod_{j=1}^n \frac{1_{f_{u \cdot z_j \cdot yg}}}{C(Z_j|X_i)} \prod_{k=1}^n \frac{B_{gj}(X_i) \hat{L}_g^{tc}(Z_j|X_i)}{B_{gk}(X_i) \hat{L}_g^{tc}(Z_k|X_i)}}{\prod_{j=1}^n \frac{\hat{L}_g^{tc}(Z_j|j, X_i) \prod_{k=1}^n \frac{B_{gj}(X_i) \hat{L}_g^{tc}(Z_j|X_i)}{B_{gk}(X_i) \hat{L}_g^{tc}(Z_k|X_i)}}} \\ &= \prod_{j=1}^n \frac{1_{f_{u \cdot z_j \cdot yg}} \prod_{k=1}^n \frac{B_{gj}(X_i)}{B_{gk}(X_i) \hat{L}_g^{tc}(Z_k|j, X_i)}}{C(Z_j|j, X_i) \hat{L}_g^{tc}(Z_j|j, X_i)} = \int_u^Z \frac{d\hat{H}_{1g}^a(z|j, X_i)}{C(z|j, X_i) \hat{L}_g^{tc}(z|j, X_i)} \end{aligned}$$

y por tanto

$$C_i^a = \int_0^Z \int_{\mu}^y \frac{\hat{L}_g^{tc}(u|j, X_i)}{\hat{L}_g^{tc}(u|j, X_i)} \int_u^Z \frac{d\hat{H}_{1g}^a(z|j, X_i)}{C(z|j, X_i) \hat{L}_g^{tc}(z|j, X_i)} \frac{dH_1^a(u|j, X_i)}{(C(u|j, X_i))^2}$$

Ahora, utilizando nuevamente que $\mathbb{1}_i \int_0^1 (X_i) L(u|j, X_i) (1 - H(u|j, X_i)) = C(u|j, X_i)$ y $\mathbb{1}_g \int_0^1 (X_i) \int_0^1 \hat{H}_g^{tc}(u|j, X_i) \hat{L}_g^{tc}(u|j, X_i) = \hat{C}_g(u|j, X_i)$; se tiene que C_{ij} C_i^a es igual a

$$\begin{aligned} & \int_0^Z \int_{\mu}^y \frac{C(u|j, X_i)}{\mathbb{1}_i \int_0^1 (X_i) (1 - H(u|j, X_i))} \int_u^Z \frac{\mathbb{1}_g \int_0^1 (X_i) (1 - H(v|j, X_i)) dH_1^a(v|j, X_i)}{C(v|j, X_i) C(v|j, X_i)} \frac{dH_1^a(u|j, X_i)}{(C(u|j, X_i))^2} \\ &= \int_0^Z \int_{\mu}^y \frac{\hat{C}_g(u|j, X_i)}{\mathbb{1}_g \int_0^1 (X_i) \int_0^1 \hat{H}_g^{tc}(u|j, X_i)} \int_u^Z \frac{\mathbb{1}_g \int_0^1 (X_i) \int_0^1 \hat{H}_g^{tc}(v|j, X_i) d\hat{H}_{1g}^a(v|j, X_i)}{C(v|j, X_i) \hat{C}_g(v|j, X_i)} \frac{dH_1^a(u|j, X_i)}{(C(u|j, X_i))^2} \\ &= \int_0^Z \int_{\mu}^y \frac{C(u|j, X_i)}{\mathbb{1}_i \int_0^1 (X_i) \int_0^1 \hat{H}_g^{tc}(u|j, X_i)} \int_u^Z \frac{\hat{C}_g(u|j, X_i)}{\mathbb{1}_g \int_0^1 (X_i) \int_0^1 \hat{H}_g^{tc}(u|j, X_i)} \frac{dH_1^a(u|j, X_i)}{(C(u|j, X_i))^2} \end{aligned}$$

donde

$$\hat{v}_2(u | j, X_i) = \int_u^y \frac{(1 - H(v | j, X_i)) dH_1^n(v | j, X_i)}{C(v | j, X_i)C(v | j, x)}$$

y

$$\hat{v}_{2g}(u | j, X_i) = \int_u^y \frac{(1 - \hat{H}_g^{tc}(u | j, X_i)) d\hat{H}_{1g}^n(z | j, X_i)}{C(z | j, x)\hat{C}_g(z | j, X_i)}$$

Teniendo en cuenta que cuando $n \rightarrow \infty$ podemos encontrar cierto intervalo $I_{\pm 1}$ donde

$$\sup_{[a;b] \in I_{\pm 1}} \frac{|\hat{H}_g^{tc}(y | j, \tilde{x}) - H(y | j, \tilde{x})|}{H(y | j, \tilde{x})} = O\left(\frac{\tilde{A}_\mu \ln n}{nh} \mathbb{P}_{1=2}^{\dagger}\right) \quad \text{c:s:}$$

$$\sup_{[a;b] \in I_{\pm 1}} \frac{|\hat{C}_g(u | j, \tilde{x}) - C(u | j, \tilde{x})|}{C(u | j, \tilde{x})} = O\left(\frac{\tilde{A}_\mu \ln n}{nh} \mathbb{P}_{1=2}^{\dagger}\right) \quad \text{c:s:} \quad (3.62)$$

$$\sup_{[a;b] \in I_{\pm 1}} \frac{|\hat{H}_{1g}^n(y | j, \tilde{x}) - H_1^n(y | j, \tilde{x})|}{H_1^n(y | j, \tilde{x})} = O\left(\frac{\tilde{A}_\mu \ln n}{nh} \mathbb{P}_{1=2}^{\dagger}\right) \quad \text{c:s:} \quad (3.63)$$

y que uniformemente en $[a; b] \in I$; $C(y | j, x)$ y $1 - H(y | j, x)$ están por encima de 0, se llega a

$$\sup_{[a;b] \in I} \sum_{i=1}^X B_{hi}(x) (C_i^n - C_i) = O\left(\frac{\tilde{A}_\mu \ln n}{nh} \mathbb{P}_{1=2}^{\dagger}\right) \quad \text{c:s:} \quad (3.64)$$

Finalmente, (3.60), (3.61), (3.64) y (3.59) prueban que el primer sumando de $v_n^n(y | j, x) - v_n(y | j, x)$; dado por (3.56) verifica:

$$(3.56) = O\left(\frac{\tilde{A}_\mu \ln n}{nh} \mathbb{P}_{1=2}^{\dagger}\right) \quad \text{c:s:}$$

En cuanto al segundo sumando, dado por (3.57), se acota de la siguiente forma:

$$(3.57) = \sum_{i=1}^X [O(1)] B_{hi}(x) \int \frac{h^3}{E^{Y|X}} \left[E^{Y|X} \left[\frac{v_{2g}^n(y | j, X_i) - v_{2g}(y | j, X_i)}{E^{Y|X} v_{2g}(y | j, X_i)} \right] + E^{Y|X} \left[\frac{v_{2g}^n(y | j, X_i) - v_{2g}(y | j, X_i)}{E^{Y|X} v_{2g}(y | j, X_i)} \right] \right] \quad \text{c:s:} \quad (3.65)$$

Se tiene que:

$$E^{n, \mu} \gg_i(y; x) = E^n \frac{\mu_{1fZ_i^a \cdot y; \pm_i = 1g}}{C(Z_i^a j x)} \int_0^1 \int_0^1 \frac{E^n 1_{fT_i^a \cdot u \cdot Z_i^a g}}{(C(u j x))^2} dH_1^n(u j x) =$$

$$= \int_0^1 \int_0^1 \frac{d\hat{H}_{1g}^n(u j X_i)}{C(u j x)} \int_0^1 \int_0^1 \frac{\hat{C}_g(u j X_i)}{(C(u j x))^2} dH_1^n(u j x)$$

y

$$E^{YjX} \gg_i(y; x) = E^{YjX} \frac{\mu_{1fZ_i^a \cdot y; \pm_i = 1g}}{C(Z_i^a j x)} \int_0^1 \int_0^1 \frac{E^{YjX} 1_{fT_i^a \cdot u \cdot Z_i^a g}}{(C(u j x))^2} dH_1^n(u j x) =$$

$$= \int_0^1 \int_0^1 \frac{dH_1^n(u j X_i)}{C(u j x)} \int_0^1 \int_0^1 \frac{C(u j X_i)}{(C(u j x))^2} dH_1^n(u j x):$$

Por tanto

$$E^{YjX} \gg_i(y; x) - E^{n, \mu} \gg_i(y; x) = \int_0^1 \int_0^1 \frac{dH_1^n(u j X_i) - d\hat{H}_{1g}^n(u j X_i)}{C(u j x)} \int_0^1 \int_0^1 \frac{C(u j X_i) - \hat{C}_g(u j X_i)}{(C(u j x))^2} dH_1^n(u j x)$$

Utilizando integración por partes, teniendo en cuenta (3.62) y (3.63) y razonando como en el estudio de $\sup_{[a;b] \in I} \sum_{i=1}^n B_{hi}(x) j A_i^a j A_i^j$; llegamos a

$$\sup_{[a;b] \in I} \sum_{i=1}^n B_{hi}(x) \cdot E^{YjX} \gg_i(y; x) - E^{n, \mu} \gg_i(y; x) = O \left(\frac{\tilde{A} \mu_{\ln n}}{nh} \right) \quad c:s:$$

Finalmente, usando este último resultado, junto con el hecho de que la suma $E^{YjX} \gg_i(y; x) + E^{n, \mu} \gg_i(y; x)$ es uniformemente acotada en $[a; b] \in I$ y (3.65) se obtiene:

$$(3:57) = O \left(\frac{\tilde{A} \mu_{\ln n}}{nh} \right) \quad c:s:$$

De este modo, hemos probado que

$$v_n^a(y j x) - v_n(y j x) = O \left(\frac{\tilde{A} \mu_{\ln n}}{nh} \right) \quad c:s:$$

Lema 3.15 Bajo las hipótesis H1-H15, para $x \in I$ e $y \in [a; b]$, se verifica que

$$\sup_{t \in \mathbb{R}} \frac{h}{(nh)^{1/2}} \left(\hat{F}_n^{Boot}(y j x) - \hat{F}_g^{tc}(y j x) \right) \cdot t \in \Phi_N \left(\frac{\mu_{t_i b(y j x)}}{s(y j x)} \right) \quad c:s:$$

Demostración. Utilizando la desigualdad triangular obtenemos:

$$\begin{aligned} & \sup_{t \in \mathbb{R}} P^n \left[(nh)^{1-2} \hat{F}_h^{\text{Boot}}(y_j, x) - \hat{F}_g^{\text{tc}}(y_j, x) \right] \cdot t_i \Phi_N \frac{\mu_{t_i} b(y_j, x)}{s(y_j, x)} \\ & + \sup_{t \in \mathbb{R}} P^n \left[(nh)^{1-2} \hat{F}_h^{\text{Boot}}(y_j, x) - \hat{F}_g^{\text{tc}}(y_j, x) \right] \cdot t_i \Phi_N \frac{\tilde{A} \mu_{t_i} b_n^a(y_j, x)}{(v_n^a(y_j, x))^{1-2}} \\ & + \sup_{t \in \mathbb{R}} \Phi_N \frac{\tilde{A} \mu_{t_i} b_n^a(y_j, x)}{(v_n^a(y_j, x))^{1-2}} - \Phi_N \frac{\mu_{t_i} b(y_j, x)}{s(y_j, x)} \end{aligned} \quad (3.66)$$

Comenzamos estudiando el segundo sumando del segundo término de la desigualdad (3.66).

Para ello utilizamos el Lema 2.6 que nos lleva a

$$\begin{aligned} & \sup_{t \in \mathbb{R}} \Phi_N \frac{\tilde{A} \mu_{t_i} b_n^a(y_j, x)}{(v_n^a(y_j, x))^{1-2}} - \Phi_N \frac{\mu_{t_i} b(y_j, x)}{s(y_j, x)} \\ & = O(b_n^a(y_j, x) - b(y_j, x)) + O\left(\frac{1}{(v_n^a(y_j, x))^{1-2}} - \frac{1}{s(y_j, x)}\right) \quad \text{c.s.} \end{aligned}$$

Combinando los Lemas 2.8 y 3.14 obtenemos que

$$b_n^a(y_j, x) - b(y_j, x) = o_P(1); \quad v_n^a(y_j, x) - s^2(y_j, x) = o(1) \quad \text{c.s.}$$

y por lo tanto

$$\sup_{t \in \mathbb{R}} \Phi_N \frac{\tilde{A} \mu_{t_i} b_n^a(y_j, x)}{(v_n^a(y_j, x))^{1-2}} - \Phi_N \frac{\mu_{t_i} b(y_j, x)}{s(y_j, x)} = o_P(1) \quad (3.67)$$

Estudiamos ahora, el primer sumando del segundo término de la desigualdad (3.66). Teniendo en cuenta la representación casi segura bootstrap dada en el Teorema 3.6 y el esquema de la acotación del término

$$\sup_{t \in \mathbb{R}} P^n \left[(nh)^{1-2} \hat{F}_h(y_j, x) - F(y_j, x) \right] \cdot t_i \Phi_N \frac{\mu_{t_i} b(y_j, x)}{s(y_j, x)}$$

en la demostración del Teorema 2.5, obtenemos:

$$\begin{aligned} & \sup_{t \in \mathbb{R}} P^n \left[(nh)^{1-2} \hat{F}_h^{\text{Boot}}(y_j, x) - \hat{F}_g^{\text{tc}}(y_j, x) \right] \cdot t_i \Phi_N \frac{\tilde{A} \mu_{t_i} b_n^a(y_j, x)}{(v_n^a(y_j, x))^{1-2}} \\ & \leq \sup_{t \in \mathbb{R}} P^n \left[(nh)^{1-2} (1 - F(y_j, x)) \varepsilon \right] \\ & \leq \sum_{i=1}^{\infty} (B_{hi}(x) \mathbb{1}_{\{Z_i^a \leq T_i^a \leq \pm_i^a; y; x\}} + B_{gi}(x) \mathbb{1}_{\{Z_i \leq T_i \leq \pm_i; y; x\}}) \cdot t_i + \frac{1}{n} \end{aligned}$$

$$\sup_{t \in \mathbb{R}} \sum_{i=1}^n \Phi_N \frac{\tilde{A}_i(t + \frac{b_n^a(y_j, x)}{h})}{(v_n^a(y_j, x))^{1/2}} \quad (3.68)$$

$$+ \sup_{t \in \mathbb{R}} \sum_{i=1}^n \Phi_N \frac{\tilde{A}_i(t + \frac{b_n^a(y_j, x)}{h})}{(v_n^a(y_j, x))^{1/2}} - \sum_{i=1}^n \Phi_N \frac{\tilde{A}_i(t + \frac{b_n^a(y_j, x)}{h})}{(v_n^a(y_j, x))^{1/2}} \quad (3.69)$$

$$+ \sup_{t \in \mathbb{R}} \sum_{i=1}^n P_n^a (nh)^{1/2} R_n^{0\text{Boot}}(y_j, x) \quad (3.70)$$

El término (3.68) tiende a cero de forma casi segura por el Lema 3.13. El término (3.69) tiende a cero con n y (3.70) tiende a cero puesto que

$$\sup_{[a;b] \in I} \sum_{j \in I} R_n^{0\text{Boot}}(y_j, x) = O_P \left(\frac{\mu_{\ln n}^{1/3}}{nh} \right) \quad c.s.(P)$$

Estos resultados junto con (3.67) prueban el presente lema.

4. Demostración del Teorema 3.7.

Teniendo en cuenta la desigualdad triangular, se tiene

$$\begin{aligned} & \sup_{t \in \mathbb{R}} \sum_{i=1}^n P_n^a (nh)^{1/2} \left| \hat{F}_h^{\text{Boot}}(y_j, x) - \hat{F}_g^{\text{tc}}(y_j, x) \right| \cdot t_i \\ & \leq \sup_{t \in \mathbb{R}} \sum_{i=1}^n P_n^a (nh)^{1/2} \left| \hat{F}_h^{\text{tc}}(y_j, x) - F(y_j, x) \right| \cdot t_i \\ & \quad + \sup_{t \in \mathbb{R}} \sum_{i=1}^n P_n^a (nh)^{1/2} \left| \hat{F}_h^{\text{Boot}}(y_j, x) - \hat{F}_h^{\text{tc}}(y_j, x) \right| \cdot t_i \Phi_N \frac{\mu_{t_i, b(y_j, x)}}{s(y_j, x)} \\ & \quad + \sup_{t \in \mathbb{R}} \sum_{i=1}^n P_n^a (nh)^{1/2} \left| \hat{F}_h^{\text{tc}}(y_j, x) - F(y_j, x) \right| \cdot t_i \Phi_N \frac{\mu_{t_i, b(y_j, x)}}{s(y_j, x)} \end{aligned}$$

Aplicando ahora, el Teorema 2.5 y el Lema 3.15 se prueban las convergencias a 0 de cada uno de los sumandos, y por tanto queda demostrado este teorema.

Capítulo 4

Aplicación práctica.

En los capítulos anteriores hemos revisado los principales modelos para el tratamiento de datos obtenidos en presencia de censura y truncamiento, centrándonos de modo especial en la definición y estudio de un estimador para la función de distribución condicional. Este capítulo pretende ilustrar el comportamiento de dicho estimador en un problema con datos reales, problema que ha sido objeto de diversos estudios en el campo del análisis de supervivencia (Andersen, Borgan, Gill y Keiding (1993), pag. 14). Se trata de analizar los datos recopilados por el doctor Anders Green sobre la mortalidad de diabéticos en el condado de Fyn (Dinamarca). Concretamente, la situación es la siguiente:

A 1 de julio de 1973, el condado de Fyn tenía alrededor de 450.000 habitantes, de los cuales 1.499 padecían una diabetes mellitus dependiente de insulina. Esta información fue obtenida mediante la recopilación en los registros del Servicio Nacional de Salud de todas las prescripciones de insulina, durante un periodo de 5 meses (cubriendo la fecha arriba citada) y, posteriormente, chequeando el historial médico de cada paciente que proporcionó una serie de variables relevantes. A 1 de enero de 1982 (8 años y medio después del inicio del estudio) se evaluó el estado de supervivencia (fallo, no fallo) de cada uno de los pacientes, mediante el registro de habitantes de la zona. De todo ello, se obtuvo un registro de 1.499 casos, con la siguiente información para cada caso:

1. Nº de caso o de identificación del paciente (de 1 a 1.499).
2. Sexo: Hombre=1, Mujer=0.
3. Estado de supervivencia: Fallo=1, No fallo=0.
4. Fecha de salida del estudio: ddmmaa (día-mes-año).
5. Edad (en años) a la salida.
6. Edad (en años) a la entrada del estudio (010773).
7. Edad (en años) de diagnóstico de la enfermedad.

8. Duración (en meses) de la enfermedad a la entrada (010773).

Estas variables nos proporcionan unas primeras consideraciones importantes:

- La presencia de censura por la derecha aparece de forma natural dado que al ...nal del estudio hay muchos pacientes que todavía viven y por tanto, la variable "edad a la salida" es el mínimo entre el tiempo de fallo (variable de interés) y el tiempo de censura. Con la notación utilizada en los capítulos anteriores se tiene que la variable observada es $Z = \text{"Edad a la salida del estudio"}$ y la variable indicadora de fallo es $\pm = \text{"Estado de supervivencia"}$.

- Al estar interesados en el estudio del tiempo de fallo de pacientes diabéticos, sólo tiene sentido el seguimiento de pacientes después del comienzo de su enfermedad, esto es, de pacientes que hayan sido diagnosticados. Pero debido al sistema de recopilación utilizado, sólo observaremos datos de aquellos pacientes diagnosticados que estén vivos en la fecha de inicio del estudio. Es decir, de...niendo la variable truncamiento $T = \text{"Edad a la entrada del estudio"}$, sólo observaremos pacientes con $T \cdot Z$ y perderemos la información relativa a diabéticos que se han muerto o perdido antes del inicio del estudio. Esto da lugar a un truncamiento por la izquierda que debe ser tenido en cuenta para evitar el sesgo de estimación que se cometería al utilizar una muestra de datos que pierde los pacientes con tiempos de fallo cortos (en los cuales Z no llega a T).

- Tal como vienen dados los datos, las variables Y , C y T miden la edad de cada individuo en diferentes fechas del calendario: la de su fallo, la de su pérdida de seguimiento o fecha ...nal del estudio (si sobrevive) y la de inicio del estudio, respectivamente. Otra posibilidad, más acorde con los modelos de...nidos en los capítulos anteriores, es la de de...nir las variables Y , C y T como tiempos de fallo, de censura y de truncamiento, entendiendo como tales los tiempos transcurridos desde el diagnóstico de la enfermedad hasta el fallo, la pérdida de seguimiento y la entrada en el estudio, respectivamente.

- Posibles covariables en este estudio pueden ser el sexo, la edad de diagnóstico de la enfermedad o la duración de la misma.

A continuación presentamos algunos datos de interés, fruto de un sencillo análisis descriptivo:

- De los 1.499 pacientes diabéticos, 783 son hombres (52,23%) y 716 mujeres (47,77%).

- El porcentaje global de censura es del 67,24% (1.008 casos censurados: 3 por emigración y el resto por sobrevivir al estudio). Entre los hombres el porcentaje de censura es del 67,56% (529 casos) y entre las mujeres del 66,9% (479 casos).

- Las medidas resumen de la covariable $X = \text{"Edad (en años) de diagnóstico"}$ por sexos, son las siguientes

| | Hombres | Mujeres |
|----------------|---------|---------|
| Media = | 30; 65 | 35; 12 |
| D: típica = | 18; 66 | 20; 88 |
| Percentil 25 = | 15 | 15 |
| Mediana = | 29 | 35 |
| Percentil 75 = | 45 | 52 |
| Mínimo = | 0 | 1 |
| Máximo = | 85 | 87 |

- La población de diabéticos por edades a 1 de julio de 1973 (fecha de comienzo del estudio) y la proporción de muertes durante los 8 años y medio de estudio, desglosados por sexos (Análisis 1), se reflejan en las siguientes tablas:

Hombres

| Edad a 1-7-73 (en años) | Pacientes a 1-7-73 | Muertes durante el estudio | Proporción de muertes |
|----------------------------|-----------------------|-------------------------------|--------------------------|
| 0 ; 29 | 207 | 9 | 0; 043 |
| 30 ; 39 | 115 | 22 | 0; 191 |
| 40 ; 49 | 124 | 25 | 0; 202 |
| 50 ; 59 | 122 | 43 | 0; 352 |
| 60 ; 69 | 126 | 78 | 0; 619 |
| ≥ 70 | 89 | 77 | 0; 865 |
| Total | 783 | 254 | 0,324 |

Mujeres

| Edad a 1-7-73 (en años) | Pacientes a 1-7-73 | Muertes durante el estudio | Proporción de muertes |
|----------------------------|-----------------------|-------------------------------|--------------------------|
| 0 ; 29 | 146 | 6 | 0; 041 |
| 30 ; 39 | 89 | 10 | 0; 112 |
| 40 ; 49 | 83 | 10 | 0; 120 |
| 50 ; 59 | 107 | 29 | 0; 271 |
| 60 ; 69 | 143 | 69 | 0; 483 |
| ≥ 70 | 148 | 113 | 0; 769 |
| Total | 716 | 237 | 0,331 |

De ellas, al menos de forma preliminar, parece deducirse que la mortalidad aumenta con la edad y que los hombres tienen una mortalidad más alta que las mujeres.

Esta última información, que hemos llamado Análisis 1, se puede ver en la pag. 15 de la referencia anteriormente citada (Andersen, Borgan, Gill y Keiding (1993)).

Estos autores realizan, además, otros análisis, desglosados por sexos, en el contexto de los modelos de censura y truncamiento sin covariables entre los que nos interesa destacar:

- Análisis 2: Estudio del porcentaje de individuos a riesgo en y , esto es $C_n(y) = \frac{1}{n} \sum_{i=1}^n 1_{T_i \cdot y \cdot Z_i} g$, estimación de la función que hemos denotado en el capítulo 1 por $C(y) = P(T \cdot y \cdot Z | T \cdot Z)$ (ver (1.25)).

- Análisis 3: Cálculo del estimador de la función razón de fallo acumulada (ver (1.31)), dado por

$$\Lambda_n^{tc}(y) = \sum_{i=1}^n \frac{1_{fZ_i \cdot y; \pm_i = 1} g}{nC_n(Z_i)}$$

-Análisis 4: Cálculo del estimador de la función de supervivencia (ver (1.32)):

$$S_n^{tc}(y) = 1 - \sum_{i=1}^n \frac{1_{fZ_i \cdot y; \pm_i = 1} g}{nC_n(Z_i)}$$

Nosotros vamos a realizar los cuatro análisis descritos, pero de forma condicional, tomando como covariable X la edad de diagnóstico de la enfermedad y desglosando también los resultados por sexos. Entre todos los posibles valores de la variable X , hemos elegido tres: 15, 30 y 50 años, por coincidir (más o menos según el sexo) con los cuartiles de la variable.

En cualquiera de los análisis condicionales que pretendemos realizar, nos encontramos con un primer problema que hemos de solucionar: el de la elección del parámetro ventana. Nos centraremos, a continuación, en elegir una ventana o parámetro de suavización que proporcione un "buen" estimador de la función de distribución condicional.

4.1 Selección de la ventana.

La selección del parámetro de suavización es un problema muy relevante y complejo de la estimación no paramétrica de curvas.

En el contexto de la estimación tipo núcleo de la función de distribución, donde el estimador de la distribución F viene dado por

$$\hat{F}_h(x) = \frac{1}{n} \sum_{i=1}^n W \left(\frac{x_i - x}{h} \right)$$

siendo W una función de distribución, cabe citar algunas propuestas recientes debidas a Sarda (1993), Altman y Leger (1995) y Bowman, Hall y Prvan (1998), entre

otros. Estos últimos autores proponen un método de validación cruzada, adaptado a la estimación núcleo de funciones de distribución, que consiste en minimizar la función dada por:

$$CV(h) = \frac{1}{n} \sum_{i=1}^n \int_{F_{X_i} - h}^{F_{X_i} + h} \hat{F}_h^{(i)}(x)^2 dx$$

donde $\hat{F}_h^{(i)}(x)$ denota la estimación no paramétrica de F en el punto x construida sin la observación i -ésima.

Además dichos autores prueban que la función

$$H(h) = CV(h) + \frac{1}{n} \sum_{i=1}^n \int_{F_{X_i} - h}^{F_{X_i} + h} F(x)^2 dx$$

tiene esperanza igual a

$$E[H(h)] = E \int_{F_{X_i} - h}^{F_{X_i} + h} \hat{F}_h^{(i)}(x)^2 dx = MISE_{n_i-1}(h)$$

donde $\hat{F}_h^{(i)}(x)$ denota el estimador núcleo construido con una muestra de tamaño $n_i - 1$.

De lo anterior se deduce que la función $CV(h)$ proporciona un estimador insesgado de la curva $MISE(h)$ para tamaño muestral $n_i - 1$, trasladada verticalmente por una constante desconocida. Como dicha constante no depende de h no afecta a la elección de la ventana óptima.

Nosotros hemos generalizado este criterio a nuestro contexto de estimación de la función de distribución condicional con censura y truncamiento.

De este modo, hemos seleccionado el parámetro ventana que minimiza el siguiente criterio:

$$CVG(h) = \sum_{i=1}^n w_i \int_{F_{Y_i} - h}^{F_{Y_i} + h} \hat{F}_h^{tc(i)}(y|x)^2 dy \quad (4.1)$$

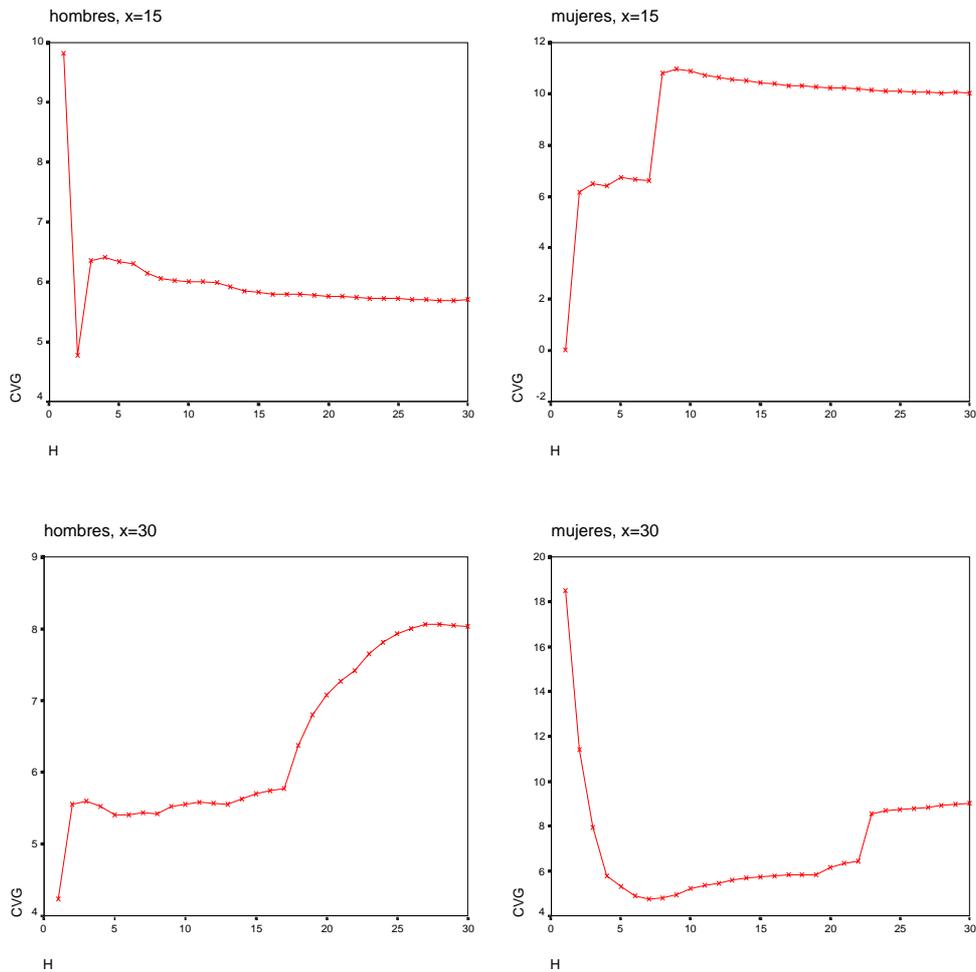
donde $\hat{F}_h^{tc(i)}(y|x)$ denota el estimador de la función de distribución condicional con censura y truncamiento construido sin los datos del individuo i -ésimo y w_i es el salto de $\hat{F}_h^{tc}(y|x)$ en Z_i , dado por $w_i B_{hi}(x) + \hat{F}_h^{tc}(Z_i^-|x) = \hat{C}_h(Z_i|x)$:

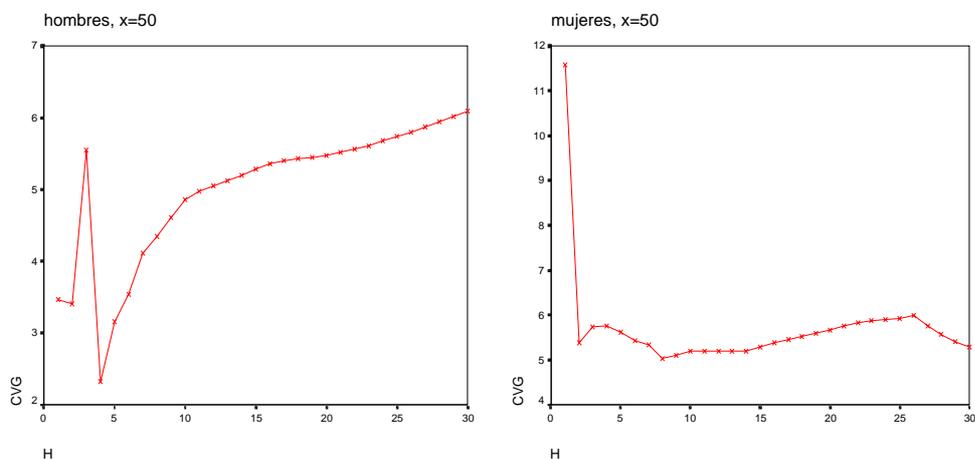
El funcionamiento de este criterio, así como el planteamiento de cualquier otro que proporcione un parámetro ventana en el presente contexto, excede el alcance de esta memoria y deberá ser objeto de posteriores y más exhaustivos estudios.

Una vez decidido el modo de elegir el parámetro ventana, lo hemos aplicado a nuestros datos. Como ya hemos comentado anteriormente, para cada individuo

se han considerado las variables Z = "Tiempo transcurrido desde el diagnóstico de la enfermedad hasta la salida del estudio", T = "Tiempo transcurrido desde el diagnóstico de la enfermedad hasta la entrada en el estudio" y X = "Edad (en años) de diagnóstico". Para las variables Z y T hemos elegido como unidad de tiempo el año, si bien ha sido recalculada a partir del tiempo medido en meses y por ello no presenta un aspecto tan discretizado como el de la variable X . Finalmente, decir que hemos utilizado la función núcleo de Epanechnikov $K(x) = 0.75 \cdot 1 - |x|^3$; si $|x| \leq 1$:

El comportamiento del criterio $CVG(h)$ (expuesto en (4.1)) para los distintos valores de x (15, 30 y 50), según los sexos, se presenta en las siguientes gráficas:





Observación. El valor cero alcanzado por el criterio $CVG(h)$ para $h = 1$ en el grupo de las mujeres con $x = 15$ no debe ser tenido en cuenta al proceder de un sumatorio con todos los términos nulos. Ello se debe a que los 8 casos de mujeres diagnosticadas a los 15 años son censurados.

Como es habitual cuando se trabaja con un criterio de validación cruzada, hemos elegido la ventana óptima como el mayor de los mínimos locales. De este modo, los resultados para la ventana óptima según el criterio $CVG(h)$ (dado en (4.1)) son:

| | Hombres | Mujeres |
|----------|---------|---------|
| $x = 15$ | $h = 2$ | $h = 7$ |
| $x = 30$ | $h = 6$ | $h = 7$ |
| $x = 50$ | $h = 4$ | $h = 8$ |

(4.2)

4.2 Análisis condicionales.

ANÁLISIS 1: Presentamos las tablas descriptivas con la población de diabéticos por edades a 1 de julio de 1973 y el número de muertes durante los 8 años y medio de estudio, desglosados por sexos, para aquellos pacientes que cumplen la condición $jx_j \cdot X_{ij} \cdot h$, esto es, que han sido diagnosticados en torno a los 15, 30 y 50 años, respectivamente. Para ello, hemos utilizado alrededor de cada una de las tres edades de diagnóstico las ventanas de la tabla expuesta en (4.2).

Para $x = 15$ se tiene:

Hombres ($h = 2$)

| Edad a 1-7-73 (en años) | Pacientes a 1-7-73 | Muertes durante el estudio | Proporción de muertes |
|----------------------------|-----------------------|-------------------------------|--------------------------|
| 0 j 29 | 28 | 1 | 0,036 |
| 30 j 39 | 13 | 5 | 0,385 |
| 40 j 49 | 3 | 2 | 0,667 |
| 50 j 59 | 7 | 1 | 0,143 |
| 60 j 69 | 2 | 0 | 0 |
| ≥ 70 | 0 | 0 | 0 |
| Total | 53 | 9 | 0,170 |

Mujeres ($h = 7$)

| Edad a 1-7-73 (en años) | Pacientes a 1-7-73 | Muertes durante el estudio | Proporción de muertes |
|----------------------------|-----------------------|-------------------------------|--------------------------|
| 0 j 29 | 78 | 4 | 0,051 |
| 30 j 39 | 37 | 5 | 0,135 |
| 40 j 49 | 23 | 4 | 0,174 |
| 50 j 59 | 14 | 4 | 0,286 |
| 60 j 69 | 4 | 2 | 0,500 |
| ≥ 70 | 0 | 0 | 0 |
| Total | 156 | 19 | 0,122 |

Para $x = 30$:

Hombres ($h = 6$)

| Edad a 1-7-73 (en años) | Pacientes a 1-7-73 | Muertes durante el estudio | Proporción de muertes |
|----------------------------|-----------------------|-------------------------------|--------------------------|
| 0 j 29 | 7 | 0 | 0 |
| 30 j 39 | 41 | 2 | 0,049 |
| 40 j 49 | 54 | 10 | 0,185 |
| 50 j 59 | 32 | 10 | 0,312 |
| 60 j 69 | 12 | 8 | 0,667 |
| ≥ 70 | 7 | 5 | 0,714 |
| Total | 153 | 35 | 0,229 |

Mujeres (h = 7)

| Edad a 1-7-73 (en años) | Pacientes a 1-7-73 | Muertes durante el estudio | Proporción de muertes |
|----------------------------|-----------------------|-------------------------------|--------------------------|
| 0 i 29 | 4 | 0 | 0 |
| 30 i 39 | 28 | 1 | 0,036 |
| 40 i 49 | 28 | 2 | 0,071 |
| 50 i 59 | 24 | 6 | 0,25 |
| 60 i 69 | 18 | 10 | 0,556 |
| ≥ 70 | 2 | 1 | 0,50 |
| Total | 104 | 20 | 0,192 |

Y para x = 50:

Hombres (h = 4)

| Edad a 1-7-73 (en años) | Pacientes a 1-7-73 | Muertes durante el estudio | Proporción de muertes |
|----------------------------|-----------------------|-------------------------------|--------------------------|
| 40 i 49 | 3 | 1 | 0,333 |
| 50 i 59 | 18 | 8 | 0,444 |
| 60 i 69 | 37 | 25 | 0,676 |
| ≥ 70 | 10 | 9 | 0,90 |
| Total | 68 | 43 | 0,632 |

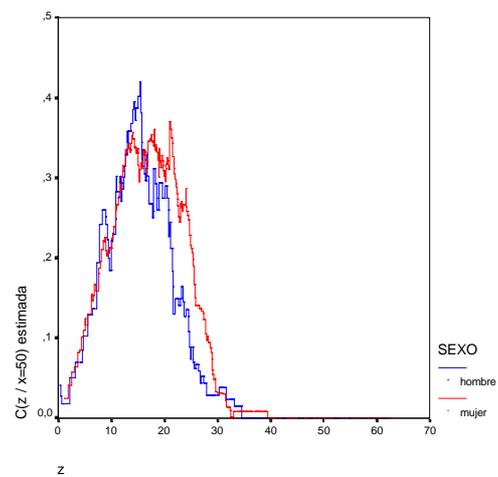
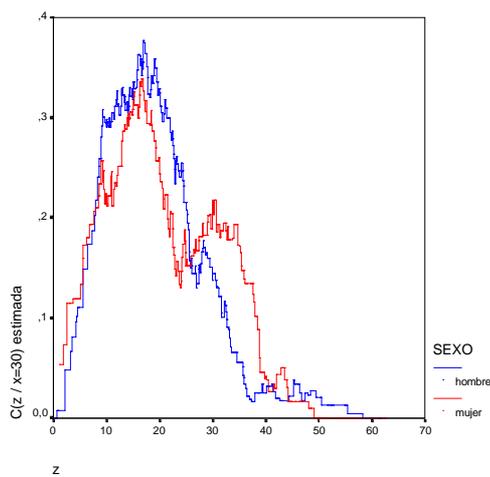
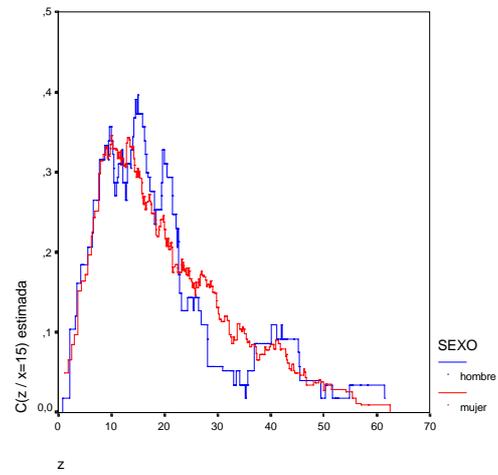
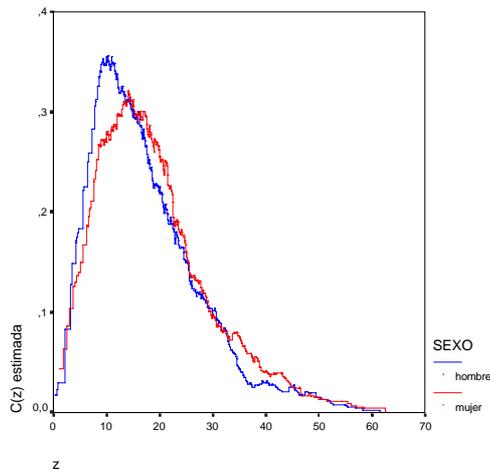
Mujeres (h = 8)

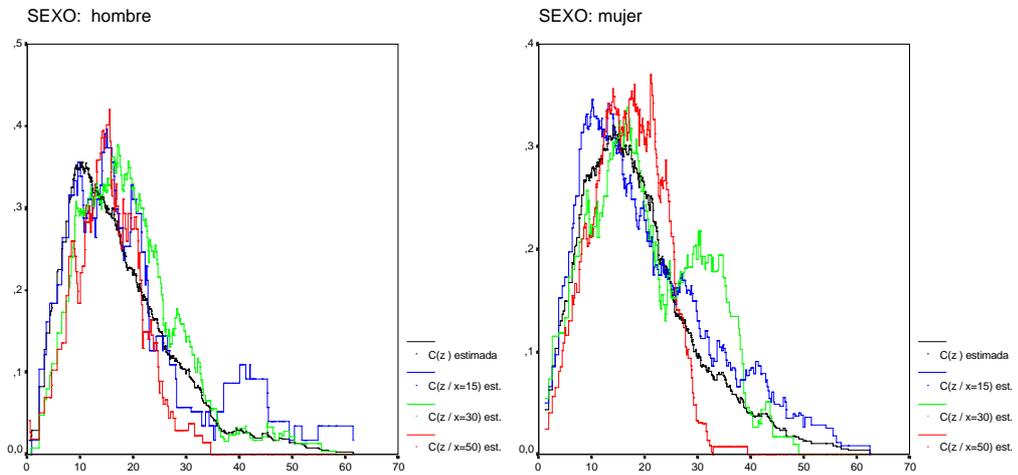
| Edad a 1-7-73 (en años) | Pacientes a 1-7-73 | Muertes durante el estudio | Proporción de muertes |
|----------------------------|-----------------------|-------------------------------|--------------------------|
| 40 i 49 | 5 | 0 | 0 |
| 50 i 59 | 41 | 12 | 0,293 |
| 60 i 69 | 74 | 34 | 0,459 |
| ≥ 70 | 50 | 41 | 0,820 |
| Total | 170 | 87 | 0,512 |

ANÁLISIS 2: En este apartado estudiaremos, para cada sexo, el porcentaje de individuos a riesgo en función del tiempo (en años) transcurrido desde el diagnóstico de la enfermedad, condicionado a los valores $x = 15$; $x = 30$ y $x = 50$; respectivamente, es decir $\hat{C}_h(y | x) = \prod_{i=1}^n 1_{fT_i} \cdot y \cdot z_i g B_{hi}(x)$:

Como en el análisis anterior utilizaremos las ventanas de la tabla (4.2), que según el criterio (4.1) son las óptimas para la estimación de la función de distribución condicional (objetivo de la aplicación práctica de este capítulo), y que han sido calculadas en función del sexo y del valor x de la edad de diagnóstico.

Dichas estimaciones se comparan con la curva incondicional cuya ecuación viene dada por $C_n(y) = \frac{1}{n} \sum_{i=1}^n 1_{F_{T_i}} \cdot y \cdot Z_i g_n^1$:





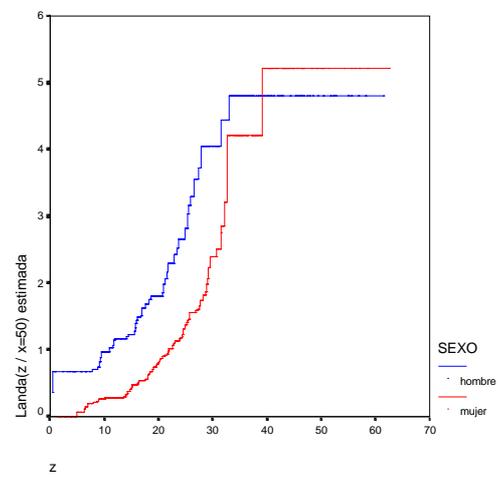
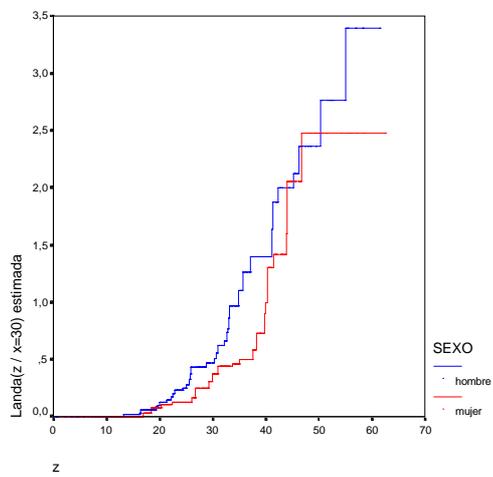
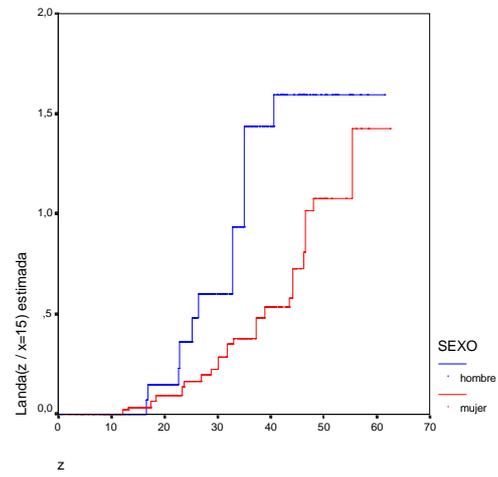
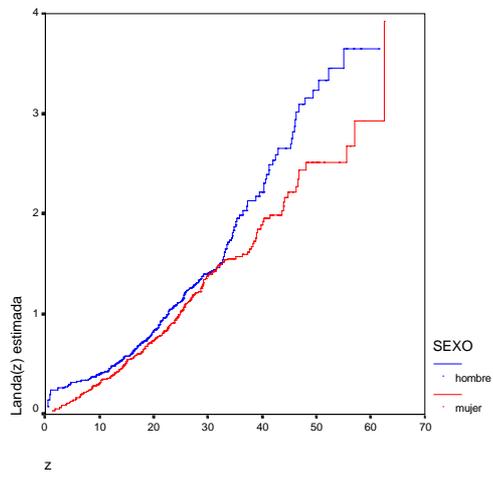
ANÁLISIS 3: En este apartado se calcula, para cada sexo, el estimador de la función razón de fallo acumulada condicionada a los valores para la edad de diagnóstico de $x = 15$; $x = 30$ y $x = 50$ respectivamente, dado por

$$\hat{\Lambda}_n^{tc}(y | x) = \prod_{i=1}^x \frac{1_{fZ_i \cdot y; \pm_i = 1g} B_{hi}(x)}{\hat{C}_n(Z_i | x)}$$

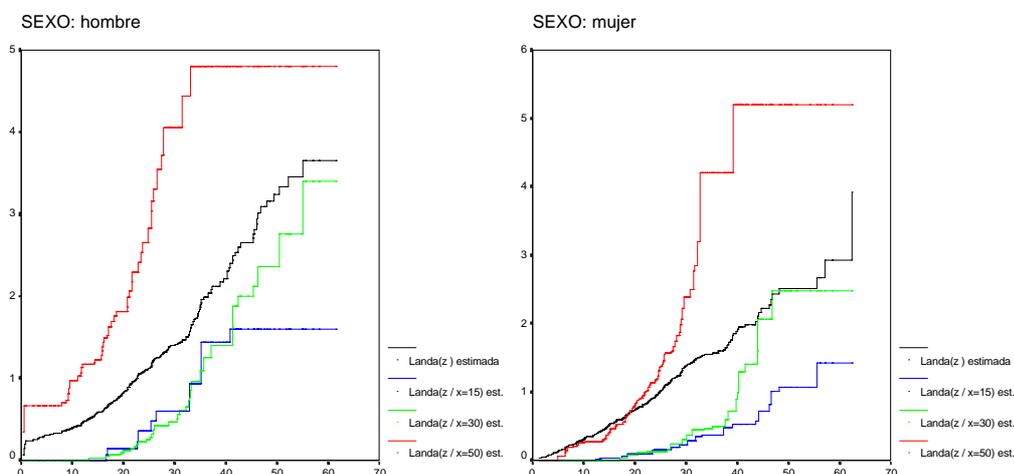
Como en todos los análisis de este capítulo, vamos a utilizar las ventanas de la tabla (4.2).

También presentamos el estimador de la función razón de fallo acumulada para el caso incondicional cuya expresión es

$$\Lambda_n^{tc}(y) = \prod_{i=1}^x \frac{1_{fZ_i \cdot y; \pm_i = 1g}}{nC_n(Z_i)}$$



Agrupando las curvas por sexos, se tiene:



La relación existente entre las funciones razón de fallo acumulada y razón de fallo (la segunda es la derivada de la primera), nos permiten interpretar estas curvas en términos del riesgo de mortalidad para los años transcurridos desde el diagnóstico. Así, un comportamiento convexo de la función razón de fallo acumulada indicará un crecimiento en el riesgo de mortalidad y un comportamiento cóncavo un decrecimiento en el mismo.

Para la edad de diagnóstico de $x = 15$ años se puede observar, tanto para hombres como para mujeres, un comportamiento inicial y bastante prolongado de riesgo de mortalidad creciente conforme aumenta el tiempo desde el diagnóstico, riesgo que ...nalmente tiende a decrecer (a partir de los 35-40 años desde el diagnóstico para los hombres y casi los 50 para las mujeres). Además, la función razón de fallo acumulada en los hombres está por encima de la misma en el grupo de mujeres y la diferencia entre ellas aumenta progresivamente, de lo que se deduce un riesgo de mortalidad mayor en los hombres que en las mujeres.

Para $x = 30$ años el comportamiento para el riesgo de muerte sigue unas pautas similares a las descritas para $x = 15$, en cuanto que para hombres y mujeres se observa un crecimiento inicial y un posterior decrecimiento (el cambio en el crecimiento se puede situar alrededor de los 40 años desde el diagnóstico (70 años de edad) en ambos sexos). Sin embargo, las curvas $\hat{\Lambda}_h^{tc}(y | x)$ por sexos están mucho más próximas que en el caso anterior. Como diferencia más destacable cabe citar un crecimiento del riesgo algo más rápido en los hombres.

Para $x = 50$ años, hay que destacar un riesgo de muerte en periodos cercanos a la edad de diagnóstico bastante importante en el grupo de los hombres que va decreciendo durante los 10 primeros años desde el diagnóstico para posteriormente iniciar una remontada. En las mujeres también se observa un riesgo de mortalidad

que decrece inicialmente, si bien empieza más tardíamente y ocurre de forma más moderada. A partir de los 10-15 años desde el diagnóstico se produce un crecimiento en el riesgo de mortalidad prácticamente similar en ambos sexos.

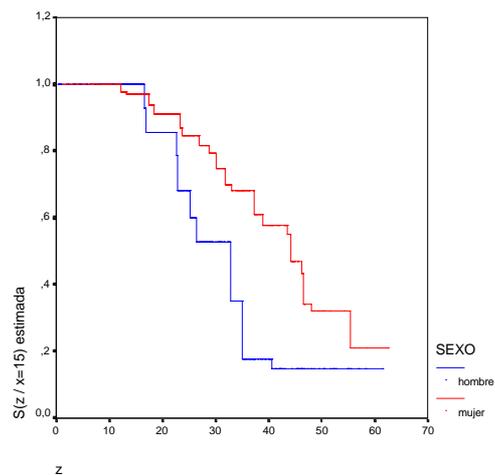
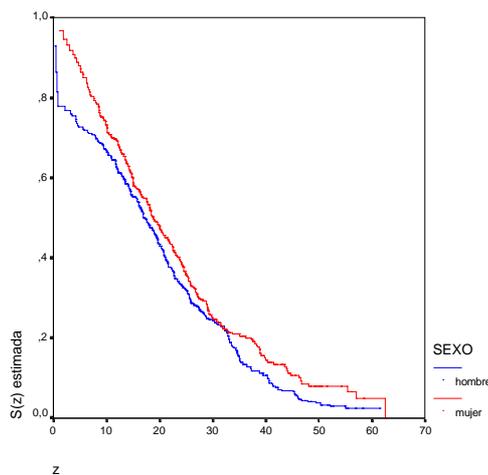
Como última característica destacable mencionaremos que, tanto en hombres como en mujeres, el riesgo de mortalidad comienza antes y crece más rápidamente si la edad de comienzo de la enfermedad es a los 50 años que si es a los 15 o 30 años. Esto se observa fácilmente en las últimas gráficas de este apartado, correspondientes a las funciones de razón de fallo acumuladas condicionales por sexos.

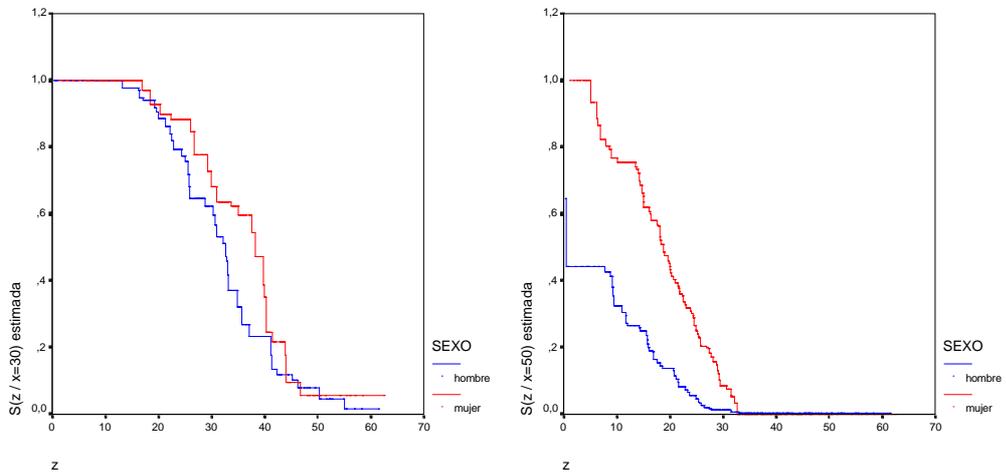
ANÁLISIS 4: En este apartado se calcula, para cada sexo, el estimador de la función de supervivencia condicionada a los valores $x = 15$; $x = 30$ y $x = 50$; respectivamente, dado por

$$\hat{S}_h^{tc}(y | x) = 1 - \prod_{i=1}^n \frac{1_{fZ_i, y; \pm i=1g} B_{hi}(x)}{\hat{C}_h(Z_i | x)}$$

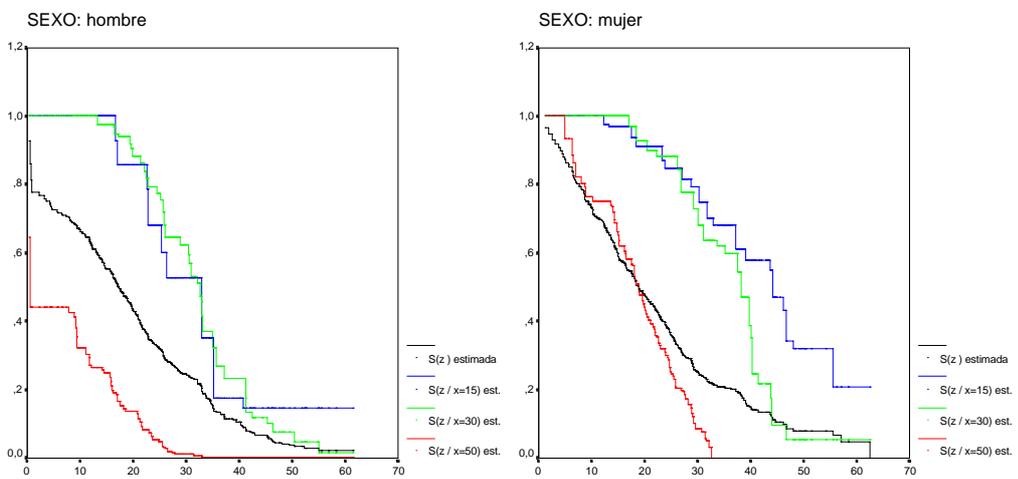
También, presentamos el estimador de la función de supervivencia para el caso incondicional, es decir

$$S_n^{tc}(y) = 1 - \prod_{i=1}^n \frac{1_{fZ_i, y; \pm i=1g}}{nC_n(Z_i)}$$

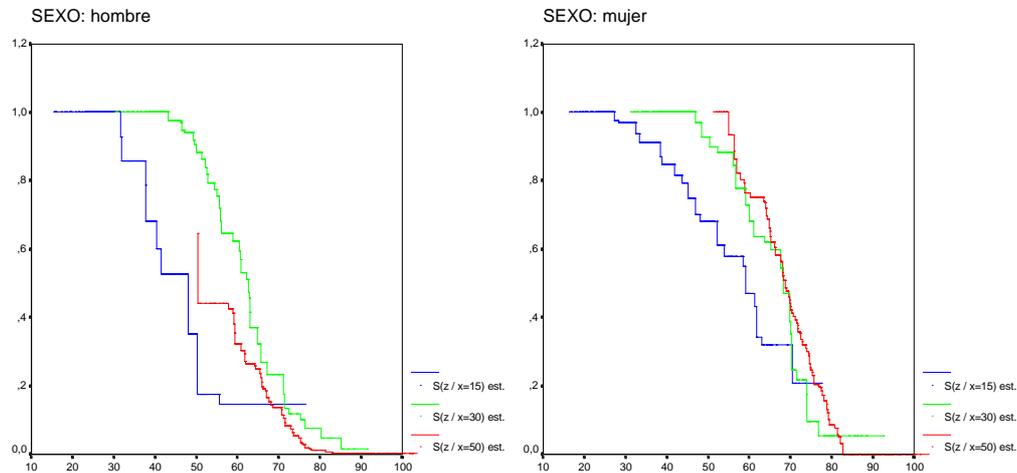




Las curvas de supervivencia para cada sexo, se muestran en las dos gráficas siguientes:



Si consideramos como medida del eje X los años de vida, en lugar de los transcurridos desde el diagnóstico, las curvas condicionales anteriores se ven de la siguiente forma:



Conclusiones:

La primera conclusión que podemos destacar es que el sexo influye en la supervivencia (mortalidad) de los diabéticos. En el caso incondicional está claro que las mujeres presentan una supervivencia más alta que los hombres. Particularizando por edades de diagnóstico, vemos que esta tendencia también se mantiene, aunque podemos profundizar un poco más.

Para $x = 15$ años la diferencia entre las curvas de supervivencia a favor de las mujeres aparece de forma muy marcada, salvo en un pequeño intervalo inicial, y se va incrementando considerablemente conforme aumenta el tiempo transcurrido desde el diagnóstico de la enfermedad. Así, por ejemplo, la probabilidad de sobrevivir 10 años desde el diagnóstico, esto es, de superar los 25 años de vida es del 100% para hombres y mujeres; la probabilidad de sobrevivir 20 años desde el diagnóstico, esto es, de superar los 35 años de vida es del 85,71% para los hombres y sube al 91,22% para las mujeres, mientras que la probabilidad de sobrevivir 30 años desde el diagnóstico (superar los 45) es del 52,86% en los hombres frente al 79,45% de las mujeres y la de sobrevivir 40 años desde el diagnóstico o de superar los 55 años de vida es del 17,62% para los hombres y del 57,77% para las mujeres. Además, los valores medianos de ambas distribuciones son 32,92 y 44,17 años desde el diagnóstico para hombres y mujeres, respectivamente. Prácticamente 11 años de diferencia, a favor de las mujeres.

Para $x = 50$ años, también se observa una diferencia muy marcada entre las curvas de supervivencia de hombres y mujeres, pero aquí la diferencia, que empieza siendo muy importante, disminuye conforme aumenta el tiempo transcurrido desde el diagnóstico. Así, el año de vida desde el diagnóstico sólo lo supera el 44,12% de los hombres frente al 100% de las mujeres, los 10 años de vida desde el diag-

nóstico (60 años de edad) equivalen a los percentiles 67,57 y 23,40 para hombres y mujeres, respectivamente, y los 20 años desde el diagnóstico (70 años de edad) se corresponden con el percentil 86,29 en hombres y el 56,77 en mujeres. En cuanto a las medianas de la distribución son de 0,50 y 18,83 años (desde el diagnóstico) en hombres y en mujeres, respectivamente.

Aunque para $x = 30$ años se mantiene la tendencia de mayor supervivencia en mujeres que en hombres, se observa mucha más cercanía entre las curvas de supervivencia de ambos sexos que para las edades de diagnóstico de 15 y 50 años, e incluso se puede ver la existencia de algunos solapamientos en torno a los 20 y 45 años desde el diagnóstico (50 y 75 años de edad, respectivamente). Los tiempos de 10, 20, 30 y 40 años desde el diagnóstico (edades de 40, 50, 60 y 70 años) corresponden a los percentiles 0; 11,60; 37,70 y 76,70 en hombres y 0; 7,33; 31,82 y 61,26 en mujeres. Las medianas toman valores, en años desde el diagnóstico, de 32,67 para los hombres y 38,25 para las mujeres.

La segunda conclusión importante es que la edad de diagnóstico también influye en el comportamiento de la supervivencia (mortalidad) de los diabéticos. Además de los comentarios efectuados previamente, se puede ver que, en general, la supervivencia en un cierto tiempo desde el diagnóstico disminuye al incrementar la edad de comienzo de la enfermedad: es más difícil sobrevivir 10 años desde el diagnóstico partiendo de 50 años que partiendo de 15 años, pero también es bastante evidente el hecho de que no es lo mismo alcanzar una edad de vida de 60 años que de 25. Por ello, conviene matizar que la afirmación anterior no significa que la supervivencia para cierta edad disminuya al incrementar la edad de comienzo de la enfermedad, es más, de los últimos gráficos se inferiría más bien lo contrario: los diabéticos diagnosticados a los 15 años presentan, tanto en hombres como en mujeres, una supervivencia menor (en edad) que los diagnosticados a los 30 o 50 años.

Finalmente, es importante destacar la fuerte influencia del fenómeno de la censura que se manifiesta de forma muy clara en la parte derecha de las curvas de supervivencia (no acaban valiendo cero), sobre todo en el caso de $x = 15$ años y de forma más débil para $x = 30$ años.

Bibliografía

- Akritis, M.G. (1986). Bootstrapping the Kaplan-Meier estimator. *J. Amer. Statist. Assoc.* 81, 1032-1038.
- Akritis, M.G. (1994). Nearest neighbor estimation of a bivariate distribution under random censoring. *Ann. Statist.* 22, 1299-1327.
- Altman, N. and Leger, C. (1995). Bandwidth selection for kernel distribution function estimation. *J. Statist. Plan. Inf.* 46, 195-214.
- Andersen, P.K., Borgan, O., Gill, R.D. and Keiding, N. (1993). *Statistical Models Based on Counting Processes*. Springer-Verlag, New York.
- Arcones, M.A. and Giné, E. (1995). On the law of the iterated logarithm for canonical U-statistics and processes. *Stoch. Proc. Appl.* 58, 217-245.
- Araujo, A. and Giné, E. (1980). *The central limit theorem for real and Banach valued random variables*. Wiley, New York.
- Beran, R. (1981). *Nonparametric regression with randomly censored data*. Technical report. Univ. California, Berkeley.
- Bickel, P.J. and Wichura, M.J. (1971). Convergence criteria for multiparameter stochastic processes and some applications. *Ann. Math. Statist.* 42, 1656-1670.
- Bilker, W.B. and Wang, M.C. (1997). Bootstrapping left truncated and right censored data. *Commun. Statist. Simula.* 26, 141-171.
- Billingsley, P. (1968). *Convergence of probability measures*. Wiley, New York.
- Bowman, A., Hall, P. and Prvan, T. (1998). Bandwidth selection for the smoothing of distribution functions. *Biometrika.* 85, 799—808.
- Breslow, N.E. and Crowley, J.J. (1974). A large sample study of the life table and product-limit estimates under random censorship. *Ann. Statist.* 2, 437-453.
- Burke, M.D., Csörgö, S. and Horváth, L. (1988). A correction to and improvement of strong approximations of some biometric estimates under random censorship. *Prob. Theory Rel. Fields* 79, 51-57.
- Cao Abad, R. (1988). *El bootstrap con datos censurados*. Memoria de Licenciatura. Universidad de Santiago de Compostela.
- Cao Abad, R. (1990). *Aplicaciones y nuevos resultados del método bootstrap en*

la estimación no paramétrica de curvas. Tesis Doctoral. Universidad de Santiago de Compostela.

Cao Abad, R. and González Manteiga, W. (1993). Bootstrap methods in regression smoothing. *J. Nonparametric Statist.* 2, 379-388.

Chang, M. and Rao, P. (1989). Berry-Esseen bound for the Kaplan-Meier estimator. *Commun. Stat. Theory Meth.* 18, 4647-4664.

Chao, M.T. and Lo, S.H. (1988). Some representations of the nonparametric maximum likelihood estimators with truncated data. *Ann. Statist.* 16, 661-668.

Cheng, K.F. and Cheng, P.E. (1987). Robust nonparametric estimation of a regression function. *Sankhyä, Ser. B,* 49, 9-22.

Dabrowska, D. (1989). Uniform consistency of the kernel conditional Kaplan-Meier estimate. *Ann. Statist.* 17, 1157-1167.

Dabrowska, D. (1992). Nonparametric quantile regression with censored data. *Sankhyä, Ser. A,* 54, 252-259.

Efron, B. (1979). Bootstrap methods: another look at the jackknife. *Ann. Statist.* 7, 1-26.

Efron, B. (1981). Censored data and the Bootstrap. *J. Amer. Statist. Assoc.* 76, 312-319.

Eubank, R.L. (1988). Spline smoothing and nonparametric regression. Marcel Dekker. New York.

Földes, A. and Rejto, L. (1981). A LIL type result for the product limit estimator. *Z. Wahrsch. verw. Gebiete* 56, 75-86.

Gijbels, I. and Wang, J.L. (1993). Strong representation of the survival function estimator for truncated and censored data with applications. *J. Multivariate Anal.* 47, 210-229.

Gill, R. (1983). Large sample behaviour of the product-limit estimator on the whole line. *Ann. Statist.* 11, 49-58.

González Manteiga, W. and Cadarso Suárez, C. (1994). Asymptotic properties of a generalized Kaplan-Meier estimator with some applications. *Nonparametric Statist.* 4, 65-78.

Gross, S.T. and Lai, T.L. (1996). Bootstrap methods for truncated and censored data. *Statistica Sinica,* 6, 509-530.

Härdle, W. (1990). Applied nonparametric regression. Cambridge University Press.

Härdle, W., Janssen, P. and Serfling, R. (1988). Strong uniform consistency rates for estimators of conditional functionals. *Ann. Statist.* 16, 1428-1449.

Härdle, W. and Mammen, E. (1993). Comparing nonparametric versus parametric regression ...ts. *Ann. Statist.* 21, 1926-1947.

Härdle, W. and Marron, S. (1991). Bootstrap simultaneous error bars for non-parametric regression. *Ann. Statist.* 19, 778-796.

Iglesias Pérez, M.C. and González Manteiga, W. (1999). Strong representation of a generalized product-limit estimator for truncated and censored data with some applications. *J. Nonparametric Statist.* 10, 213-244.

Kaplan, E.L. and Meier, P. (1958). Nonparametric estimation from incomplete observations. *J. Amer. Statist. Assoc.* 53, 457-481.

Keiding, N. and Gill, R.D. (1990). Random truncation models and Markov Processes. *Ann. Statist.* 18, 582-602.

Lagakos, S.W., Barraj, L.M. and De Gruttola, V. (1988). Nonparametric Analysis of truncated survival data, with application to AIDS. *Biometrika.* 75, 515-523.

Lai, T.L. and Ying, Z. (1991). Estimating a distribution function with truncated and censored data. *Ann. Statist.* 19, 417-442.

LaValley, M.P. and Akritas, M.G. (1994). Extensions of the Lynden-Bell-Woodroffe method for truncated data. Submitted for publication.

Lo, S. and Singh, K. (1985). The product-limit estimator and the bootstrap: Some asymptotic representations. *Prob. Theory Rel. Fields* 71, 455-465.

Lo, S.H., Mack, Y.P. and Wang, J.L. (1989). Density and hazard rate estimation for censored data via strong representation of the Kaplan-Meier estimator. *Prob. Theory Rel. Fields* 80, 461-473.

Lynden-Bell, D. (1971). A method of allowing for known observational selection in small samples applied to 3CR quasars. *Mon. Not. R. Astr. Soc.* 155, 95-118.

Mammen, E. (1992). When does Bootstrap works? Lecture notes in Statistics, 77. Springer-Verlag.

Nadaraya, E.A. (1964). On estimating regression. *Theory Prob. Appl.* 10, 186-190.

Parzen, E. (1962). On estimation of a probability density and mode. *Ann. Mathem. Statist.* 35, 1065-1076.

Presedo Quindimil, M.A. (1991). Un análisis mediante la función de influencia de la estimación no paramétrica de funciones de regresión. Tesis Doctoral. Universidad de Santiago de Compostela.

Rosenblatt, M. (1956). Remarks on some nonparametric estimates of a density function. *Ann. Mathem. Statist.* 27, 642-669.

Sarda, P. (1993). Smoothing parameter selection for smooth distribution functions. *J. Statist. Plan. Inf.* 35, 65-75.

Shao, J. and Tu, D. (1995). The jackknife and bootstrap. Springer-Verlag, New York.

Shorack, G.R. and Wellner, J.A. (1986). Empirical processes with applications to statistics. Wiley, New York.

Stute, W. (1976). On a generalization of the Glivenko-Cantelli theorem. *Z. Wahrsch. verw. Gebiete* 35, 167-175.

Stute, W. (1984). The oscillation behavior of empirical processes: the multivariate case. *Ann. Probab.* 12, 361-379.

Stute, W. (1993a). Almost sure representations of the product-limit estimator for truncated data. *Ann. Statist.* 21, 146-156.

Stute, W. (1993b). Integrated U-statistic processes: A martingale approach. Preprint. Univ. Giessen.

Stute, W. (1994). The bias of Kaplan-Meier integrals. *Scand. J. Statist.* 21, 475-484.

Stute, W. (1995). The central limit theorem under random censorship. *Ann. Statist.* 23, 422-439.

Stute, W. (1996). The jackknife estimate of variance of a Kaplan-Meier integral. *Ann. Statist.* 24, 2679-2704.

Stute, W. (1997). Nonparametric model checks for regression. *Ann. Statist.* 25, 613-641.

Stute, W. and Wang, J.L. (1993). The strong law under random censorship. *Ann. Statist.* 21, 1591-1607.

Tsai, W.Y., Jewell, N.P. and Wang, M.C. (1987). A note on the product-limit estimator under right censoring and left truncation. *Biometrika.* 74, 883-886.

Turnbull, B.W. (1976). The empirical distribution function with arbitrarily grouped, censored and truncated data. *J. R. Statist. Soc. B*, 38, 290-295.

Van Keilegom, I. and Veraverbeke, N. (1996). Uniform strong convergence results for the conditional Kaplan-Meier estimator and its quantiles. *Commun. Statist. Theory Meth.* 25, 2251-2265.

Van Keilegom, I. and Veraverbeke, N. (1997a). Estimation and bootstrap with censored data in ...xed design nonparametric regression. *Ann. Inst. Statist. Math.* 49, 467-491.

Van Keilegom, I. and Veraverbeke, N. (1997b). Weak convergence of the bootstrapped conditional Kaplan-Meier process and its quantile process. *Commun. Statist. Theory Meth.* 26, 853-869.

Wang, M.C. (1987). Product limit estimates: A generalized maximum likelihood study. *Commun. Statist. Theory Meth.* 16, 3117-3132.

Wang, M.C. (1991). Nonparametric estimation from cross-sectional survival data. *J. Amer. Statist. Assoc.* 86, 130-143.

Wang, M.C., Jewell, N.P. and Tsai, W.Y. (1986). Asymptotic properties of the product limit estimate under random truncation. *Ann. Statist.* 14, 1597-1605.

Watson, G.M. (1964). Smooth regression analysis. *Sankhyä, Ser. A*, 26, 359-372.

Woodroffe, M. (1985). Estimating a distribution function with truncated data. *Ann. Statist.* 13, 163-177.

Zhou, Y. (1996). A note on the TJW product-limit estimator for truncated and censored data. *Statist. and Probab. Letters.* 26, 381-387.

Zhou, Y. and Yip, P.S.F. (1999). A strong representation of the product-limit estimator for left truncated and right censored data. *J. Multivariate Anal.* 69, 261-280.