



UNIVERSIDADE DA CORUÑA

DEPARTAMENTO DE MATEMÁTICAS

**ESTIMACIÓN PRESUAVIZADA DE LAS
FUNCIONES DE DENSIDAD Y DISTRIBUCIÓN
CON DATOS CENSURADOS**

M^a Amalia Jácome Pumar
Julio 2005

Tesis Doctoral

Realizado el acto público de defensa y mantenimiento de esta Tesis Doctoral el día 11 de julio de 2005 en la Facultad de Informática de la Universidad de La Coruña, ante el tribunal formado por:

Presidente: Dr. D. Wenceslao González Manteiga

Vocales: Dr. D. Jacobo de Úña Álvarez
Dra. Dña. Irène Gijbels
Dr. D. Noël Veraverbeke

Secretario: Dr. D. Juan Manuel Vilar Fernández

obtuvo la máxima calificación de SOBRESALIENTE CUM LAUDE, siendo director de la misma el Dr. D. Ricardo Cao Abad.

*A Manolo
y a mis padres*

Agradecimientos

En primer lugar, deseo expresar mi más profundo agradecimiento al director de esta tesis, el Profesor Ricardo Cao Abad, por su paciencia y estimulante apoyo. Siento por él no sólo gratitud, sino una gran admiración.

Agradezco también a la Profesora Irène Gijbels su calurosa acogida durante mi visita al Institut de Statistique (Louvain-la-Neuve, Bélgica), y su inestimable ayuda en la realización del capítulo 5 de este trabajo.

Gracias al Departamento de Matemáticas de la Universidad de la Coruña, en particular a los integrantes del área de Estadística e Investigación Operativa, responsables en gran medida de mi formación estadística.

A los miembros del Departamento de Estadística e Investigación Operativa de la Universidad de Vigo, y en especial a los compañeros del grupo de Inferencia no Paramétrica.

Así mismo, quiero agradecer el apoyo de mis compañeros en el campus de Ourense, y especialmente a Ana Pérez, Tomás Cotos y Rubén Fernández, que han conseguido que la elaboración de esta tesis fuese más agradable. Aunque sólo sea por el tiempo que pasamos juntos, son casi mi segunda familia.

Quisiera dar las gracias también a Alberto por su optimismo, a Juan Carlos (sin él Louvain-la-Neuve no es lo mismo), a Leticia, María, Javier... Aunque no nos veamos tanto como yo quisiera, me he sentido apoyada en todo momento.

A toda mi familia, amigos y muy especialmente a mis padres y hermanos, por su cariño, comprensión y ánimo durante todos estos años.

A Manolo, pues parte de esta tesis es suya.

Finalmente, hago constar que parte de la investigación ha sido subvencionada por los proyectos PGIDT03PXIC10505PN (financiado por la Xunta de Galicia), BFM2002-00265 (financiado por el Ministerio de Ciencia y Tecnología), la Dirección Xeral de Investigación e Desenvolvemento de la Xunta de Galicia, el proyecto “Análisis de datos censurados: aplicaciones en fiabilidad industrial” (financiado por Caixa Galicia) y los proyectos “Modelos estadísticos flexibles para el análisis de datos” y “Modelos estadísticos flexibles para el análisis de datos II”, financiados por la Universidad de Vigo.

Índice general

Introducción	1
1. Estimación en el modelo de censura aleatoria por la derecha	5
1.1. Introducción a los modelos de censura	5
1.1.1. El modelo de censura aleatoria por la derecha	7
1.2. Estimación de la función de distribución	8
1.2.1. El estimador de Kaplan-Meier	8
1.2.2. Propiedades del estimador de Kaplan-Meier	11
1.2.3. El estimador semiparamétrico de Dikta	19
1.3. Estimación de la función de densidad.	25
1.3.1. El estimador basado en el estimador de Kaplan-Meier	25
2. Estimación presuavizada de la función de distribución	37
2.1. Introducción a la estimación presuavizada	37
2.2. Definición del estimador	38
2.3. Propiedades generales del estimador	39
2.4. Propiedades asintóticas del estimador	43
2.4.1. Lemas previos	43
2.4.2. Representación asintótica	52
2.4.3. Normalidad asintótica	80
2.5. Estudio de simulación	84
3. Estimación presuavizada de la función de densidad	93
3.1. Definición del estimador. Propiedades	93
3.2. Propiedades asintóticas del estimador	94
3.2.1. Representación asintótica del estimador	94
3.2.2. Normalidad asintótica	153
3.2.3. Representación asintótica del MISE	159
3.3. Estudio de simulación	167

4. Selección del parámetro de suavización	177
4.1. Representación de las ventanas MISE	178
4.2. Comparación de las ventanas MISE y AMISE por simulación	189
4.3. Selectores de las ventanas de tipo plug-in	197
4.3.1. Consistencia de las ventanas plug-in	199
4.3.2. Estudio de simulación	234
4.4. Selectores de las ventanas de tipo bootstrap	241
4.4.1. Remuestreos bootstrap para datos con censura	242
4.4.2. Estudio de simulación	245
5. Comparación de los estimadores de Nadaraya-Watson y local lineal en la estimación presuavizada de la densidad	257
5.1. Los estimadores de Nadaraya-Watson y local lineal de la regresión . .	257
5.2. Estimación presuavizada de la densidad con los diferentes ajustes de p	259
5.2.1. Estimación local logística de p	261
5.3. Estudio de simulación	263
6. Aplicación a datos reales	271
6.1. Estimación de la función de distribución	274
6.2. Estimación de la función de densidad	279
Bibliografía	289

Introducción

Esta memoria se enmarca dentro del análisis de supervivencia, un campo de la Estadística de creciente interés en los últimos años. Su objeto es el estudio de variables aleatorias positivas que miden, en general, el tiempo que transcurre desde un cierto suceso inicial hasta un suceso final. Estas variables, denominadas *tiempos de vida*, pueden ser el tiempo que transcurre hasta que una componente física (mecánica o eléctrica) experimenta el primer fallo, o el tiempo hasta la muerte de una unidad biológica (paciente, animal, célula, etc.). Sus importantes aplicaciones en ciencias como la Medicina o la Biología, así como en Ingeniería (estudios de fiabilidad) y en Economía, han provocado una preocupación por la mejora de los métodos de Inferencia Estadística propuestos en este contexto.

Los métodos de inferencia estadística se basan en la información proporcionada por una muestra aleatoria procedente de una distribución desconocida, y se desea estimar una característica de la población, que puede ser un parámetro o una curva de interés. Este problema puede enfocarse desde dos puntos de vista diferentes:

- **Paramétricamente**, donde se supone que la distribución teórica pertenece a una familia de distribuciones conocida, $F_{\theta}(x)$, indicada por un parámetro finito dimensional θ . El problema se reduce a estimar el parámetro a partir de la información contenida en la muestra. El mayor inconveniente es el riesgo de cometer un error de especificación del modelo, que podría llegar a invalidar los resultados obtenidos.
- **No Paramétricamente**, tratando de establecer el modelo probabilístico que ha generado los datos sin ninguna suposición inicial sobre la distribución. La única información es la proporcionada por la muestra.

Nuestro interés se centrará en la estimación no paramétrica de curvas, una de las áreas más activas de la inferencia estadística, puesto que es una metodología muy flexible que puede usarse en multitud de situaciones prácticas. Permite relajar las hipótesis sobre el modelo teórico y es muy apropiada en un estudio exploratorio previo

de los datos. En particular, los métodos no paramétricos de estimación de la función de densidad son extremadamente útiles para determinar las características de una población a partir de una muestra, y tienen aplicaciones directas en muchos problemas de inferencia.

El primer estimador no paramétrico de la función de densidad considerado fue el histograma, pero su construcción es muy subjetiva. Los métodos modernos de estimación de la densidad se han ido desarrollando a partir de los años 50, dando lugar a estimadores continuos, más adecuados que los simples histogramas: el método del vecino más próximo, el estimador de series, el estimador spline y los wavelets son algunos de los métodos no paramétricos para estimar la densidad. Sin embargo, a partir de los trabajos de Rosenblatt (1956) y Parzen (1962), los estimadores tipo núcleo han sido quizás los más populares en la estimación de la densidad. Dada una muestra aleatoria simple $\{X_i\}_{i=1}^n$ de la variable de interés, el estimador tipo núcleo de la función de densidad viene dado por:

$$\hat{f}_h(x) = \frac{1}{n} \sum_{i=1}^n K_h(x - X_i)$$

donde $K(\cdot)$ es una función real, llamada núcleo, con integral igual a 1, y $K_h(\cdot) = h^{-1}K(\cdot/h)$ es la función núcleo reescalada por un parámetro de suavizado, h , positivo, llamado comúnmente *ventana*.

Éste es el procedimiento que hemos elegido para estimar la función de densidad, por ser uno de los más utilizados y estudiados en la práctica. Además, presenta la ventaja de una formulación algebraica relativamente fácil de tratar. El problema principal de este estimador sería la elección del parámetro ventana, h , porque regula el grado de suavización:

- Una ventana pequeña conduce a que el estimador utilice pocas observaciones en la ponderación, disminuyendo el sesgo, pero haciendo que las estimaciones dependan demasiado de la variabilidad muestral. Esta situación recibe el nombre de *infrasuavización*.
- Una ventana grande produce un aumento del sesgo, puesto que hace uso de muchas observaciones en el promedio en el punto x de interés, lo que se conoce como *sobresuavización*.

Por lo tanto, es imprescindible el desarrollo y uso de métodos automáticos de elección de la ventana de acuerdo a algún criterio específico.

Nuestro objeto es, por tanto, estimar de forma no paramétrica las funciones de distribución y densidad en el campo del análisis de supervivencia. Sin embargo, el

procedimiento de obtención de información empírica de datos que miden el tiempo transcurrido hasta un suceso puntual no siempre es fácil. Para ello sería necesario observar el instante del suceso inicial de todos los individuos de la muestra, y esperar a que se produzca el suceso final igualmente para todos ellos. No obstante, por consideraciones de tiempo o de coste, en muchas ocasiones el investigador finaliza el estudio antes de que se produzca el suceso final para todos los individuos. Esta situación se conoce como *censura por la derecha*, y es en la que nos centraremos en esta memoria. Es frecuente, por ejemplo, en estudios de fiabilidad sobre la duración de una determinada componente, en los que no se espera al fallo de la misma. También en Medicina, puesto que algunos pacientes pueden sobrevivir al final del ensayo clínico, pasar a otro tratamiento, o morir por otra causa ajena a la enfermedad.

Aunque no todos los datos están completos, está claro que todos los datos disponibles aportan información sobre la distribución de la variable de interés. Se sabe que los tiempos de vida de los individuos censurados (aquellos para los que no se había producido el suceso final en el momento de finalizar el estudio) es siempre superior al valor observado de la variable. Por ello, los métodos clásicos de inferencia estadística no son eficientes en este contexto, y se hace necesario considerar nuevos métodos que se ajusten a esta situación y que optimicen el uso de la información contenida en la muestra.

En esta memoria se presenta un nuevo método de estimación, para las funciones de distribución y densidad, en presencia de censura aleatoria por la derecha. La novedad de este método radica en la estimación previa, mediante procedimientos no paramétricos, de la función probabilidad condicional de no censura. Este paso preliminar, llamado *presuavización*, da lugar a un uso más eficiente de la información aportada en los datos que mediante los métodos clásicos que se aplican en este contexto.

La memoria consta de seis capítulos, que se organizan de la siguiente manera. En el primero de ellos, se presenta el modelo de censura con el que trabajaremos, y se hace una revisión de los principales estimadores de las funciones de distribución y densidad para datos censurados, y de sus propiedades más relevantes.

El segundo capítulo introduce la presuavización como paso previo para estimar la función de distribución. Para este nuevo estimador de la función de distribución, se exponen sus principales propiedades asintóticas y se compara, mediante un estudio de simulación, con el estimador clásico para datos censurados, el llamado estimador límite producto o de Kaplan-Meier.

En el tercer capítulo se propone el estimador presuavizado tipo núcleo de la función de densidad, obtenido por convolución del estimador presuavizado de la función de

distribución, y se presentan sus propiedades asintóticas. Su comportamiento práctico se ilustra por medio de un estudio de simulación.

El capítulo 4 aborda el problema de selección de los parámetros ventana para el estimador presuavizado de la función de densidad. En él se proponen dos tipos de selectores, unos de tipo plug-in y otros bootstrap, se demuestra que el selector plug-in es consistente, y finalmente se estudia el comportamiento de todos ellos en un estudio de simulación.

La presuavización consiste en la estimación no paramétrica de la probabilidad condicional de no censura, que es una función de regresión. En el capítulo 5 se analiza el efecto, en la estimación final de la densidad, de presuavizar utilizando dos métodos diferentes para ajustar la función de regresión. En concreto, los métodos considerados han sido el estimador de Nadaraya-Watson y el estimador local lineal.

El sexto y último capítulo de este trabajo ilustra el comportamiento de los estimadores presuavizados de las funciones de distribución y densidad, con los datos de trasplante de corazón de Stanford. Se estudia, además, el efecto de distintas covariables en el tiempo de vida de los pacientes.

Capítulo 1

Estimación en el modelo de censura aleatoria por la derecha

1.1. Introducción a los modelos de censura

Tal como se manifestó en la introducción de esta memoria, el término de censura hace referencia a un tipo de pérdida de información en situaciones en las que la variable de interés es un tiempo de vida. La censura surge en las ocasiones en las que hay individuos de la muestra para los que no se conoce exactamente su tiempo de vida, sino que únicamente se sabe que éste ha ocurrido dentro de un cierto intervalo de valores. De esta forma se pueden considerar varios tipos de censura: censura por la derecha, por la izquierda y censura dentro de un intervalo.

Si la variable de interés es el tiempo transcurrido desde que ocurre un suceso inicial hasta que ocurre un suceso final o suceso de interés, comúnmente llamado *fallo*, entonces se asume que hay *censura por la derecha* cuando en el momento en que finaliza el estudio hay sujetos para los que no se conoce el instante exacto de fallo, sino que solamente se conoce que ha sido posterior a un momento dado. Por tanto, el valor exacto del tiempo de vida será superior al valor observado. Este problema es habitual cuando se analizan tiempos de vida, puesto que los estudios pueden terminar antes del fallo de todos los individuos de la muestra. Aquellos sujetos que no hayan fallado antes del fin del estudio serán censurados por la derecha, puesto que sólo sabremos que, de continuar el estudio, su instante de fallo sería posterior al instante final del estudio, y por tanto su tiempo de vida sería mayor que el observado. Lo mismo ocurre cuando no se puede observar el instante de fallo debido a la pérdida de seguimiento del individuo. Esto puede ser, entre otras causas, por un fallo debido a alguna razón ajena a la de interés, el abandono del estudio por parte del individuo, cambio de domicilio, etc.

Análogamente, el tiempo de vida asociado a un individuo en estudio se considera *censurado por la izquierda* si es menor que cierto valor dado, es decir, si el momento exacto en el que ocurrió el fallo es desconocido, sabiendo tan sólo que ha ocurrido antes de que el individuo se incluya en el estudio. Por ejemplo, es posible encontrarse en la muestra con sujetos que ya hayan fallado antes del comienzo del estudio, sin saber exactamente cuándo.

Un tipo más general de censura que generaliza los dos anteriores surge cuando de alguno de los tiempos de vida sólo se conoce que pertenecen a cierto intervalo. Se habla entonces de *censura de tipo intervalo*. Es común en estudios donde se hace un seguimiento periódico a los individuos. En este caso, cuando se encuentra un fallo para un individuo, solamente se sabe que el suceso de interés, el fallo, ocurrió entre dos revisiones periódicas.

De todos estos tipos de censura, el más común y en el que nos centraremos a partir de ahora, es el de censura por la derecha. Ahora bien, como hemos visto en los ejemplos anteriores, las causas que originan la censura de una observación pueden ser aleatorias o controladas; esto hace que se distinga entre tres clases de censura:

Censura tipo I: El suceso se observa si ocurre antes de un momento fijo predefinido C . En este caso, C es una constante prefijada por el investigador para todas las unidades muestrales. Este tipo de censura es común cuando, por diversas causas, el investigador finaliza el estudio antes de que todos los individuos hayan experimentado el suceso de interés. Si no hay pérdidas accidentales, todas las observaciones censuradas son iguales a la longitud del periodo en estudio.

Censura tipo II: Este tipo de censura surge cuando se fija el final del estudio en el momento en que un número $r < n$ predeterminado de individuos falla. Los tiempos de vida observados son los r menores valores de la muestra, de forma que C se convierte en la variable aleatoria $C = T_{(r)}$.

Censura tipo III: En la mayoría de los estudios, se fija la duración y los individuos entran a formar parte de la muestra a lo largo de ese periodo. Para los individuos que fallan antes del final del estudio, se conocen exactamente sus tiempos de vida. Para los que no han experimentado el suceso al final del estudio, la censura de sus tiempos de vida es semejante a la de tipo I. En ocasiones, algunos sujetos experimentan otros sucesos independientes del de interés que provocan su eliminación del estudio. Esta situación se denomina también *censura aleatoria*. En este tipo de censura, C es una variable aleatoria que se supone independiente de la variable de interés.

En esta memoria nos centraremos en el estudio de la censura aleatoria por la derecha.

1.1.1. El modelo de censura aleatoria por la derecha

Denotaremos por Y la variable aleatoria de interés tiempo de vida con función de densidad $f(\cdot)$ y función de distribución $F(\cdot)$. En presencia de censura aleatoria por la derecha, lo que realmente observa el estadístico para cada unidad muestral es la variable aleatoria bidimensional (Z, δ) definida por:

$$Z = \min(Y, C) \quad \text{y} \quad \delta = \mathbf{1}_{\{Y \leq C\}},$$

siendo C la variable de censura por la derecha y $\mathbf{1}_A$ la función indicadora del suceso A , es decir:

$$Z = \begin{cases} Y & \text{si } Y \leq C \\ C & \text{si } Y \geq C \end{cases} \quad \text{y} \quad \delta = \begin{cases} 1 & \text{si la observación no es censurada } (Z = Y) \\ 0 & \text{si la observación es censurada } (Z = C) \end{cases}$$

Bajo este modelo, C es una variable aleatoria con función de distribución $G(t) = P(C \leq t)$ y función de densidad $g(\cdot)$, que supondremos independiente de Y . Por lo tanto, la variable observada Z tiene función de distribución $H(\cdot)$ dada por:

$$1 - H(t) = (1 - F(t))(1 - G(t)). \quad (1.1)$$

Supondremos además, sin pérdida de generalidad, que las variables aleatorias son positivas, y adoptaremos la siguiente notación:

$$a_F = \inf \{t > 0 : F(t) > 0\} \quad \text{y} \quad b_F = \sup \{t > 0 : F(t) < 1\}$$

para representar los extremos inferior y superior del soporte de la función de distribución $F(\cdot)$. Extenderemos esta notación no sólo para la función $F(\cdot)$, sino también para las funciones de distribución $G(\cdot)$ y $H(\cdot)$: a_G, b_G, a_H y b_H respectivamente. Bajo la relación (1.1), estos extremos verifican lo siguiente:

$$a_H = \min \{a_F, a_G\} \quad \text{y} \quad b_H = \min \{b_F, b_G\}.$$

Asociadas a este modelo necesitaremos definir algunas funciones de interés. En primer lugar, sea $p(t)$ la probabilidad condicional de que una observación sea no censurada condicionado a que $Z = t$:

$$p(t) = P(\delta = 1 | Z = t) = \mathbb{E}(\delta | Z = t).$$

Junto a esta probabilidad, definimos ahora la probabilidad incondicional γ de que una observación sea no censurada:

$$\gamma = P(\delta = 1) = \mathbb{E}(\delta) = P(Z \leq C) = \int_{a_F}^{\infty} (1 - G(v)) dF(v) = H^1(+\infty),$$

donde $H^1(t) = P(Z \leq t, \delta = 1)$ es la función de subdistribución de las observaciones sin censura, que se puede escribir de la siguiente forma

$$\begin{aligned} \int_{a_H}^t p(v) dH(v) &= P(Z \leq t, \delta = 1) = H^1(t) = P(Y \leq t, Y \leq C) \\ &= \int_{a_F}^t (1 - G(v^-)) dF(v), \end{aligned}$$

es decir,

$$dH^1(t) = p(t) dH(t).$$

Análogamente, la función de subdistribución de las observaciones censuradas es

$$H^0(t) = P(Z \leq t, \delta = 0).$$

1.2. Estimación de la función de distribución

1.2.1. El estimador de Kaplan-Meier

El estimador de la función de supervivencia para datos censurados aleatoriamente por la derecha más usado y estudiado en la práctica es el propuesto por Kaplan y Meier (1958), también conocido en la literatura como *estimador límite-producto*. Su expresión se puede derivar siguiendo distintos métodos, algunos más intuitivos que otros, aunque aquí se expondrá la que se obtiene a partir de la siguiente relación entre la función de supervivencia y la razón de fallo acumulada:

$$1 - F(t) = \exp[-\Lambda_F(t)], \quad (1.3)$$

puesto que es la idea que seguiremos para obtener la expresión del estimador pre-suavizado de la función de distribución en el siguiente capítulo. Esta relación se puede generalizar, al caso en el que la función $\Lambda_F(\cdot)$ presente discontinuidades, de la siguiente manera:

$$1 - F(t) = \exp[-\Lambda_F^c(t)] \prod_{\{a_i \in A / a_i \leq t\}} (1 - \Lambda_F\{a_i\}),$$

donde $\Lambda_F^c(\cdot)$ denota la parte continua de $\Lambda_F(\cdot)$, A el conjunto de puntos donde $\Lambda_F(\cdot)$

tiene discontinuidades de salto y $\Lambda_F \{a_i\} = \Lambda_F(a_i) - \Lambda_F(a_i^-)$ es la magnitud del salto de $\Lambda_F(\cdot)$ en a_i . Puede verse la demostración de esta relación en la página 301 de Shorack y Wellner (1986).

En este modelo de censura, podemos escribir la razón de fallo acumulada $\Lambda_F(\cdot)$ en función de cantidades estimables empíricamente de la siguiente manera:

$$\begin{aligned}\Lambda_F(t) &= \int_0^t \lambda(v) dv = \int_0^t \frac{dF(v)}{1 - F(v^-)} = \int_0^t \frac{1 - G(v^-)}{1 - H(v^-)} dF(v) \\ &= \int_0^t \frac{dH^1(v)}{1 - H(v^-)} = \int_0^t \frac{p(v)}{1 - H(v^-)} dH(v)\end{aligned}$$

para todo $t < b_H$. Si sustituimos las funciones $H(\cdot)$ y $H^1(\cdot)$ en la expresión anterior por sus respectivas estimaciones empíricas,

$$H_n(t) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{\{Z_i \leq t\}} \quad \text{y} \quad H_n^1(t) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{\{Z_i \leq t\}} \delta_i,$$

y tomamos $H_n(t^-) = \lim_{x \uparrow t} H_n(x)$, obtenemos el estimador no paramétrico de la razón de fallo acumulada más utilizado en este contexto, el conocido estimador de Nelson-Aalen:

$$\Lambda_n^{NA}(t) = \int_0^t \frac{dH_n^1(v)}{1 - H_n(v^-)} = \sum_{i=1}^n \frac{\mathbf{1}_{\{Z_i \leq t, \delta_i=1\}}}{n(1 - H_n(Z_i^-))} = \sum_{Z_{(i)} \leq t} \frac{\delta_{[i]}}{n - i + 1}, \quad (1.4)$$

donde $\{Z_{(i)}, \delta_{[i]}\}_{i=1}^n$ son las observaciones ordenadas, y $\delta_{[i]}$'s son los concomitantes correspondientes a los indicadores de no censura. Notemos que precisamente $n\delta_{[i]}/(n - i + 1)$ es la estimación empírica de $p(v)/(1 - H(v^-))$. Retomaremos esta idea para introducir los estimadores presuavizados en la siguiente sección.

Este estimador fue sugerido por primera vez por Nelson (1972) en el contexto de la fiabilidad, y redescubierto por Aalen (1978), quien lo obtuvo usando técnicas de procesos de contar. Detalles de estas técnicas pueden encontrarse también en Andersen *et al.* (1993) y en Fleming y Harrington (1991).

Se puede modificar la expresión (1.4) del estimador de Nelson-Aalen para permitir más de un fallo en un instante t . Supongamos que los sucesos ocurren en D tiempos distintos $t_1 < \dots < t_D$, y que en el instante t_i hay d_i sucesos o fallos, siendo N_i el número de individuos en riesgo, es decir, el número de individuos vivos en t_i . En el caso de que sólo haya un fallo en cada instante, entonces $d_i = \delta_{[i]}$ y $N_i = n - i + 1$.

El cociente d_i/N_i proporciona una estimación de la probabilidad condicionada de que un individuo que sobrevive hasta justo antes del instante t_i , falle en el instante t_i . Ésta es la cantidad básica a partir de la cual se construyen los estimadores de la función de supervivencia y de la razón de fallo acumulada. En este caso, el estimador de Nelson-Aalen se puede escribir de la forma:

$$\Lambda_n^{NA(2)}(t) = \sum_{Z_{(i)} \leq t} \frac{d_i}{N_i}.$$

A partir de la igualdad (1.3), resulta el siguiente estimador de la función de supervivencia en el instante t :

$$1 - \widehat{F}(t) = \exp[-\Lambda_n^{NA}(t)] = \exp\left[-\sum_{Z_{(i)} \leq t} \frac{\delta_{[i]}}{n-i+1}\right] = \prod_{Z_{(i)} \leq t} \exp\left[-\frac{\delta_{[i]}}{n-i+1}\right].$$

Podemos obtener una expresión de este estimador más cómoda de calcular si usamos la aproximación $e^{-t} \simeq 1 - t$ para t próximos a 0, dando lugar a la expresión final del estimador de Kaplan y Meier de la función de supervivencia:

$$1 - F_n^{KM}(t) = \prod_{Z_{(i)} \leq t} \left(1 - \frac{\delta_{[i]}}{n-i+1}\right). \quad (1.5)$$

Análogamente, en el caso de permitir más de un fallo en un instante t , el estimador de la función de supervivencia sería

$$1 - F_n^{KM(2)}(t) = \prod_{Z_{(i)} \leq t} \left(1 - \frac{d_i}{N_i}\right).$$

Se puede comprobar que el estimador de Nelson-Aalen de la razón de fallo acumulada es el primer término de la serie de Taylor de menos el logaritmo del estimador de Kaplan-Meier de la función de supervivencia. En concreto, la relación entre los estimadores de Kaplan-Meier y Nelson-Aalen es :

$$1 - F_n^{KM}(t) = \exp[-\Lambda_n^{NA}(t)] + O_P(n^{-1}).$$

Observación 1.2.1

En lugar de usar la aproximación

$$\exp\left(-\frac{\delta_{[i]}}{n-i+1}\right) \simeq 1 - \frac{\delta_{[i]}}{n-i+1}$$

de la que se obtiene el estimador (1.5), podemos considerar la aproximación

$$\exp\left(-\frac{\delta_{[i]}}{n-i+1}\right) = \left[\exp\left(-\frac{1}{n-i+1}\right)\right]^{\delta_{[i]}} \simeq \left(1 - \frac{1}{n-i+1}\right)^{\delta_{[i]}} ,$$

que da lugar a la siguiente expresión del estimador de Kaplan-Meier:

$$1 - F_n^{KM}(t) = \prod_{Z_{(i)} \leq t} \left(\frac{n-i}{n-i+1}\right)^{\delta_{[i]}} ,$$

equivalente a la dada en (1.5).

1.2.2. Propiedades del estimador de Kaplan-Meier

El estimador de Kaplan-Meier presenta varias propiedades deseables: su facilidad de cálculo y el hecho de que sea el estimador no paramétrico de máxima verosimilitud para datos censurados (Johansen (1978), Scholz (1980) o Wang (1987)) hacen que sea el estimador más utilizado y estudiado en este contexto. Además, se reduce al estimador empírico clásico en el caso de no haber censura.

Sin embargo, este estimador presenta problemas cuando la hipótesis de independencia entre los tiempos de fallo Y y los tiempos de censura C no se verifica. Por otra parte, tiene forma escalonada con saltos únicamente en las observaciones no censuradas, con pesos que aumentan desde el menor al mayor dato no censurado, puesto que dependen del número de observaciones censuradas entre ellos (ver Efron (1967)).

El estimador de Kaplan-Meier está bien definido para todos los valores de t menores que el mayor tiempo en estudio observado. No obstante, si el mayor valor observado corresponde a un tiempo de vida no censurado, entonces la curva de supervivencia para valores de t posteriores es 0. Pero si la última observación es censurada, el valor de $1 - F(\cdot)$ para tiempos posteriores es indeterminado porque no se puede saber cuándo este último individuo hubiera fallado de no haber sido censurado. Se han dado varias soluciones a esta ambigüedad: Efron (1967) propuso estimar $1 - F(t)$ para $t > Z_{(n)}$ por el valor 0. Esto equivale a suponer que el individuo con el mayor tiempo de vida hubiera fallado inmediatamente después de haber sido censurado, y da lugar a un estimador negativamente sesgado. Gill (1980) sugirió estimar $1 - F(t)$ para $t > Z_{(n)}$ por $1 - F_n^{KM}(Z_{(n)})$, que corresponde a asumir que este individuo fallaría en $t = \infty$ y conduce a un estimador con sesgo positivo. Aunque las dos propuestas tienen las

mismas propiedades asintóticas y convergen a la verdadera función de supervivencia, un estudio con tamaños muestrales finitos de ambos estimadores realizado por Klein (1991) revela un mejor comportamiento de la versión de Gill del estimador de Kaplan-Meier.

Como ya se ha comentado con anterioridad, el estimador de Kaplan-Meier ha sido el estimador de la función de supervivencia en presencia de censura más estudiado en los últimos años. Expondremos a continuación algunas de sus propiedades asintóticas más destacables.

El punto de referencia es, sin duda, el artículo de Kaplan y Meier (1958), en el que estudian y comparan el estimador límite-producto con otros dos estimadores: los estimadores de tipo actuarial y el estimador de muestra reducida.

Phadia y Shao (1999) obtuvieron las expresiones exactas de los momentos k -ésimos del estimador de Kaplan-Meier, tanto de la versión de Efron como de la de Gill, sin suponer ningún modelo previo de la distribución de la supervivencia. El principal resultado para la versión de Efron del estimador límite producto es el siguiente:

Propiedad 1.2.1 (Teorema 1 de Phadia y Shao (1999))

Sea $\overline{F}_n^{KM}(\cdot)$ el estimador de Kaplan-Meier de la función de supervivencia $\overline{F}(\cdot) = 1 - F(\cdot)$. Entonces, el momento k -ésimo de $\overline{F}_n^{KM}(\cdot)$ viene dado por

$$\mathbb{E} \left[(\overline{F}_n^{KM})^k(t) \right] = \sum_{i=0}^{n-1} \frac{n!}{(n-i)!} \overline{H}^{n-i}(t) \int_0^t \int_0^{t_i} \dots \int_0^{t_2} \prod_{j \leq i} d\phi_j(t_j), \quad (1.6)$$

donde $0 < t_1 < t_2 < \dots < t_i \leq t$ y

$$\phi_j(t) = H(t) - H^1(t) + H^1(t) \left(\frac{n-j}{n-j+1} \right)^k,$$

siendo $H^1(t) = P(Z \leq t, \delta = 1)$. El producto se calcula sobre $j = 1, 2, \dots, i$. Para $i = 0$, el producto se define como 1.

En el caso de la versión de Gill del estimador de Kaplan-Meier, el sumatorio de la expresión (1.6) se extiende hasta el sumando n .

Puesto que la fórmula anterior no es manejable en aplicaciones prácticas, se propuso la siguiente aproximación:

Propiedad 1.2.2 (Teorema 2 de Phadia y Shao (1999))

Si se aproxima cada función $\phi_j(x)$ por su componente lineal $[(\phi_j(t) - \phi_j(0)) / t] x$, en el intervalo $(0, t)$, y se sustituye en la expresión (1.6), se obtiene

$$\mathbb{E} \left[(\bar{F}_n^{KM})^k(t) \right] \simeq \sum_{i=0}^{n-1} \binom{n}{i} \bar{H}^{n-i}(t) \prod_{j \leq i} \left[H(t) - H^1(t) + H^1(t) \left(\frac{n-j}{n-j+1} \right)^k \right].$$

Meier (1975) establece para el estimador límite producto un teorema análogo al de Glivenko-Cantelli y prueba que es asintóticamente insesgado, la normalidad asintótica en cada punto y la convergencia débil a un movimiento Browniano apropiadamente transformado. Paralelamente, Breslow y Crowley (1974) obtuvieron la normalidad asintótica puntual y sobre intervalos compactos del proceso. La extensión de este teorema a toda la recta real se puede encontrar en Gill (1983).

Propiedad 1.2.3 (Normalidad asintótica puntual y sobre intervalos compactos, Breslow y Cowley (1974))

Sean las funciones de distribución $F(\cdot)$ y $G(\cdot)$ continuas. Entonces:

a) Para todo $0 < t < b_H$,

$$\sqrt{n}(F_n^{KM}(t) - F(t)) \xrightarrow{d} N(0, \sigma(t))$$

donde

$$\sigma^2(t) = (1 - F(t))^2 \int_0^t (1 - H(v))^{-2} dH^1(v).$$

b) El proceso estocástico $\sqrt{n}(F_n^{KM} - F)$ converge globalmente en $\mathcal{D}[0, T]$ para cada $T < b_H$ a un proceso Gaussiano $Z(\cdot)$

$$\sqrt{n}(F_n^{KM} - F) \xrightarrow{d} Z(\cdot)$$

con media 0 y función de covarianzas:

$$\text{Cov}(Z(x), Z(t)) = (1 - F(x))(1 - F(t)) \int_0^{x \wedge t} (1 - H(v))^{-2} dH^1(v),$$

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

También a principios de los años 80, diversos autores estudiaron la consistencia del estimador de Kaplan-Meier, entre ellos Földes y Rejtő (1980).

Propiedad 1.2.4 (*Consistencia uniforme fuerte, Földes y Rejtö (1980)*)

a) Sea $0 < T < b_H$. Entonces

$$|F_n^{KM} - F| \rightarrow 0 \quad \text{c.s. uniformemente en } [0, T].$$

b) Si además $G(b_F^-) > 0$, donde $b_F = \sup \{t : F(t) < 1\}$, entonces

$$|F_n^{KM} - F| \rightarrow 0 \quad \text{c.s. uniformemente en } \mathbb{R}.$$

Además, se establecen también los órdenes de convergencia, siendo $O(n^{-1/4}(\log n)^{1/2})$. Los mismos autores, Földes y Rejtö, en 1981 mejoran la razón de convergencia anterior cuando $G(\cdot)$ pone toda su masa en un número finito de puntos, obteniendo el siguiente resultado:

$$\sup_{0 \leq t \leq T} |F_n^{KM}(t) - F(t)| = O\left(n^{-1/2}(\log n)^{1/2}\right) \quad \text{c.s.}$$

Varios autores estudiaron la *ley del logaritmo iterado* (LIL) del estimador, entre ellos Csörgö y Horváth (1983), quienes obtuvieron el siguiente resultado:

Propiedad 1.2.5 (*Ley del logaritmo iterado, Csörgö y Horváth (1983)*)

Si $F(T) < 1$, entonces

$$\limsup_{n \rightarrow \infty} \left(\frac{n}{2 \log \log n} \right)^{1/2} \sup_{0 \leq t \leq T} |F_n^{KM}(t) - F(t)| \leq \frac{1}{1 - H(T)}.$$

Lo y Singh (1986) establecieron una aproximación fuerte uniforme de la diferencia entre estimador de Kaplan-Meier $F_n^{KM}(\cdot)$ y la función de distribución teórica $F(\cdot)$ como una media de variables aleatorias independientes idénticamente distribuidas y acotadas más un término despreciable de orden conocido. También proporcionan algunas versiones bootstrap de estas representaciones. La importancia de estos desarrollos radica en que permiten trabajar con una suma de variables i.i.d., mucho más manejable que el producto por el que viene dado el estimador $F_n^{KM}(\cdot)$, y obtener propiedades tales como la normalidad asintótica y la convergencia del proceso.

Propiedad 1.2.6 (*Representación casi segura, Lo y Singh (1986)*)

Bajo la hipótesis de que $F(\cdot)$ y $G(\cdot)$ son continuas, se puede escribir, para todo $t \leq T < b_H$:

$$F_n^{KM}(t) - F(t) = n^{-1} \sum_{i=1}^n \xi(Z_i, \delta_i, t) + r_n(t),$$

donde

$$\xi(Z, \delta, t) = (1 - F(t)) \left[g(Z \wedge t) + \frac{1}{1 - H(Z)} \mathbf{1}_{\{Z \leq t, \delta=1\}} \right],$$

siendo

$$g(t) = \int_0^t (1 - H(v))^{-2} d(1 - H^1(v))$$

y

$$\sup_{0 \leq t \leq T} |r_n(t)| = O \left(\left(\frac{\log n}{n} \right)^{3/4} \right) \quad c.s.$$

Además, para todo $\alpha \geq 1$,

$$\sup_{0 \leq t \leq T} \mathbb{E} |r_n(t)|^\alpha = O \left(\left(\frac{\log n}{n} \right)^{3\alpha/4} \right).$$

El orden del término $r_n(\cdot)$ es suficiente para probar la mayoría de las propiedades asintóticas del estimador de Kaplan-Meier, pero se queda pequeño cuando se estudian estimadores de la densidad o razón de fallo derivados de $F_n^{KM}(\cdot)$. Fue posteriormente mejorado hasta $O(n^{-1} \log n)$ *c.s.* por Lo, Mack y Wang (1989).

Mientras las propiedades asintóticas del estimador de Kaplan-Meier han sido ampliamente investigadas y estudiadas, no ha ocurrido así con sus propiedades para tamaños muestrales finitos. Una excepción ha sido el artículo de Chen, Hollander y Langberg (1982), en el que se obtuvieron los momentos exactos del estimador, aunque bajo el modelo de Koziol-Green (ver Abdushukurov (1987) y Cheng y Lin (1987)). Posteriormente Guerts (1985, 1987) consideró dos versiones del estimador límite-producto, las propuestas por Efron (1967) y Gill (1980), y estudió los comportamientos para tamaños muestrales finitos de los sesgos y de los errores cuadráticos medios en algunos modelos de riesgos proporcionales mediante simulación.

Por último, hemos de destacar que en los últimos años se han multiplicado los resultados sobre las integrales Kaplan-Meier. Éstas tienen la siguiente forma

$$\int \varphi dF_n^{KM},$$

donde $\varphi(\cdot)$ es una función Borel-medible tal que $\int |\varphi| dF_n^{KM} < \infty$. El interés de estas integrales se basa en que muchos de los parámetros o funciones de interés en Estadística se pueden escribir en forma de este tipo de integral sin más que elegir de una forma adecuada la función $\varphi(\cdot)$. Si X es una variable aleatoria con función de distribución $F(\cdot)$, tenemos lo siguiente:

1. Si $\varphi(x) = \mathbf{1}_{(-\infty, t]}(x)$, entonces $\int \varphi(v) dF(v)$ es la función de distribución $F(t)$.
2. Si $\varphi(x) = x^k$, entonces $\int \varphi(v) dF(v)$ es el momento de orden k , $\mathbb{E}(X^k)$.
3. Si $\varphi(x) = \exp(itx)$, entonces $\int \varphi(v) dF(v)$ es la función característica de X .
4. Si $\varphi(x) = (x - t) \mathbf{1}_{\{x > t\}}$ y $F(t) < 1$, entonces

$$\frac{\int \varphi(v) dF(v)}{1 - F(t)}$$

es la vida media residual en el instante t .

Stute y Wang (1993) han demostrado la convergencia casi segura y en media de estas integrales. El resultado viene dado en la siguiente propiedad:

Propiedad 1.2.7 (*Consistencia de la integral Kaplan-Meier, Stute y Wang (1993)*)

Supongamos que las funciones $F(\cdot)$ y $G(\cdot)$ no tienen saltos en común, y sea A el conjunto de todos los valores discretos de $H(\cdot)$. Entonces, con probabilidad uno y en media,

$$\begin{aligned} \lim_{n \rightarrow \infty} \int \varphi(v) dF_n^{KM}(v) &= \int_{\{s \notin A, s < b_H\}} \varphi(v) dF(v) + \sum_{a_i \in A} \varphi(a_i) F\{a_i\} \\ &= \int_{\{s < b_H\}} \varphi(v) dF(v) + \mathbf{1}_{\{b_H \in A\}} \varphi(b_H) F\{b_H\} \end{aligned}$$

donde $F\{a\} = F(a) - F(a^-)$.

En particular, si $F(\cdot)$ es continua y $b_F = b_H$, entonces la expresión anterior se reduce a

$$\lim_{n \rightarrow \infty} \int \varphi(v) dF_n^{KM}(v) = \int \varphi(v) dF(v)$$

Este resultado permite derivar propiedades tan importantes como las siguientes:

1. Si $\varphi(x) = \mathbf{1}_{(-\infty, t]}(x)$, entonces tenemos la consistencia fuerte puntual para $F_n^{KM}(t)$, con $t \leq b_H$:

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

2. La consistencia fuerte uniforme de $F_n^{KM}(\cdot)$ en el intervalo $[0, b_H]$:

$$\sup_{0 \leq t \leq b_H} |F_n^{KM}(t) - F(t)| \rightarrow 0 \quad c.s.$$

3. Si $\varphi(x) = x^k$, entonces obtendremos la consistencia fuerte de los estimadores de los momentos:

$$\lim_{n \rightarrow \infty} \int s^k dF_n^{KM}(v) = \mathbb{E}(Y^k) \quad c.s.$$

4. Si $\varphi(x) = \exp(itx)$, entonces se deriva la convergencia puntual de la función característica de Y .

5. Si $\varphi(x) = (x - t) \mathbf{1}_{\{x > t\}}$ y $F(t) < 1$, entonces podemos concluir que

$$\frac{\int_{\{v > t\}} (v - t) dF_n^{KM}(v)}{1 - F_n^{KM}(t)} \longrightarrow \frac{\int_{\{v > t\}} (v - t) dF(v)}{1 - F(t)} \quad \text{con probabilidad 1}$$

para la vida media residual en el instante t .

Con respecto a los momentos de las integrales de Kaplan-Meier, Stute (1994) obtuvo bajo ciertas condiciones sobre la función $\varphi(\cdot)$ una acotación tanto superior como inferior para el sesgo de estas integrales, junto con la expresión exacta. En lo que se refiere a la varianza, Stute (1996) demostró que bajo ciertas condiciones de integrabilidad de la función $\varphi(\cdot)$, el estimador jackknife de la varianza estima consistentemente la varianza límite. Finalmente, en Stute (1995) se puede encontrar la demostración de que la integral $\int \varphi(v) dF_n^{KM}(v)$, adecuadamente estandarizada, es asintóticamente normal.

Propiedad 1.2.8 (Normalidad asintótica de la integral Kaplan-Meier, Stute (1995))

Bajo ciertas condiciones (bastante generales) de integrabilidad de la función $\varphi(\cdot)$, se verifica que

$$\sqrt{n} \int \varphi(v) d(F_n^{KM}(v) - F(v)) \xrightarrow{d} N(0, \sigma^2)$$

siendo

$$\sigma^2 = \text{Var}[\varphi(Z) \gamma_0(Z) \delta + \gamma_1(Z)(1 - \delta) - \gamma_2(Z)]$$

donde $\gamma_0(\cdot)$, $\gamma_1(\cdot)$ y $\gamma_2(\cdot)$ tienen una expresión complicada, pero en el caso en que las funciones $F(\cdot)$ y $G(\cdot)$ sean continuas, sus expresiones se reducen a

$$\begin{aligned} \gamma_0(t) &= \frac{1}{1 - G(t)}, \\ \gamma_1(t) &= \frac{1}{1 - H(t)} \int_t^{b_H} \varphi(v) dF(v), \\ \gamma_2(t) &= \int_t^{b_H} \varphi(v) \left[\int_t^{b_H} \mathbf{1}_{\{w < t, w < v\}} \frac{1 - F(w)}{(1 - H(w))^2} dG(w) \right] dF(v). \end{aligned}$$

También Stute (1995) obtuvo una representación de la integral de Kaplan-Meier como una suma de variables independientes más un término despreciable, en función de $\gamma_0(\cdot)$, $\gamma_1(\cdot)$ y $\gamma_2(\cdot)$.

Propiedad 1.2.9 (*Representación de la integral Kaplan-Meier como suma de variables independientes, Stute (1995)*)

Bajo las mismas condiciones de integrabilidad de la función $\varphi(\cdot)$ que en la propiedad anterior, tenemos que

$$\int \varphi dF_n^{KM} = \frac{1}{n} \sum_{i=1}^n \varphi(Z_i) \gamma_0(Z_i) \delta_i + \frac{1}{n} \sum_{i=1}^n \gamma_1(Z_i) (1 - \delta_i) - \frac{1}{n} \sum_{i=1}^n \gamma_2(Z_i) + R_n,$$

donde

$$R_n = o_P\left(n^{-1/2}\right).$$

Mucho más recientemente, Bae y Kim (2003a,b) generalizaron los resultados de Stute (1995) a los siguientes procesos:

$$U_n(\varphi) = \sqrt{n} \int \varphi d\left(F_n^{KM} - \tilde{F}\right).$$

Los autores consideran en lugar de la función de distribución teórica $F(\cdot)$ la siguiente subdistribución

$$\tilde{F}(t) = F(t) \mathbf{1}_{\{t < b_H\}} + \left\{ F(b_H^-) + \mathbf{1}_{\{b_H \in A\}} (F(b_H) - F(b_H^-)) \right\} \mathbf{1}_{\{t \geq b_H\}}$$

puesto que b_H puede no ser necesariamente finito, donde A es el conjunto de valores discretos de $H(\cdot)$. En el caso en que $b_H < \infty$, entonces $\tilde{F}(\cdot)$ se reduce a la función de distribución $F(\cdot)$. Los principales resultados están resumidos en las siguientes propiedades (ver Bae y Kim (2003a,b) para el detalle de algunas definiciones).

Propiedad 1.2.10 (*Consistencia uniforme de los procesos de integrales Kaplan-Meier, Bae y Kim (2003a)*)

Si $\mathcal{F} \subset \mathcal{L}_1(F) := \{\varphi : \int |\varphi| dF < \infty\}$ tiene una entropía de recubrimiento, y $F(\cdot)$ y $G(\cdot)$ no tienen saltos en común, entonces

$$\sup_{\varphi \in \mathcal{F}} \left| \int \varphi d\left(F_n^{KM} - \tilde{F}\right) \right| \rightarrow 0$$

en probabilidad y en media.

Propiedad 1.2.11 (Teorema central del límite uniforme de los procesos de integrales Kaplan-Meier, Bae y Kim (2003b))

Sea \mathcal{F} la clase de funciones reales medibles definidas en \mathbb{R} . Sea (\mathcal{F}, d) un espacio métrico, con $N_{[\cdot]}(u, \mathcal{F}, d)$ el menor n para el que existe $\{f_{0,\delta}^l, f_{0,\delta}^u, \dots, f_{n,\delta}^l, f_{n,\delta}^u\}$ tal que, para todo $f \in \mathcal{F}$ existe algún $0 \leq i \leq n$ para el que $f_{i,\delta}^l \leq f \leq f_{i,\delta}^u$ y $d(f_{i,\delta}^l, f_{i,\delta}^u) < \delta$, y sea

$$J(\delta) := \int_0^\delta [\log N_{[\cdot]}(u, \mathcal{F}, d)]^{1/2} du \quad \text{para } 0 < \delta \leq 1$$

la integral asociada de la entropía métrica de recubrimiento $\log N_{[\cdot]}(u, \mathcal{F}, d)$. Si $J(1) < \infty$ y \mathcal{F} tiene la propiedad puntual minimal, entonces el proceso $U_n(\varphi) = \sqrt{n} \int \varphi d(F_n^{KM} - \tilde{F})$ converge como elemento de $B(\mathcal{F})$, el espacio de las funciones acotadas de \mathcal{F} , a un proceso gaussiano $W(\varphi)$ con media $\mathbb{E}[W(\varphi)] = 0$ y función de covarianzas $\text{cov}(W(\varphi_1), W(\varphi_2)) = \text{Cov}(\xi(\varphi_1), \xi(\varphi_2))$, donde

$$\xi(\varphi) = \varphi(Z) \gamma(Z) \delta - \int \varphi d\tilde{F} + \gamma_1(Z)(1 - \delta) - \gamma_2(Z)$$

y

$$\begin{aligned} \gamma(t) &= \exp\left(\int_{-\infty}^{t^-} \frac{dH^0(v)}{1 - H(t)}\right) \text{ con } H^0(t) = P(Z \leq t, \delta = 0), \\ \gamma_1(t) &= \frac{1}{1 - H(t)} \int_t^{b_H} \varphi(v) \gamma(v) dH^1(v), \\ \gamma_2(t) &= \iint \mathbf{1}_{\{w < t, w < v\}} \frac{\varphi(v) \gamma(v)}{(1 - H(w))^2} dH^0(w) dH^1(v). \end{aligned}$$

Los mismos autores (ver Bae y Kim (2003c)) obtuvieron una ley del logaritmo iterado empírica para los procesos de integrales de Kaplan-Meier, que es la extensión al modelo de censura aleatoria de la ley del logaritmo iterado para datos completos.

1.2.3. El estimador semiparamétrico de Dikta

Los estimadores de Kaplan-Meier (1.5) para la función de supervivencia y de Nelson-Aalen (1.4) para la razón de fallo acumulada se basan en considerar la versión puramente empírica de

$$\Lambda_F(t) = \int_0^t \frac{p(v) dH(v)}{1 - H(v-)},$$

es decir, en el i -ésimo salto ordenado (de tamaño n^{-1}) de $H_n(t) = n^{-1} \sum_{i=1}^n \mathbf{1}_{\{Z_i \leq t\}}$, estimar el integrando $p(v) (1 - H(v-))^{-1}$ por el valor empírico

$$\frac{\delta_{[i]}}{1 - \frac{i-1}{n}} = \frac{n\delta_{[i]}}{n-i+1}$$

y obtener por tanto

$$\Lambda_n^{NA}(t) = \sum_{Z_{(i)} \leq t} \frac{\delta_{[i]}}{n-i+1} \quad \text{y} \quad 1 - F_n^{KM}(t) = \prod_{Z_{(i)} \leq t} \left(1 - \frac{\delta_{[i]}}{n-i+1}\right).$$

La idea de los estimadores semiparamétricos consiste en sustituir $p(\cdot)$ en las expresiones

$$\Lambda_F(t) = \int_0^t \frac{p(v)dH(v)}{1-H(v-)} \quad \text{y} \quad 1 - F(t) = \exp[-\Lambda_F(t)] = \exp\left(-\int_0^t p(v)d\Lambda_H(v)\right)$$

por un estimador paramétrico de $p(\cdot)$.

La función $p(\cdot)$ viene dada por

$$p(t) = P(\delta = 1|Z = t) = \mathbb{E}(\delta|Z = t),$$

es decir, es la probabilidad condicional de que la observación sea no censurada dado $Z = t$. Su importancia ya ha sido apuntada por Stute y Wang (1993) para probar la consistencia de las integrales de Kaplan-Meier $\int \varphi dF_n^{KM}$, y es muy útil a la hora de expresar los funcionales de la variable de interés Y como función de la variable observada $Z = \min(Y, C)$, por ejemplo:

$$\begin{aligned} \Lambda_F(t) &= \int_0^t p(v) d\Lambda_H(v), \\ \lambda_F(t) &= p(t) \lambda_H(t), \\ 1 - F(t) &= \exp\left(-\int_0^t p(v) d\Lambda_H(v)\right). \end{aligned}$$

Los estimadores semiparamétricos, introducidos por Dikta (1998), se obtienen estimando $p(\cdot)$ de una forma totalmente paramétrica, es decir, asumiendo que $p(t) = p(t, \theta_0)$, donde $p(\cdot, \cdot)$ es una función continua conocida y $\theta_0 = (\theta_{0,1}, \theta_{0,2}, \dots, \theta_{0,k}) \in \Theta$ es un parámetro k -dimensional desconocido. Para la estimación del parámetro θ_0 , Dikta (1998) usa el método de máxima verosimilitud bajo censura aleatoria, es decir, $\hat{\theta}_{0,MV}$ es el parámetro que maximiza

$$L_n(\theta) = \prod_{i=1}^n p(Z_i, \theta)^{\delta_i} (1 - p(Z_i, \theta))^{1-\delta_i}.$$

Sustituyendo, por tanto, $p(\cdot)$ por $p_n(\cdot) = p(\cdot, \hat{\theta}_{0,MV})$, obtendremos los estimadores semiparamétricos de la razón de fallo acumulativa y la función de supervivencia propuestos por Dikta (1998):

$$\Lambda_n^D(t) = \int_0^t \frac{p(v, \hat{\theta}_{0,MV})}{1 - H_n(v-)} dH_n(v) = \sum_{Z_{(i)} \leq t} \frac{p(Z_{(i)}, \hat{\theta}_{0,MV})}{n - i + 1}$$

y

$$1 - F_n^D(t) = \exp[-\Lambda_n^D(t)] = \prod_{Z_{(i)} \leq t} \left(1 - \frac{1}{n - i + 1}\right)^{p(Z_{(i)}, \hat{\theta}_{0,MV})} \quad (1.7)$$

Observación 1.2.2

Análogamente a lo que ocurría con el estimador de Kaplan-Meier, el estimador $F_n^D(\cdot)$ deriva de la expresión del estimador de la razón de fallo acumulativa $\Lambda_n^D(\cdot)$ mediante la aproximación $\exp(-x) \simeq 1 - x$ para valores de x pequeños. Si en lugar de usar la aproximación

$$\exp\left(-\frac{p(Z_{(i)}, \hat{\theta}_{0,MV})}{n - i + 1}\right) \simeq \left(1 - \frac{1}{n - i + 1}\right)^{p(Z_{(i)}, \hat{\theta}_{0,MV})}$$

de la que se obtiene el estimador (1.7), consideramos la aproximación:

$$\exp\left(-\frac{p(Z_{(i)}, \hat{\theta}_{0,MV})}{n - i + 1}\right) = 1 - \frac{p(Z_{(i)}, \hat{\theta}_{0,MV})}{n - i + 1},$$

obtendremos la siguiente expresión del estimador semiparamétrico:

$$1 - F_n^{D(2)}(t) = \prod_{Z_{(i)} \leq t} \left(1 - \frac{p(Z_{(i)}, \hat{\theta}_{0,MV})}{n - i + 1}\right).$$

Ambos estimadores tienen las mismas propiedades asintóticas. La principal diferencia de ambos es el peso que otorgan a la última observación $Z_{(n)}$, puesto que $F_n^D(Z_{(n)}) = 1$, mientras que $F_n^{D(2)}(Z_{(n)}) < 1$ en la mayoría de los casos.

El siguiente teorema presenta la convergencia en norma del supremo de los estimadores semiparamétricos de la razón de fallo acumulativa y la función de distribución:

Propiedad 1.2.12 (*Consistencia de los estimadores semiparamétricos, Dikta (1998)*)

Sea $0 < T < \infty$ con $H(T) < 1$ y Θ un subconjunto abierto y conexo de \mathbb{R}^k . Supongamos que $\hat{\theta}_{0,MV} \in \Theta$ es una solución medible de la ecuación $\text{Grad}(l_n(\theta)) = (D_1 l_n(\theta), \dots, D_k l_n(\theta)) = 0$, siendo $D_r l_n(\theta_0) = [\partial l_n(\theta) / \partial \theta_r] |_{\theta=\theta_0}$ y $l_n(\theta)$ la función de logverosimilitud normalizada de los datos, tal que $\hat{\theta}_{0,MV} \rightarrow \theta_0$ con probabilidad uno. Sea $p(\cdot, \theta)$ también una función con derivadas parciales continuas con respecto a θ para todo $\theta \in \Theta$ y $t \geq 0$, donde $D_r p(\cdot, \theta)$ es medible para todo $\theta \in \Theta$ y existe un entorno $V(\theta_0) \in \Theta$ de θ_0 y una función medible M tal que $|D_r p(t, \theta)| \leq M(t)$ y $\mathbb{E}[M(X)] < \infty$ para todo $\theta \in V(\theta_0)$, $t \geq 0$, y $1 \leq r \leq k$. Entonces, con probabilidad uno, cuando $n \rightarrow \infty$,

$$\sup_{0 \leq t \leq T} |\Lambda_n^D(t) - \Lambda_F(t)| \rightarrow 0,$$

$$\sup_{0 \leq t \leq T} |F_n^D(t) - F(t)| \rightarrow 0.$$

Bajo ciertas condiciones de la función de logverosimilitud $l_n(\theta)$ y de la función $p(\cdot, \theta)$, y suponiendo que $H(\cdot)$ es continua con $H(T) < 1$, Dikta (1998) demuestra que el proceso $n^{1/2}(\Lambda_n^D(\cdot) - \Lambda_F(\cdot))$ converge débilmente a un proceso Gaussiano centrado $S(\cdot)$, donde la estructura de covarianzas de $S(\cdot)$ viene dada para $0 \leq s \leq t \leq T$ por

$$\begin{aligned} \text{Cov}(S(s), S(t)) &= \int_0^s \frac{p(v, \theta_0)}{(1 - H(v))^2} dH^1(v) \\ &+ \int_0^s \int_0^t \frac{\alpha(x, y)}{(1 - H(x))(1 - H(y))} dH(x) dH(y), \end{aligned}$$

donde

$$\alpha(x, y) = \langle \text{Grad}(p(x, \theta_0)), I^{-1}(\theta_0) \text{Grad}(p(y, \theta_0)) \rangle$$

y

$$I(\theta_0) = (-\mathbb{E}(D_{r,s} \log(p(x, \theta))))_{1 \leq r, s \leq k} = \left(\mathbb{E} \left(\frac{D_r p(X, \theta_0) D_s p(X, \theta_0)}{p(X, \theta_0)(1 - p(X, \theta_0))} \right) \right)_{1 \leq r, s \leq k}.$$

Como consecuencia de este resultado, Dikta (1998) demuestra la convergencia débil del proceso $n^{1/2}(F_n^D(\cdot) - F(\cdot))$ al proceso Gaussiano centrado $(1 - F(\cdot))S(\cdot)$.

Si denotamos la varianza asintótica del proceso $n^{1/2}(F_n^D(\cdot) - F(\cdot))$ por $v^D(\cdot)$, y la de $n^{1/2}(F_n^{KM}(\cdot) - F(\cdot))$ por $v^{KM}(\cdot)$, entonces bajo ciertas condiciones sobre la función de logverosimilitud $l_n(\theta)$ y la función $p(\cdot, \theta)$, se verifica para $0 \leq t \leq T$

$$v^{KM}(t) - v^D(t) = (1 - F(t))^2 r(t) \geq 0,$$

donde

$$r(t) = \int_0^t \frac{1 - p(v, \theta_0)}{(1 - H(v))^2} dH^u(v) - \int_0^t \int_0^t \frac{\alpha(x, y)}{(1 - H(x))(1 - H(y))} dH(x)dH(y),$$

por lo que queda demostrada la mayor eficiencia asintótica del nuevo estimador semiparamétrico de la función de supervivencia con respecto al clásico estimador de Kaplan-Meier.

Las propiedades de las integrales de funciones Borel-medibles con respecto al estimador semiparamétrico de la función de densidad, $\int \varphi dF_n^D$, han sido estudiadas por Dikta (2000), quien demostró su consistencia.

Propiedad 1.2.13 (*Consistencia de las integrales semiparamétricas $\int \varphi dF_n^D$, Dikta (2000)*)

Supongamos que $H(\cdot)$ es continua, $\widehat{\theta}_{0, MV}$ es medible y tiende a θ_0 con probabilidad 1, que para cada $\varepsilon > 0$ existe un entorno $V(\varepsilon, \theta_0) \subset \Theta$ de θ_0 tal que para todo $\theta \in V$

$$\sup_{x \geq 0} |p(x, \theta) - p(x, \theta_0)| < \varepsilon.$$

Si además

$$\int_0^{b_H} \frac{|\varphi(t)|}{p(t, \theta_0)(1 - H(t))^\varepsilon} dF(t) < \infty$$

para algún $\varepsilon > 0$, entonces

$$\lim_{n \rightarrow \infty} \int_0^\infty \varphi(t) dF_n^D(t) = \int_0^{b_H} \varphi(t) dF(t) \quad \text{con probabilidad uno.}$$

Observación 1.2.3

Si la función $\varphi(x) = \mathbf{1}_{\{0 \leq x \leq t\}}$ verifica las condiciones de la propiedad anterior, y $t \leq b_H$, entonces se obtiene la convergencia puntual del estimador $F_n^D(\cdot)$:

$$F_n^D(t) \longrightarrow F(t) \quad \text{en probabilidad.}$$

Si además la función $\varphi(x) = \mathbf{1}_{\{0 \leq x \leq t\}}$ verifica las condiciones de la propiedad anterior para todo $t \leq b_H$, entonces se obtiene la consistencia fuerte del estimador $F_n^D(\cdot)$:

$$\sup_{0 \leq t < b_H} |F_n^D(t) - F(t)| \longrightarrow 0, \quad \text{c.s.}$$

En el resultado análogo de la ley fuerte para las integrales de Kaplan-Meier de Stute y Wang (1993), no se pide la continuidad de la función $H(\cdot)$ porque se supone que las funciones de distribución $F(\cdot)$ y $G(\cdot)$ no tienen saltos en común. Sin embargo, esta suposición no es apropiada en este modelo semiparamétrico. Basta tomar el caso particular del modelo de censura proporcional o de Koziol-Green, en el que se supone que $1 - G(\cdot) = (1 - F(\cdot))^\beta$, para algún $\beta > 0$. Obviamente, todos los átomos de probabilidad de las distribuciones dadas por $F(\cdot)$ y $G(\cdot)$ son los mismos.

Con respecto al estimador semiparamétrico de la razón de fallo acumulativa, Dikta (2001) obtuvo una representación débil de $\Lambda_n^D(\cdot)$, y, basándose en ella, demostró que $|\Lambda_n^D(\cdot) - \Lambda_F(\cdot)|$ está uniformemente acotado en probabilidad para todo $t \leq Z_{(n)}$. Estos resultados están resumidos en las siguientes propiedades:

Propiedad 1.2.14 (*Representación del estimador semiparamétrico de la razón de fallo acumulativa $\Lambda_n^D(\cdot)$, Dikta (2001)*)

Bajo ciertas condiciones sobre el estimador $\hat{\theta}_{0,MV}$, la función $p(\cdot, \theta)$, y suponiendo que $H(\cdot)$ es continua con $1 \leq r_n < n$, entonces

$$\Lambda_n^D(t) = \int_0^t \frac{dH_n^1(v)}{1 - H(v)} + \int_0^t \frac{H_n(v^-) - H(v)}{(1 - H(v))^2} dH^1(v) + R_n^0(t) + R_n^1(t),$$

donde

$$\sup_{0 \leq t \leq Z_{[r_n]}} |R_n^0(t)| = O_P\left(\frac{1}{n - r_n}\right) \quad y \quad \sup_{0 \leq t \leq Z_{[r_n]}} |R_n^1(t)| = O_P\left(n^{1/2} \sum_{i=1}^{r_n} \frac{1}{n - i + 1}\right).$$

Propiedad 1.2.15 (*Acotación uniforme del estimador semiparamétrico de la razón de fallo acumulativa $\Lambda_n^D(\cdot)$, Dikta (2001)*)

Bajo ciertas condiciones sobre el estimador $\hat{\theta}_{0,MV}$, la función $p(t, \theta)$, y suponiendo que $H(\cdot)$ es continua con $1 \leq r_n < n$, entonces

$$\sup_{0 \leq t \leq Z_{[r_n]}} |\Lambda_n^D(t) - \Lambda_F(t)| = O_P(1).$$

Recientemente, Zhu, Yuen y Tang (2002) han propuesto un test de bondad de ajuste para este modelo semiparamétrico propuesto por Dikta (1998). Debido a la complejidad de la distribución del estadístico de contraste bajo la hipótesis nula, los autores aproximaron los valores críticos mediante el método bootstrap y el método de simetrización aleatoria. Ambos métodos proporcionaron buenos resultados para tamaños muestrales moderados.

1.3. Estimación de la función de densidad.

1.3.1. El estimador basado en el estimador de Kaplan-Meier

El estimador tipo núcleo es el estimador no paramétrico de la función de densidad más estudiado y utilizado en los últimos años, y es el método en el que nos centraremos para estimar la densidad en presencia de censura aleatoria por la derecha. Desde los trabajos de Rosenblatt (1956) y Parzen (1962), los estimadores de tipo núcleo de la densidad han sido quizás los más populares, aunque los primeros trabajos sobre la versión para datos censurados por la derecha no aparecieron hasta 1980 (ver Blum y Susarla (1980)). En general, los estimadores tipo núcleo de la función de densidad son de la forma:

$$\hat{f}_h(t) = h^{-1} \int K\left(\frac{t-v}{h}\right) d\hat{F}(v) = \int K_h(t-v) d\hat{F}(v) = (K_h * \hat{F})(t) \quad (1.8)$$

donde $K_h(\cdot) = h^{-1}K(\cdot/h)$ es la función núcleo reescalada, $h > 0$ es el parámetro de suavización o ventana, el símbolo $*$ denota la convolución, y $\hat{F}(\cdot)$ es un estimador de la función de distribución, en nuestro caso, el estimador de Kaplan-Meier. Las condiciones que normalmente ha de verificar la función núcleo es que sea positiva, simétrica y que $\int K(u)du = 1$. Por tanto, en general, ha de ser una función de densidad, pues así $\hat{f}_h(\cdot)$ también lo será ya que hereda todas las propiedades analíticas del núcleo.

Observación 1.3.1

Si $F(\cdot)$ y $G(\cdot)$ son dos funciones de distribución, entonces se define la convolución $F * G(\cdot)$ como la función

$$H(t) = \int_{-\infty}^{\infty} G(t-v) dF(v) \quad \forall t \in \mathbb{R}.$$

La operación convolución verifica las propiedades conmutativa, asociativa y distributiva respecto a la suma. Si $G(\cdot)$ es absolutamente continua con densidad $g(\cdot)$, entonces $H(\cdot) = F * G(\cdot)$ es absolutamente continua con densidad $h(\cdot)$ dada por

$$h(t) = \int_{-\infty}^{\infty} g(t-v) dF(v) \quad \forall t \in \mathbb{R}.$$

Si además $F(\cdot)$ es absolutamente continua con densidad $f(\cdot)$, entonces

$$h(t) = \int_{-\infty}^{\infty} g(t-v)f(v)dv \quad \forall t \in \mathbb{R}.$$

La importancia de la convolución de funciones de distribución viene dada por el

hecho de que si X e Y son dos variables aleatorias independientes con funciones de distribución $F(\cdot)$ y $G(\cdot)$ respectivamente, entonces la función de distribución de la variable suma $S = X + Y$ es

$$H(t) = \int_{-\infty}^{\infty} G(t-v)dF(v) \quad \forall t \in \mathbb{R}.$$

Como ya se comentó antes, no fue hasta principios de los años 80 cuando se publicaron las primeras propiedades del estimador tipo núcleo de la función de densidad con datos censurados por la derecha. Földes, Rejtő y Winter (1981) obtuvieron, tomando como $\hat{F}(\cdot)$ en la expresión (1.8) el estimador de Kaplan-Meier, la convergencia fuerte para el estimador núcleo de la densidad $f_n^{KM}(\cdot)$, que se reduce al estimador usual tipo núcleo de Parzen (1962) en el caso de no censura (puesto que $F_n^{KM}(\cdot)$, en tal caso, se convierte en la función de distribución empírica usual, $F_n(\cdot)$).

Propiedad 1.3.1 (*Consistencia del estimador, Földes, Rejtő y Winter (1981)*)

Si $F(\cdot)$ es acotada, $G(b_F^-) < 1$, $K(\cdot)$ es continua por la derecha y de variación acotada en \mathbb{R} y la sucesión de ventanas h_n verifica que $h_n \rightarrow 0$ pero $h_n(n \log n)^{1/8} \rightarrow \infty$, entonces:

(a) Si $f(\cdot)$ es continua en t , entonces

$$f_n^{KM}(t) \rightarrow f(t) \quad \text{c.s. cuando } n \rightarrow \infty.$$

(b) Si $-\infty \leq a < b \leq \infty$ y $f(\cdot)$ es uniformemente continua en (a, b) , entonces $\forall \varepsilon > 0$:

$$\sup_{a+\varepsilon < t < b-\varepsilon} |f_n^{KM}(t) - f(t)| \rightarrow 0 \quad \text{c.s.}$$

(c) Si $-\infty \leq a < b \leq \infty$ y $f(\cdot)$ tiene una derivada acotada en (a, b) , entonces:

$$\sup_{a < t < b} |f_n^{KM}(t) - f(t)| \rightarrow 0 \quad \text{c.s.}$$

La consistencia uniforme de este estimador ha sido estudiada por diversos autores: Mielniczuk (1986), Stute y Wang (1993), Stute (1995), Chaubey y Sen (1996) y Gannoun y Saracco (2002) entre otros, abordando la demostración de muy diferentes formas. Como consecuencia, las condiciones sobre la sucesión de ventanas se han ido relajando, pidiendo finalmente tan sólo que $h_n \rightarrow 0$ y $nh_n^2(\log n)^{-1} \rightarrow \infty$.

Diehl y Stute (1988) representaron el estimador tipo núcleo de la función de densidad con pesos Kaplan-Meier $f_n^{KM}(\cdot)$ como una suma de variables aleatorias independientes más un término de error despreciable en probabilidad y de forma casi

segura. Este resultado es muy importante, ya que permite determinar la razón exacta de convergencia puntual (o ley del logaritmo iterado), la razón exacta de convergencia uniforme y obtener resultados puntuales sobre la distribución asintótica. Además, va a permitir relajar las condiciones sobre la sucesión de ventanas en la normalidad asintótica obtenida por Mielniczuk (1986), pidiendo tan sólo que $h_n \rightarrow 0$ y $nh_n \rightarrow \infty$. De los resultados de Diehl y Stute (1988) se puede obtener además una tasa para la consistencia casi segura del estimador (ver Xiang (1994)). Este orden es óptimo, en contraste con el obtenido por Karunamuni y Yang (1991).

Propiedad 1.3.2 (*Tasa de consistencia fuerte, Diehl y Stute (1988)*)

Sea $J = [a, b]$ un intervalo, y sea $K(\cdot)$ una función núcleo continuamente diferenciable con soporte compacto. Supongamos que $f(\cdot) = F'(\cdot)$ y $g(\cdot) = G'(\cdot)$ son acotadas en el intervalo $[0, T']$ para algún $T < T'$ donde $0 < T < \infty$ es un punto fijo tal que $H(T) < 1$, si $f^{(r)}(\cdot)$ es continua en $J_\varepsilon = [a - \varepsilon, b + \varepsilon]$, para algún $\varepsilon > 0$, entonces

$$\sup_{t \in J} |f_n^{KM}(t) - f(t)| = O\left(\left(\frac{\log n}{n}\right)^{r/(2r+1)}\right) \quad \text{c.s. con } h_n = O\left(\left(\frac{\log n}{n}\right)^{1/(2r+1)}\right).$$

Este resultado ha sido extendido a la derivada de orden m por Xiang (1994).

Propiedad 1.3.3 (*Tasa de consistencia fuerte del estimador de la derivada de orden m , Xiang (1994)*)

Sea $J = [a, b]$ un intervalo, y sea $K(\cdot)$ una función núcleo continuamente diferenciable hasta el orden $m+1$, para algún $m \geq 1$, con soporte compacto, verificando que $\int K(v) dv = 1$, $\int s^j K(v) dv = 0$ con $j = 1, \dots, r-1$ y $\int s^r K(v) dv \neq 0$. Supongamos que $f^{(m+r)}(\cdot)$ es continua en $J_\varepsilon = [a - \varepsilon, b + \varepsilon]$, con $0 < \delta \leq f(\cdot) \leq M < \infty$, para todo $t \in J_\varepsilon$, para algún $\varepsilon > 0$ y para algún $r \geq 2$.

Entonces, si $h_n = O\left(\left(\frac{\log n}{n}\right)^{1/[2(r+m)+1]}\right)$

$$\sup_{t \in J} |f_n^{KM(m)}(t) - f^{(m)}(t)| = O\left(\left(\frac{\log n}{n}\right)^{r/[2(r+m)+1]}\right) \quad \text{c.s.}$$

Por su parte, Ghorai y Pattanaik (1990) estudiaron la consistencia \mathcal{L}_1 del estimador. El principal resultado hace referencia a la distancia \mathcal{L}_1 , definida por

$$J_n(T) = \int_{-\infty}^T |f_n^{KM}(v) - f(v)| dv,$$

donde T es un número real convenientemente elegido. Suponiendo que la sucesión de

ventanas verifica que $h_n \rightarrow 0$ y $nh_n \rightarrow \infty$, los autores demostraron que $J_n(T) \rightarrow 0$ completamente cuando $n \rightarrow \infty$, así como la convergencia débil y fuerte de $J_n(T)$ a cero. En el caso de no censura, este resultado se reduce al obtenido por Devroye (1983).

La generalización de la distancia $J_n(T)$ a la distancia en \mathcal{L}_p dada por

$$J_n(T, p) = \int_{-\infty}^T |f_n^{KM}(v) - f(v)|^p dv$$

ha sido estudiada, entre otros, por Csörgo, Gombay y Horváth (1991), quienes demostraron que bajo ciertas condiciones impuestas sobre la función núcleo $K(\cdot)$, la sucesión de ventanas h_n y la función de densidad teórica $f(\cdot)$, entonces la distancia $J_n(T, p)$ es asintóticamente normal. El caso particular de $p = 2$ lo aborda también Zhang (1998), obteniendo un desarrollo de la función $J_n(t - \varepsilon)$ para todo $\varepsilon > 0$ y $t > 0$ tal que $H(t) < 1$.

La distribución asintótica del estimador $f_n^{KM}(\cdot)$ fue estudiada por, entre otros, Diehl y Stute (1988), Lo, Mack y Wang (1989) y más recientemente, por Zhang (1996) y Louani (1998), aunque los primeros resultados se deben a Mielniczuk (1986).

Propiedad 1.3.4 (Normalidad asintótica, Mielniczuk (1986))

Si $K(\cdot)$ es una función par, la función de densidad $f(\cdot)$ tiene segunda derivada acotada en un entorno de t , y la sucesión de ventanas h_n es de orden $O(n^{-1/3})$, entonces

$$\sqrt{nh_n} (f_n^{KM}(t) - f(t)) \xrightarrow{d} N(0, \sigma^2(t))$$

con varianza

$$\sigma^2(t) = \frac{f(t)}{1 - G(t)} \int K^2(v) dv.$$

Lo, Mack y Wang (1989) obtuvieron el sesgo y la varianza asintóticos del estimador y, basándose en la representación casi segura del estimador de Kaplan-Meier de la función de distribución obtenida por Lo y Singh (1986), establecieron la representación casi segura del estimador de la función de densidad con una cota para el término de error no sólo de forma casi segura, sino también en media. Finalmente establecieron, bajo ciertas hipótesis sobre la función núcleo y la sucesión de ventanas h_n , la ley del logaritmo iterado.

Propiedad 1.3.5 (*Sesgo y varianza, Lo, Mack y Wang (1989)*)

Sea $K(\cdot)$ una función de densidad continua, simétrica, de variación acotada y con soporte compacto. Sea h_n una sucesión de ventanas tales que $h_n \rightarrow 0$ y $(\log n)^2 (nh_n)^{-1} \rightarrow 0$. Si $f(\cdot)$ es una función de densidad dos veces continuamente diferenciable en t , entonces

$$\mathbb{E} [f_n^{KM}(t)] = f(t) + \frac{1}{2} f''(t) \left(\int v^2 K(v) dv \right) h_n^2 + o(h_n^2),$$

$$\text{Var} (f_n^{KM}(t)) = \frac{1}{nh_n} \frac{f(t)}{1-G(t)} \int K^2(v) dv + o((nh_n)^{-1}).$$

El término dominante del error cuadrático medio puede obtenerse directamente a partir de estas expresiones, de manera que el clásico balance entre el sesgo y la varianza proporcionará ventanas óptimas del orden $n^{-1/5}$, al igual que en el caso de no censura.

Propiedad 1.3.6 (*Representación casi segura, Lo, Mack y Wang (1989)*)

Sea $K(\cdot)$ una función núcleo de variación acotada y con soporte compacto. Entonces el estimador $f_n^{KM}(\cdot)$ admite la siguiente representación

$$f_n^{KM}(t) = f(t) + \beta_n(t) + \sigma_n(t) + e_n(t),$$

donde

$$\sup_{0 \leq t \leq T} |e_n(t)| = O\left((nh_n)^{-1} \log n\right) \quad \text{c.s.}$$

y

$$\sup_{0 \leq t \leq T} \mathbb{E} |e_n(t)|^\alpha = O\left(\left[(nh_n)^{-1} \log n\right]^\alpha\right) \quad \text{para todo } \alpha \geq 1$$

siendo

$$\beta_n(t) = \int f(t - h_n v) K(v) dv - f(t),$$

$$\sigma_n(t) = \frac{1}{nh_n} \sum_{i=1}^n \int \xi_i(t - h_n v) dK(v),$$

con

$$\xi_i(t) = (1 - F(t)) \left[- \int_0^{Z_i \wedge t} \frac{dH^1(v)}{(1 - H(v))^2} + \frac{1}{1 - H(Z_i)} \mathbf{1}_{\{Z_i \leq t, \delta_i = 1\}} \right].$$

El término $\beta_n(\cdot)$ es esencialmente el sesgo del estimador, mientras que $\sigma_n(\cdot)$ contendría la parte de la varianza. La expresión del error sería entonces $e_n(t) = h_n^{-1} \int r_n(t - h_n v) dK(v)$, donde $r_n(\cdot)$ es el término del error de la representación casi segura del estimador de Kaplan-Meier obtenida por Lo y Singh (1986).

Diehl y Stute (1988) habían obtenido previamente otra representación para el estimador de la función de densidad, en función de

$$\bar{f}_h(t) = \int K_h(t-v) dF(v).$$

En el caso de no censura, $\mathbb{E}(f_n^{KM}(t)) = \bar{f}_h(t)$; sin embargo, en presencia de censura, ambas funciones pueden diferir.

Propiedad 1.3.7 (Diehl y Stute (1988))

Sea $K(\cdot)$ una función núcleo continuamente diferenciable con soporte compacto, y supongamos que $f(\cdot) = F'(\cdot)$ y $g(\cdot) = G'(\cdot)$ son acotadas en el intervalo $[0, T']$ para algún $T < T'$ donde $0 < T < \infty$ es un punto fijo tal que $H(T) < 1$. Entonces

$$\sqrt{nh_n} \left| f_n^{KM}(t) - \bar{f}_h(t) - \frac{h_n^1(t) - \mathbb{E}[h_n^1(t)]}{1 - G(t)} \right| = R_n(t),$$

donde

$$\sup_{0 \leq t \leq T} |R_n(t)| = O\left((nh_n)^{-1/2}\right) + O\left(h_n^{1/2}\right) \text{ en probabilidad}$$

y

$$\sup_{0 \leq t \leq T} |R_n(t)| = O\left((\log \log n)(nh_n)^{-1/2}\right) + O\left(\log \log nh_n^{1/2}\right) \text{ de forma casi segura,}$$

siendo

$$\bar{f}_h(t) = \int K_h(t-v) dF(v)$$

y

$$h_n^1(t) = \int K_h(t-v) dH_n^1(v)$$

con

$$H_n^1(t) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{\{Z \leq t, \delta=1\}}.$$

Propiedad 1.3.8 (Ley del logaritmo iterado, Lo, Mack y Wang (1989))

Sea $K(\cdot)$ una función de densidad continua, simétrica, de variación acotada y con soporte compacto. Sea h_n una sucesión de ventanas tales que $h_n \rightarrow 0$, $h_n/h_m \rightarrow 1$ cuando $n, m \rightarrow \infty$ con $n/m \rightarrow 1$, $(\log n)^4 = o(nh_n(\log \log n))$ y $h_n = o\left((n^{-1} \log \log n)^{1/5}\right)$. Si $f(\cdot)$ es una función de densidad dos veces continuamente diferenciable en t , entonces

$$\limsup_{n \rightarrow \infty} \left(\frac{nh_n}{2 \log \log n} \right)^{1/2} |f_n^{KM}(t) - f(t)| = \left(\frac{f(t)}{1 - G(t)} \int K^2(v) dv \right)^{1/2} \text{ c.s.}$$

Diversos autores estudiaron esta propiedad para la diferencia en valor absoluto $|f_n^{KM}(\cdot) - \bar{f}_h(\cdot)|$. Bajo ciertas condiciones en la sucesión de ventanas h_n , Diehl y Stute (1988) demostraron que

$$\begin{aligned} \lim_{n \rightarrow \infty} \left(\frac{nh_n}{2 \log(1/h_n)} \right)^{1/2} \sup_{a \leq t \leq b} |f_n^{KM}(t) - \bar{f}_h(t)| \left(\frac{1 - G(t)}{f(t)} \right)^{1/2} \\ = \left(\int K^2(v) dv \right)^{1/2} \quad c.s. \end{aligned}$$

Este resultado ha sido posteriormente extendido por Xiang (1994), bajo la condición de que $h_n = O(n^{-\gamma})$ con $0 < \gamma < 1$, y Giné y Guillou (1999) demostraron, a su vez, este resultado para valores de t en el intervalo $(0, Z_{[n(1-\varepsilon_n)]}]$ con $\varepsilon_n \rightarrow 0$. Deheuvels y Einmahl (2000) generalizaron el resultado de Xiang (1994) bajo condiciones menos restrictivas e introduciendo una función concreta $\Psi(\cdot)$ continua y estrictamente positiva, siendo $\Psi_n(\cdot)$ un estimador de $\Psi(\cdot)$ tal que

$$\sup_{a \leq t \leq b} \left| \frac{\Psi_n(t)}{\Psi(t)} - 1 \right| \rightarrow 0 \text{ en probabilidad y de forma casi segura.}$$

En concreto, estos autores probaron que

$$\begin{aligned} \lim_{n \rightarrow \infty} \sup \left(\frac{nh_n}{2(\log(1/h_n) + \log \log n)} \right)^{1/2} \\ \times \sup_{a \leq t \leq b} |f_n^{KM}(t) - \bar{f}_h(t)| \left(\Psi_n(t) \frac{1 - G(t)}{f(t)} \right)^{1/2} = \left(\sup_{a \leq t \leq b} \Psi(t) \int K^2(v) dv \right)^{1/2} \quad c.s. \end{aligned}$$

Finalmente, la ley del logaritmo iterado para la diferencia $|f_n^{KM}(t) - \mathbb{E}[f_n^{KM}(t)]|$ ha sido obtenida por Zhang (1999).

En la estimación tipo núcleo de la función de densidad (con o sin censura), es bien conocido que la elección de la función núcleo tiene una importancia menor, y son comunes en la práctica los núcleos dados en la tabla 1.1.

Sin embargo, la bondad de cualquier procedimiento estadístico que implique suavización y, en concreto, la estimación tipo núcleo de la función de densidad, dependen de manera crucial de la elección del parámetro de suavización o ventana, ya que regula el grado de suavización del estimador. Ventanas demasiado pequeñas producen estimaciones con mucha variabilidad (infrasuavización), mientras que ventanas grandes dan lugar a estimaciones sesgadas (sobresuavización).

Tabla 1.1 Principales funciones núcleo usadas en la estimación no paramétrica tipo núcleo de la función de densidad.

$K(u) = \frac{3}{4} (1 - u^2) \mathbf{1}_{\{ u \leq 1\}}$	Núcleo de Epanechnikov
$K(u) = \frac{1}{\sqrt{2\pi}} e^{-u^2/2}$	Núcleo Gaussiano
$K(u) = (1 - u) \mathbf{1}_{\{ u \leq 1\}}$	Núcleo Triangular
$K(u) = \frac{15}{16} (1 - u^2)^2 \mathbf{1}_{\{ u \leq 1\}}$	Núcleo Cuártico
$K(u) = \frac{1}{2} \mathbf{1}_{\{ u \leq 1\}}$	Núcleo Uniforme

La Figura 1.1 ilustra, mediante un ejemplo, la sensibilidad del estimador de Parzen-Rosenblatt a la elección de la ventana de suavización. Se muestran distintas estimaciones de la función de densidad correspondiente a una mixtura de las distribuciones normales $N(2, 1)$ y $N(-2, 1)$ con probabilidades $p_1 = 0.6$ y $p_2 = 0.4$ respectivamente, que aparece dibujada en las cuatro gráficas con trazo continuo. Las líneas discontinuas representan estimaciones tipo núcleo de la función de densidad teórica obtenidas a partir de una muestra de 300 datos simulada de dicho modelo, utilizando la función núcleo de Epanechnikov, y considerando cuatro parámetros de suavización diferentes:

$$\begin{array}{ll}
 h = 0.2 \text{ y } h = 0.4 & \textit{infrasuavización} \\
 h = 0.925 & \textit{suavización adecuada} \\
 h = 2.5 & \textit{sobresuavización}
 \end{array}$$

Se intuye fácilmente, a partir de las gráficas, qué parámetro de suavización es el apropiado, debido a que la función subyacente a los datos es conocida; pero en el caso en el que la verdadera estructura bajo la que se obtuvieron los datos está oculta, la elección de la ventana puede ser bastante complicada. Por lo tanto, es imprescindible el desarrollo y uso de métodos automáticos de elección de la ventana h que permitan elegirla de acuerdo a algún criterio específico.

Podemos encontrar revisiones sobre los distintos métodos de estimación del parámetro de suavización en ausencia de censura en Cao, Cuevas y González Manteiga (1994), Jones, Marron y Sheather (1996) y Devroye (1997) entre otros. No obstante, cuando la muestra contiene datos censurados, la elección del parámetro ventana es un tema relativamente poco estudiado todavía.

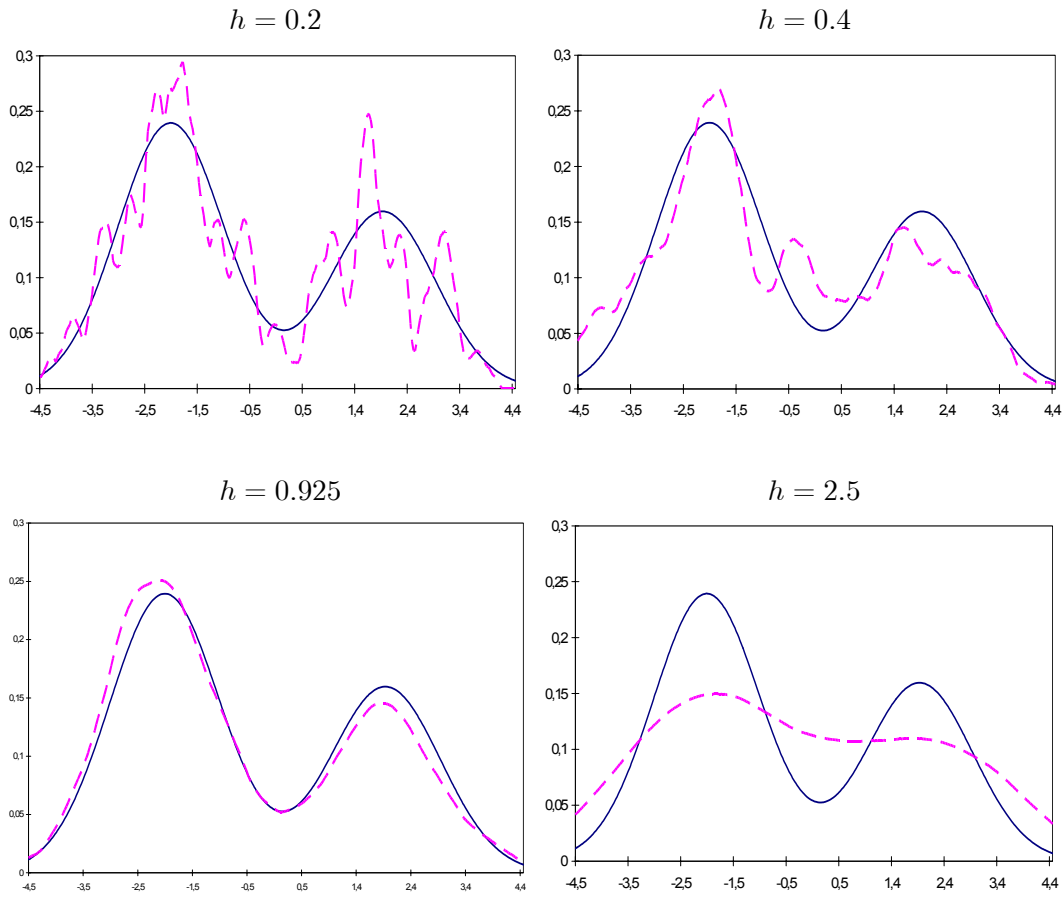


Figura 1.1. Función de densidad teórica de una mezcla de normales $N(2, 1)$ y $N(-2, 1)$ con probabilidades $p_1 = 0.6$ y $p_2 = 0.4$ respectivamente (línea continua) y su estimación mediante el método núcleo (línea discontinua).

La primera referencia en la literatura sobre el estudio de la selección de la ventana en presencia de censura aleatoria por la derecha es de Stute (1985), quien obtuvo la expresión de la ventana óptima local que minimiza la parte asintótica del error cuadrático medio ($AMSE$) de $f_n^{KM}(\cdot)$:

$$AMSE(f_n^{KM}(t)) = \left(\frac{1}{2} h^2 f''(t) \int v^2 K(v) dv \right)^2 + \frac{f(t)}{nh(1-H(t))} \int K^2(v) dv,$$

donde

$$MSE(f_n^{KM}(t)) = \mathbb{E} \left[(f_n^{KM}(t) - f(t))^2 \right].$$

La expresión de este parámetro ventana local asintóticamente óptimo es igual a

$$h_{AMSE}(t) = \left(\frac{f(t)}{4(1-H(t))} \int K^2(v) dv \right)^{1/5} \left(\frac{1}{2} f''(t) \int v^2 K(v) dv \right)^{-2/5} n^{-1/5}.$$

Paralelamente, Marron y Padgett (1987) proponen el estimador de la ventana de validación cruzada h_{CV} , introducido por primera vez para muestras completas por Rudemo (1982) y Bowman (1984), tomando como distancia entre $f_n^{KM}(\cdot)$ y $f(\cdot)$ el error cuadrático integrado:

$$ISE(f_n^{KM}, h_0) = \int (f_{n,h_0}^{KM}(v) - f(v))^2 dv$$

siendo $f_{n,h_0}^{KM}(\cdot)$ el estimador tipo núcleo con pesos Kaplan-Meier y ventana h_0 , y demuestran que el selector h_{CV} es asintóticamente óptimo en el siguiente sentido:

$$\frac{ISE(f_n^{KM}, h_{CV})}{\inf_h ISE(f_n^{KM}, h)} \rightarrow 1 \quad c.s.$$

Patil (1993) probó que este selector de la ventana converge al minimizador del error cuadrático integrado (ISE) con un orden relativo de $n^{-1/10}$.

Más recientemente, Kuhn y Padgett (1997) propusieron elegir la ventana localmente de manera que minimice la parte asintótica del error medio absoluto:

$$MAE(f_n^{KM}(t)) = \mathbb{E} |f_n^{KM}(t) - f(t)|,$$

cuya expresión final es

$$h_{AMAE}(t) = \left(\frac{4\alpha^2 f(t) \int K^2(v) dv}{\left(\int v^2 K(v) dv \right)^2 f''(t)^2 (1-H(t))} \right)^{1/5} n^{-1/5}$$

donde α es el valor para el cual $4\alpha(\Phi(\alpha) - 0.5) - \phi(\alpha) = 0$ siendo $\Phi(\cdot)$ y $\phi(\cdot)$ las funciones de distribución y de densidad respectivamente de la distribución normal estándar. Scott y Wand (1991) demostraron que $\alpha = 0.4809489$. La relación entre esta ventana y la obtenida por Stute (1985) es

$$\frac{h_{AMAE}}{h_{AMSE}} = (4\alpha^2)^{1/5} = 0.98458,$$

es decir, el valor es muy parecido.

Recientemente, Sánchez Sellero, González Manteiga y Cao (1999) han propuesto un selector tipo plug-in y otro bootstrap para la ventana en la estimación de la función de densidad en el caso más general de datos censurados y truncados.

Capítulo 2

Estimación presuavizada de la función de distribución

2.1. Introducción a la estimación presuavizada

La idea básica de los estimadores presuavizados consiste en estimar la función $p(t) = \mathbb{E}(\delta|Z = t)$ en las expresiones

$$\Lambda_F(t) = \int_0^t \frac{p(v)dH(v)}{1 - H(v-)} \quad \text{y} \quad 1 - F(t) = \exp[-\Lambda_F(t)] = \exp\left(-\int_0^t p(v)d\Lambda_H(v)\right)$$

de forma totalmente suave y no paramétrica. Dado que la función $p(t)$ es la función de regresión de δ condicionada a que $Z = t$, una posibilidad será estimar $p(Z_{(i)})$ por el estimador núcleo de Nadaraya-Watson $p_n(Z_{(i)})$ basado en las respuestas binarias δ_i con covariables Z_i , $i = 1, \dots, n$, es decir:

$$p_n(t) = \frac{(nb)^{-1} \sum_{i=1}^n K\left(\frac{t - Z_i}{b}\right) \delta_i}{(nb)^{-1} \sum_{i=1}^n K\left(\frac{x - Z_i}{b}\right)} = \frac{n^{-1} \sum_{i=1}^n K_b(t - Z_i) \delta_i}{n^{-1} \sum_{i=1}^n K_b(t - Z_i)}, \quad (2.1)$$

donde $K(\cdot)$ es una función núcleo, $K_b(\cdot) = b^{-1}K(\cdot/b)$ es la función núcleo reescalada, y $b \equiv b_n$, $n = 1, 2, \dots$ es la sucesión de ventanas.

Este estimador se puede interpretar como el de máxima verosimilitud local de $p(\cdot)$. Si la logverosimilitud local de los datos $(Z_i, \delta_i)_{i=1}^n$ es

$$\sum_{i=1}^n [\delta_i \log p(Z_i) + (1 - \delta_i) \log(1 - p(Z_i))] K_b(t - Z_i),$$

y para los Z_i de un entorno de t se usa la aproximación local constante

$$p(Z_i) \simeq a(t) \equiv a,$$

el estimador (2.1) es la solución al problema de maximizar

$$\sum_{i=1}^n [\delta_i \log(a) + (1 - \delta_i) \log(1 - a)] K_b(t - Z_i)$$

con respecto a a . Esta relación con los estimadores de máxima verosimilitud local da pie a usar otros estimadores no paramétricos de $p(\cdot)$, como por ejemplo los estimadores polinómicos locales, aunque en este capítulo nos restringiremos al estudio de los estimadores presuavizados con el estimador de Nadaraya-Watson.

El nombre de estimación presuavizada viene del hecho de que la suavización se usa únicamente para obtener una versión suavizada de los pesos de Kaplan-Meier, pero el estimador resultante de la función de distribución no es suave.

2.2. Definición del estimador

El estimador presuavizado de la función de distribución en presencia de censura aleatoria por la derecha se obtiene directamente del clásico estimador de Kaplan-Meier

$$1 - F_n^{KM}(t) = \prod_{Z_{(i)} \leq t} \left(1 - \frac{\delta_{[i]}}{n - i + 1} \right),$$

sin más que sustituir $\delta_{[i]}$ por un estimador suave de $p(Z_{(i)})$. Este estimador, propuesto por Cao, López de Ullibarri, Janssen y Veraverbeke (2005), es

$$1 - F_n^P(t) = \prod_{Z_{(i)} \leq t} \left(1 - \frac{p_n(Z_{(i)})}{n - i + 1} \right), \quad (2.2)$$

donde $p_n(\cdot)$ es el estimador tipo núcleo de Nadaraya-Watson basado en las respuestas binarias δ_i con covariables Z_i , $i = 1, \dots, n$ y dado por la expresión (2.1). Un estimador relacionado con éste es el propuesto por Ziegler (1995):

$$1 - F_n^Z(t) = \prod_{Z_{(i)} \leq t} \left(1 - \frac{\xi_{[i]}}{n - i + 1} \right),$$

donde $\xi_{[i]}$ está definido por

$$\xi_{[i]} := \begin{cases} p_n(Z_{(i)}) & \text{cuando } Z_{(i)} > b_n \\ \delta_{[i]} & \text{cuando } Z_{(i)} \leq b_n \end{cases}$$

es decir, este estimador coincide con el estimador de Kaplan-Meier en un entorno del cero, y con el estimador presuavizado en los demás puntos con el fin de evitar posibles problemas estocásticos en un entorno del cero. Esta autora establece la convergencia débil del proceso estocástico $\sqrt{n} (F_n^Z(\cdot) - F(\cdot))$, caracteriza su límite, y basándose en el trabajo de Hall y Wellner (1980), estudia y proporciona distintas bandas de confianza para la función de supervivencia.

2.3. Propiedades generales del estimador

A continuación exponemos algunas de las propiedades y observaciones sobre el estimador presuavizado de la función de distribución, que se derivan fácilmente de su propia definición.

Propiedad 2.3.1

La expresión del estimador presuavizado de la función de distribución $F_n^P(\cdot)$ se deriva de la del estimador presuavizado de la razón de fallo acumulativa $\Lambda_n^P(\cdot)$, tal como ocurre sin presuavizar con el estimador de Kaplan-Meier de la función de distribución y de Nelson-Aalen de la razón de fallo acumulativa.

La relación entre la razón de fallo acumulativa $\Lambda_F(\cdot)$ y la función de supervivencia $1 - F(\cdot)$ es, en general,

$$1 - F(t) = \exp[-\Lambda_F^c(t)] \prod_{\substack{a_i \in A \\ a_i \leq t}} (1 - \Lambda\{a_i\}), \quad t \in \mathbb{R}^+ \quad (2.3)$$

donde A es la masa de átomos de $F(\cdot)$,

$$\Lambda\{t\} := \Lambda(t) - \Lambda(t-)$$

y

$$\Lambda_F^c(t) = \Lambda(t) - \sum_{\substack{a_i \in A \\ a_i \leq t}} \Lambda\{a_i\}$$

es la “parte continua” de $\Lambda(\cdot)$ para todo $t \in \mathbb{R}^+$.

El estimador presuavizado de la razón de fallo acumulativa es una función constante a trozos de forma que

$$\Lambda_n^P\{Z_i\} = \frac{p_n(Z_i)}{n-i+1}$$

para cada $i = 1, \dots, n$. Esto motiva, por la relación (2.3), el siguiente estimador de la función de supervivencia $1 - F(\cdot)$:

$$1 - \widehat{F}(t) = \prod_{Z_{(i)} \leq t} \left(1 - \frac{p_n(Z_{(i)})}{n-i+1}\right) = 1 - F_n^P(t),$$

es decir, el estimador de Kaplan-Meier presuavizado.

Propiedad 2.3.2

El estimador presuavizado $1 - F_n^P(\cdot)$ puede tener saltos en todos los datos, sin limitarse únicamente a los datos no censurados, tal como ocurre con el estimador de Kaplan-Meier.

El estimador de Kaplan-Meier se puede expresar de la forma

$$F_n^{KM}(t) = \sum_{i=1}^n W_i^{KM} \mathbf{1}_{\{Z_{(i)} \leq t\}},$$

donde los pesos son

$$W_{(i)}^{KM} = \frac{\delta_{[i]}}{n-i+1} \prod_{j=1}^{i-1} \left(1 - \frac{\delta_{[j]}}{n-j+1}\right) \quad i = 1, \dots, n$$

es decir, no nulos únicamente en los datos con $\delta_{[i]} = 1$. Sin embargo, el estimador presuavizado es igual a

$$F_n^P(t) = \sum_{i=1}^n W_i^P \mathbf{1}_{\{Z_{(i)} \leq t\}},$$

con

$$W_{(i)}^P = \frac{p_n(Z_{(i)})}{n-i+1} \prod_{j=1}^{i-1} \left(1 - \frac{p_n(Z_{(j)})}{n-j+1}\right) \quad i = 1, \dots, n \quad (2.4)$$

en el que los productos vacíos valen 1.

Propiedad 2.3.3

Cuando no se presuaviza en absoluto, es decir, cuando se toma una ventana b muy pequeña ($b \rightarrow 0^+$), el estimador presuavizado $F_n^P(\cdot)$ de la función de distribución coincide con el clásico estimador de Kaplan-Meier $F_n^{KM}(\cdot)$.

Esto es debido a que, con una ventana de presuavizado b muy pequeña, el límite del estimador de la función $p(\cdot)$ en cada punto Z_i es

$$p_n(Z_{(i)}) = \frac{n^{-1} \sum_{j=1}^n K_b(Z_{(i)} - Z_j) \delta_j}{n^{-1} \sum_{j=1}^n K_b(Z_{(i)} - Z_j)} \rightarrow \frac{K(0)\delta_{[i]}}{K(0)} = \delta_{[i]},$$

Propiedad 2.3.4

En el caso de no censura, el estimador presuavizado $F_n^P(\cdot)$ de la función de distribución coincide con la función de distribución empírica.

En el caso de no censura, i.e., cuando $\delta_{[i]} = 1$ para $i = 1, \dots, n$, se verifica que

$$p_n(t) = \frac{(nb)^{-1} \sum_{j=1}^n K\left(\frac{t - Z_j}{b}\right) \delta_{[j]}}{(nb)^{-1} \sum_{j=1}^n K\left(\frac{t - Z_j}{b}\right)} = 1 = \delta_{[i]} \quad i = 1, \dots, n,$$

es decir, los estimadores de Kaplan-Meier y Kaplan-Meier presuavizado son iguales y coinciden con el estimador empírico de la función de supervivencia de la muestra:

$$\begin{aligned} 1 - F_n^P(t) &= \prod_{Z_{(i)} \leq t} \left(1 - \frac{p_n(Z_{(i)})}{n - i + 1}\right) = \prod_{Z_{(i)} \leq t} \left(1 - \frac{1}{n - i + 1}\right) = \prod_{Z_{(i)} \leq t} \left(\frac{n - i}{n - i + 1}\right) \\ &= \frac{n - 1}{n} \frac{n - 2}{n - 1} \dots \frac{n - k}{n - k + 1} = \frac{n - k}{n} = 1 - \sum_{i=1}^n \frac{\mathbf{1}_{\{Z_{(i)} \leq t\}}}{n}, \end{aligned}$$

donde $k = \max\{i : Z_{(i)} \leq t\}$ y por tanto $k = \sum_{i=1}^n \mathbf{1}_{\{Z_{(i)} \leq t\}}$.

Propiedad 2.3.5

Tal como ocurre en el caso del estimador clásico de Kaplan-Meier, otro estimador presuavizado de la función de supervivencia $1 - F_n^P(\cdot)$ también se podría expresar de la forma

$$1 - F_n^{P(2)}(t) = \prod_{Z_{(i)} \leq t} \left(1 - \frac{1}{n - i + 1}\right)^{p_n(Z_{(i)})},$$

si en lugar de usar la aproximación

$$\exp\left(-\frac{p_n(Z_{(i)})}{n - i + 1}\right) \simeq 1 - \frac{p_n(Z_{(i)})}{n - i + 1}$$

a la hora de construir el estimador, se usa la siguiente

$$\exp\left(-\frac{p_n(Z_{(i)})}{n-i+1}\right) = \exp\left(-\frac{1}{n-i+1}\right)^{p_n(Z_{(i)})} \simeq \left(1 - \frac{1}{n-i+1}\right)^{p_n(Z_{(i)})}.$$

Este estimador es asintóticamente equivalente a (2.2), puesto que

$$F_n^P(t) - F_n^{P(2)}(t) = O_P(n^{-1}).$$

Por simplicidad, sólo trabajaremos con el estimador $F_n^P(\cdot)$, ya que ambos proporcionan resultados casi idénticos en la mayoría de las situaciones prácticas.

Propiedad 2.3.6

El estimador presuavizado para la función de distribución $G(\cdot)$ de la variable de censura C viene dado por

$$1 - G_n^P(t) = \prod_{Z_{(i)} \leq t} \left(1 - \frac{1 - p_n(Z_{(i)})}{n - i + 1}\right),$$

o bien

$$1 - G_n^{P(2)}(t) = \prod_{Z_{(i)} \leq t} \left(1 - \frac{1}{n - i + 1}\right)^{1 - p_n(Z_{(i)})}.$$

Es inmediato, dado que una observación censurada de la variable de interés Y corresponde precisamente con una no censurada de la variable C .

Propiedad 2.3.7

El producto de los estimadores presuavizados de las funciones de supervivencia $1 - F(\cdot)$ y $1 - G(\cdot)$, en su segunda versión, coincide con la supervivencia empírica de la variable observada $1 - H(\cdot)$, es decir,

$$\left(1 - F_n^{P(2)}(t)\right) \left(1 - G_n^{P(2)}(t)\right) = 1 - H_n(t).$$

El producto de ambas funciones es

$$\begin{aligned} & \left(1 - F_n^{P(2)}(t)\right) \left(1 - G_n^{P(2)}(t)\right) \\ &= \prod_{Z_{(i)} \leq t} \left(1 - \frac{1}{n - i + 1}\right)^{p_n(Z_{(i)})} \prod_{Z_{(i)} \leq t} \left(1 - \frac{1}{n - i + 1}\right)^{1 - p_n(Z_{(i)})} \\ &= \prod_{Z_{(i)} \leq t} \left(1 - \frac{1}{n - i + 1}\right) = 1 - \sum_{i=1}^n \frac{\mathbf{1}_{\{Z_{(i)} \leq t\}}}{n} = 1 - H_n(t). \end{aligned}$$

Esta igualdad se convierte en sólo una aproximación cuando se usa la versión original del estimador presuavizado.

2.4. Propiedades asintóticas del estimador

En este apartado obtenemos una representación casi segura del estimador pre-suavizado $F_n^P(\cdot)$ como suma de variables independientes e idénticamente distribuidas, más un término de orden despreciable. Ésta es una de las propiedades más importantes de este estimador, pues a partir de ella se pueden demostrar otras como la consistencia fuerte uniforme o la normalidad asintótica del estimador.

Esta representación ya fue obtenida por Cao, López de Ullibarri, Janssen y Veraverbeke (2005), aunque únicamente en probabilidad. Nosotros recorreremos su demostración para obtener, al igual que en la representación obtenida en Lo, Mack y Wang (1989) para el estimador de Kaplan-Meier, el comportamiento asintótico del resto de forma casi segura y en media, ya que se necesitará para obtener resultados análogos en el caso de la función de densidad y en los procedimientos para la selección de las ventanas. Para ello establecemos ciertos lemas previos.

2.4.1. Lemas previos

Lema 2.4.1

$$\max_{i \in \{1, \dots, n-1\}} \frac{1 - H(Z_{(i)})}{1 - H_n(Z_{(i)})} = O(\log n) \quad \text{casi seguro.} \quad (2.5)$$

Demostración.

En el lema 1 de Sánchez Sellero *et al.* (2005). ■

Lema 2.4.2

$$\mathbb{E} \left[\left(\max_{i \in \{1, \dots, n-1\}} \frac{1 - H(Z_{(i)})}{1 - H_n(Z_{(i)})} \right)^k \right] = O(1).$$

Demostración.

En la página 61 de Petrov (1995) se puede encontrar la siguiente propiedad:

$$\mathbb{E}(|X|^k) = k \int_0^\infty P(|X| \geq x) x^{k-1} dx \quad (2.6)$$

para cualquier $k > 0$ y para una variable aleatoria arbitraria X . Este resultado, junto con la desigualdad de Shorack y Wellner (1986), página 415,

$$P \left(\max_{i \in \{1, \dots, n-1\}} \frac{1 - H(Z_{(i)})}{1 - H_n(Z_{(i)})} \geq \lambda \right) \leq e\lambda e^{-\lambda} \quad \forall \lambda \geq 1 \quad \forall n \in \mathbb{N}$$

conduce a

$$\begin{aligned} \mathbb{E} \left[\left(\max_{i \in \{1, \dots, n-1\}} \frac{1 - H(Z_{(i)})}{1 - H_n(Z_{(i)})} \right)^k \right] \\ = k \int_0^\infty P \left(\max_{i \in \{1, \dots, n-1\}} \frac{1 - H(Z_{(i)})}{1 - H_n(Z_{(i)})} \geq x \right) x^{k-1} dx \\ \leq k \int_0^\infty e x e^{-x} x^{k-1} dx = ek \int_0^\infty e^{-x} x^k dx = ek k! \end{aligned}$$

llegando así a la cota uniforme con el orden que se buscaba. ■

Lema 2.4.3 Sea $H(\cdot)$ una función de distribución arbitraria y sea $H_n(\cdot)$ su correspondiente distribución empírica. Entonces, para todo $k > 0$, se verifica que

$$\mathbb{E} \left[\left(\sup_t |H_n(t) - H(t)| \right)^k \right] = O(n^{-k/2}).$$

Demostración.

Por la desigualdad de Dvoretzky, Kiefer y Wolfowitz (1956) se tiene que

$$P \left(\sqrt{n} \sup_t |H_n(t) - H(t)| > \lambda \right) \leq 2ce^{-2\lambda^2} \quad \forall \lambda \geq 0 \quad \forall n \in \mathbb{N}$$

siendo c una constante universal. Combinando este resultado con la igualdad (2.6) podemos concluir que

$$\begin{aligned} \mathbb{E} \left[\left(\sup_t |H_n(t) - H(t)| \right)^k \right] &= k \int_0^\infty P \left(\sup_t |H_n(t) - H(t)| \geq x \right) x^{k-1} dx \\ &= k \int_0^\infty P \left(\sqrt{n} \sup_t |H_n(t) - H(t)| \geq x\sqrt{n} \right) x^{k-1} dx \\ &\leq k \int_0^\infty 2ce^{-2nx^2} x^{k-1} dx = ck \int_0^\infty e^{-2nz} z^{k/2-1} dz = ck \frac{\Gamma(\frac{k}{2})}{(2n)^{k/2}}, \end{aligned}$$

quedando de este modo demostrado el enunciado del lema. ■

Lema 2.4.4 Sea el núcleo $K(\cdot)$ una función de densidad con $\int |v| K(v) dv < \infty$, con derivada $K^{(j)}(\cdot)$ continua y de variación acotada para $j = 0, 1, \dots, N$. Sea $h_n(\cdot)$ el estimador de Parzen-Rosenblatt de la función de densidad $h(\cdot)$, con sucesión de ventanas b_n . Si $h(\cdot)$ y sus $(N+1)$ primeras derivadas son acotadas, y ε_n es una sucesión de números positivos tales que $\lim_{n \rightarrow \infty} b_n/\varepsilon_n = 0$, entonces existen dos constantes C_1 y C_2 tales que

$$P\left(\sup_t \left| h_n^{(j)}(t) - h^{(j)}(t) \right| > \varepsilon_n\right) \leq C_1 \exp\left(-C_2 n \varepsilon_n^2 b_n^{2j+2}\right) \text{ para } j = 0, 1, \dots, N.$$

Demostración.

Es el lema 2.4 de Schuster (1969). ■

Lema 2.4.5 Sea $h(\cdot)$ una función de densidad arbitraria y sea $h_n(\cdot)$ su estimador de Parzen-Rosenblatt con sucesión de ventanas b_n . Entonces, bajo las condiciones del lema anterior, se verifica para todo $k > 0$

$$\mathbb{E} \left[\left(\sup_t |h_n(t) - h(t)| \right)^k \right] = O\left((nb^2)^{-k/2}\right).$$

Demostración.

Usando de nuevo (2.6) y el lema anterior obtenemos:

$$\begin{aligned} \mathbb{E} \left[\left(\sup_t |h_n(t) - h(t)| \right)^k \right] &= k \int_0^\infty P\left(\sup_t |h_n(t) - h(t)| \geq x\right) x^{k-1} dx \\ &\leq k C_1 \int_0^\infty \exp(-C_2 n b^2 x^2) x^{k-1} dx \\ &= \frac{1}{2} k C_1 \int_0^\infty \exp(-C_2 n b^2 v) v^{k/2-1} dv \\ &= \frac{1}{2} k C_1 \frac{\Gamma\left(\frac{k}{2}\right)}{(C_2 n b^2)^{k/2}} = \frac{1}{2} k \frac{C_1}{C_2^{k/2}} \Gamma\left(\frac{k}{2}\right) (nb^2)^{-k/2}, \end{aligned}$$

como queríamos demostrar. ■

Lema 2.4.6 Sea $m(t) = \mathbb{E}(Y|X=t)$ la función de regresión de Y sobre X , y sea $m_n(\cdot)$ el estimador de Nadaraya-Watson con sucesión de ventanas b_n . Supongamos que el núcleo $K(\cdot)$ es una función de densidad acotada, simétrica, de variación acotada con $\int_0^t |v| K(v) dv < \infty$, y $\int_0^t v^2 K(v) dv < \infty$. Si las funciones $h(\cdot)$ y $m(\cdot)$ son continuas en toda la recta real, y $\min_{-\infty < a \leq t \leq b < \infty} h(t) = \mu > 0$, entonces

$$\begin{aligned} P\left(\sup_{a \leq t \leq b} |m_n(t) - m(t)| > \varepsilon\right) &\leq C_0 \left[\exp\left(-\gamma_1 [\varepsilon(\mu - \varepsilon)]^2 n b_n^2\right) + \exp\left(-\gamma_2 [\varepsilon(\mu - \varepsilon)]^2 n b_n^2\right) \right. \\ &\quad \left. + \exp\left(-\gamma_3 [\varepsilon(\mu - \varepsilon)]^2 n b_n^2\right) \right] + C_1 \exp\left(-\gamma_4 [\varepsilon(\mu - \varepsilon)]^2 n b_n^2\right). \end{aligned}$$

donde las constantes $C_0, C_1, \gamma_1, \gamma_2, \gamma_3$ y γ_4 son positivas y dependen de $l_1 = \max_{a \leq t \leq b} m(t) < \infty$ y de la función núcleo $K(\cdot)$.

Demostración.

Es el teorema 2.1 de Nadaraya (1989). ■

Observación 2.4.1

Dado μ , para un n suficientemente grande de modo que $\varepsilon < \mu/2$, de la desigualdad anterior se puede deducir la siguiente

$$P \left(\sup_{a \leq t \leq b} |m_n(t) - m(t)| > \varepsilon \right) \leq C_2 \exp(-\gamma_5 \varepsilon^2 n b_n^2)$$

siendo C_2 y γ_5 dos constantes dependientes de $C_0, C_1, \gamma_1, \gamma_2, \gamma_3$ y γ_4 , puesto que si $\varepsilon < \mu/2$, entonces $[\varepsilon(\mu - \varepsilon)]^2 > \varepsilon^2 \mu^2 / 4$, y por tanto

$$\exp\left(-\gamma[\varepsilon(\mu - \varepsilon)]^2 n b_n^2\right) \leq \exp\left(-\gamma \frac{\mu^2}{4} \varepsilon^2 n b_n^2\right).$$

Lema 2.4.7 Sea $p(t) = \mathbb{E}(\delta|X=t)$ la función de regresión de δ sobre X , y sea $p_n(\cdot)$ su estimador de Nadaraya-Watson con ventana b_n . Entonces, bajo las condiciones del lema anterior, para todo $k > 0$ se verifica que

$$\mathbb{E} \left[\left(\sup_t |p_n(t) - p(t)| \right)^k \right] = O \left((n b_n^2)^{-k/2} \right).$$

Demostración.

Idéntica a la del lema 2.4.5. ■

Lema 2.4.8 Sea $h(\cdot)$ una función de densidad, y sea $h_n(\cdot)$ el estimador de Parzen-Rosenblatt con ventana b y función núcleo $K(\cdot)$. Si $K(\cdot)$ es una función de densidad simétrica y $h(\cdot)$ es cuatro veces diferenciable con $h^{(4)}(\cdot)$ acotada, entonces

$$\begin{aligned} \mathbb{E} \left[(h_n(t) - h(t))^4 \right] &= \frac{3}{(nb)^2} h^2(t) c_K^2 + \frac{3}{2} \frac{b^3}{n} h''(t)^2 h(t) c_K d_K^2 \\ &\quad + \frac{1}{16} b^8 h''(t)^4 d_K^4 + \frac{1}{n^2 b} Q_n(t), \end{aligned}$$

donde

$$c_K = \int K^2(v) dv, \quad d_K = \int v^2 K(v) dv \quad (2.7)$$

y $Q_n(\cdot)$ depende del núcleo $K(\cdot)$, de $h(\cdot)$, de $h''(\cdot)$, de $h^{(4)}(\cdot)$ y del tamaño muestral n a través de la ventana b .

Demostración.

Denotamos por $\mu(\cdot)$ la siguiente esperanza

$$\mu(t) = \mathbb{E}[h_n(t)] = \mathbb{E}[K_b(t - Z)] = h(t) + b^2 Q_{n1}(t),$$

donde

$$Q_{ni}(t) = \int v^2 K^i(v) \int_0^1 (1-x) h''(t - vbx) dx dv.$$

El primer paso es sumar y restar la esperanza del estimador, $\mu(\cdot)$, para así descomponer la esperanza de la potencia cuarta de $h_n(\cdot) - h(\cdot)$ en los siguientes términos:

$$\begin{aligned} \mathbb{E}[(h_n(t) - h(t))^4] &= \mathbb{E}[(h_n(t) - \mu(t)) + (\mu(t) - h(t))]^4 \\ &= \mathbb{E}[(h_n(t) - \mu(t))^4] + 4(\mu(t) - h(t)) \mathbb{E}[(h_n(t) - \mu(t))^3] \\ &\quad + 6(\mu(t) - h(t))^2 \mathbb{E}[(h_n(t) - \mu(t))^2] \\ &\quad + 4(\mu(t) - h(t))^3 \mathbb{E}[h_n(t) - \mu(t)] + (\mu(t) - h(t))^4. \end{aligned}$$

La diferencia $\mu(t) - h(t)$ es igual a

$$\mu(t) - h(t) = b^2 Q_{n1}(t) = \frac{1}{2} b^2 d_K h''(t) + b^4 R_n(t),$$

con

$$R_n(t) = \frac{1}{6} \int v^4 K(v) \int_0^1 (1-x)^3 h^{(4)}(t - vbx) dx dv,$$

de modo que el último de los sumandos es igual a

$$(\mathbb{E}[h_n(t)] - h(t))^4 = \frac{1}{16} b^8 d_K^4 h''(t)^4 + b^{10} \tilde{Q}_{n,5}(t),$$

donde $\tilde{Q}_{n,5}(\cdot)$ depende de $R_n(\cdot)$, de $h''(\cdot)$ y del tamaño muestral n a través de b .

Así mismo, por definición de la función $\mu(\cdot)$, el penúltimo sumando es cero. A continuación, demostraremos que el primero de ellos es igual a

$$\mathbb{E}[(h_n(t) - \mathbb{E}[h_n(t)])^4] = 6 \frac{1}{(nb)^2} h^2(t) c_K^2 + \frac{1}{n^2 b} \tilde{Q}_{n,1}(t),$$

donde $\tilde{Q}_{n,1}(\cdot)$ es una función que depende de $h(\cdot)$, $Q_{n1}(\cdot)$, $Q_{n2}(\cdot)$, $Q_{n3}(\cdot)$ y $Q_{n4}(\cdot)$.

En efecto,

$$\begin{aligned} \mathbb{E} \left[(h_n(t) - \mathbb{E}[h_n(t)])^4 \right] &= \mathbb{E} \left[\left(\frac{1}{n} \sum_{i=1}^n (K_b(t - Z_i) - \mu(t)) \right)^4 \right] \\ &= \frac{1}{n^4} \sum_{i,j,k,l=1}^n \mathbb{E} [(K_b(t - Z_i) - \mu(t)) (K_b(t - Z_j) - \mu(t)) \\ &\quad \times (K_b(t - Z_k) - \mu(t)) (K_b(t - Z_l) - \mu(t))] \\ &= \frac{1}{n^4} \sum_{i,j,k,l=1}^n E_{ijkl}. \end{aligned}$$

La idea consiste en separar la suma en varios términos, clasificando los sumandos en función de las coincidencias entre los cuatro índices. Denotaremos por $N = \{1, \dots, n\}$ al conjunto de valores que puede tomar cada índice y formalizamos la clasificación así

$$N^4 = \{(i, j, k, l) / i, j, k, l \in N\} = I_1 \cup I_2 \cup I_3 \cup I_4 \cup I_5$$

siendo

$$\begin{aligned} I_1 &= \{(i, j, k, l) \in N^4 / \# \{i, j, k, l\} = 4\} \\ I_2 &= \{(i, j, k, l) \in N^4 / \# \{i, j, k, l\} = 3\} \\ I_3 &= \{(i, j, k, l) \in N^4 / \# \{i, j, k, l\} = 2 \text{ empatados dos a dos}\} \\ I_4 &= \{(i, j, k, l) \in N^4 / \# \{i, j, k, l\} = 2 \text{ y tres de ellos empatados}\} \\ I_5 &= \{(i, j, k, l) \in N^4 / \# \{i, j, k, l\} = 1\} \end{aligned}$$

Podemos separar la suma anterior en los siguientes términos:

$$\begin{aligned} \mathbb{E} \left[(h_n(t) - \mathbb{E}[h_n(t)])^4 \right] &= \frac{1}{n^4} \sum_{(i,j,k,l) \in I_1} E_{ijkl} + \frac{1}{n^4} \sum_{(i,j,k,l) \in I_2} E_{ijkl} \\ &\quad + \frac{1}{n^4} \sum_{(i,j,k,l) \in I_3} E_{ijkl} + \frac{1}{n^4} \sum_{(i,j,k,l) \in I_4} E_{ijkl} + \frac{1}{n^4} \sum_{(i,j,k,l) \in I_5} E_{ijkl} \end{aligned}$$

Serán cero las esperanzas de aquellos términos que tengan al menos un índice de los cuatro no empatado con los restantes, de modo que, únicamente, consideraremos los casos en los que o bien los cuatro índices coinciden, o bien haya exactamente dos pares de coincidencias, es decir,

$$\mathbb{E} \left[(h_n(t) - \mathbb{E}[h_n(t)])^4 \right] = \frac{1}{n^4} \sum_{i,j,k,l \in I_3} E_{ijkl} + \frac{1}{n^4} \sum_{i,j,k,l \in I_5} E_{ijkl}.$$

Empezando por I_3 , podemos ver que $\text{cardinal}(I_3) = 3n(n-1)$, de modo que

$$\begin{aligned}
\frac{1}{n^4} \sum_{i,j,k,l \in I_3} E_{ijkl} &= \frac{3n(n-1)}{n^4} E_{1122} = 3 \frac{n-1}{n^3} \left(\mathbb{E} [K_b(t - Z_1) - \mu(t)]^2 \right)^2 \\
&= 3 \frac{n-1}{n^3} \left(\mathbb{E} [K_b^2(t - Z_1) - 2\mu(t) K_b(t - Z_1) + \mu^2(t)] \right)^2 \\
&= 3 \frac{n-1}{n^3} \left(\frac{1}{b} h(t) c_K + b Q_{n2}(t) - (h(t) + b^2 Q_{n1}(t))^2 \right)^2 \\
&= \frac{3}{n^2 b^2} h^2(t) c_K^2 + \frac{3}{n^2 b} Q_{n,I3}(t),
\end{aligned}$$

y

$$Q_{n,I3}(t) = 2h(t) c_K \tilde{Q}_{n,I3}(t) + b \tilde{Q}_{n,I3}^2(t) - \frac{1}{nb} \left(h(t) c_K + b \tilde{Q}_{n,I3}(t) \right)^2,$$

siendo

$$\tilde{Q}_{n,I3}(t) = -h^2(t) + b^2 (Q_{n2}(t) - 2h(t) Q_{n1}(t)) - b^4 Q_{n1}^4(t).$$

Por otro lado, $\text{cardinal}(I_5) = n$, de modo que

$$\begin{aligned}
\frac{1}{n^4} \sum_{i,j,k,l \in I_5} E_{ijkl} &= \frac{1}{n^3} E_{1111} = \frac{1}{n^3} \mathbb{E} \left[(K_b(t - Z_1) - \mu(t))^4 \right] \\
&= \frac{1}{n^3} \left(\mathbb{E} [K_b^4(t - Z_1)] - 4\mu(t) \mathbb{E} [K_b^3(t - Z_1)] + 6\mu^2(t) \mathbb{E} [K_b^2(t - Z_1)] - 3\mu^4(t) \right) \\
&= \frac{1}{n^3} \left[\frac{1}{b^3} h(t) \int_0^t K^4(v) dv + \frac{1}{b} Q_{n4}(t) - 4\mu(t) \left(\frac{1}{b^2} h(t) \int K^3(v) dv + Q_{n3}(t) \right) \right. \\
&\quad \left. + 6\mu^2(t) \left(\frac{1}{b} h(t) c_K + b Q_{n2}(t) \right) - 3\mu^4(t) \right] \\
&= \frac{1}{(nb)^3} h(t) \int K^4(v) dv + \frac{1}{n^3 b^2} Q_{n,I5}(t),
\end{aligned}$$

donde $Q_{n,I5}(\cdot)$ depende de $Q_{n1}(\cdot)$, $Q_{n2}(\cdot)$, $Q_{n3}(\cdot)$, $Q_{n4}(\cdot)$ y $h(\cdot)$.

En resumen,

$$\mathbb{E} \left[(h_n(t) - \mathbb{E} [h_n(t)])^4 \right] = \frac{3}{(nb)^2} h^2(t) c_K^2 + \frac{1}{n^2 b} \tilde{Q}_{n,1}(t),$$

donde

$$\tilde{Q}_{n,1}(t) = 3Q_{n,I3}(t) + \frac{1}{nb^2} h(t) \int K^4(v) dv + \frac{1}{nb} Q_{n,I5}(t).$$

Mediante cálculos análogos, podemos obtener la esperanza de la potencia tercera de $h_n(t) - h(t)$:

$$\begin{aligned}
\mathbb{E} \left[(h_n(t) - \mathbb{E}[h_n(t)])^3 \right] &= \mathbb{E} \left[\left(\frac{1}{n} \sum_{i=1}^n (K_b(t - Z_i) - \mu(t)) \right)^3 \right] \\
&= \frac{1}{n^3} \sum_{i,j,k=1}^n \mathbb{E} [(K_b(t - Z_i) - \mu(t)) (K_b(t - Z_j) - \mu(t)) \\
&\quad \times (K_b(t - Z_k) - \mu(t))] = \frac{1}{n^3} \sum_{i,j,k=1}^n E_{ijk} = \frac{1}{n^2} E_{111}
\end{aligned}$$

El término E_{111} es igual a

$$\begin{aligned}
E_{111} &= \mathbb{E} \left[(K_b(t - Z_1) - \mu(t))^3 \right] \\
&= \mathbb{E} \left[K_b^3(t - Z_1) \right] - 3\mu(t) \mathbb{E} \left[K_b^2(t - Z_1) \right] + 2\mu^3(t) \\
&= \left(\frac{1}{b^2} h(t) \int K^3(v) dv + Q_{n3}(t) \right) - 3\mu(t) \left(\frac{1}{b} h(t) c_K + bQ_{n2}(t) \right) + 2\mu^3(t) \\
&= \frac{1}{b^2} h(t) \int K^3(v) dv + \tilde{Q}_{n,111}(t),
\end{aligned}$$

donde $\tilde{Q}_{n,111}(\cdot)$ es una función que depende de $Q_{n2}(\cdot)$, $Q_{n3}(\cdot)$ y de $h(\cdot)$.

Por tanto, el segundo término de la descomposición de $\mathbb{E} \left[(h_n(t) - h(t))^4 \right]$ es despreciable frente al primero, puesto que es igual a

$$\begin{aligned}
&4(\mu(t) - h(t)) \mathbb{E} \left[(h_n(t) - \mu(t))^3 \right] \\
&= 4 \left(\frac{1}{2} b^2 d_K h''(t) + b^4 R_n(t) \right) \left(\frac{1}{(nb)^2} h(t) \int K^3(v) dv + \frac{1}{n^2} \tilde{Q}_{n,111}(t) \right) \\
&= \frac{2}{n^2} d_K h''(t) h(t) \int K^3(v) dv + \left(\frac{b}{n} \right)^2 \tilde{Q}_{n,2}(t),
\end{aligned}$$

con $\tilde{Q}_{n,2}(\cdot)$ dependiendo de $h(\cdot)$, $h''(\cdot)$, $R_n(\cdot)$ y $\tilde{Q}_{n,111}(\cdot)$.

La esperanza de $(h_n(t) - \mu(t))^2$ es igual a

$$\begin{aligned}
\mathbb{E} \left[(h_n(t) - \mu(t))^2 \right] &= \mathbb{E} \left[\left(\frac{1}{n} \sum_{i=1}^n (K_b(t - Z_i) - \mu(t)) \right)^2 \right] \\
&= \frac{1}{n} \mathbb{E} \left[(K_b(t - Z_1) - \mu(t))^2 \right] = \frac{1}{n} (\mathbb{E} [K_b^2(t - Z_1)] - \mu^2(t)) \\
&= \frac{1}{n} \left[\left(\frac{1}{b} h(t) c_K + bQ_{n2}(t) \right) - (h(t) + b^2 Q_{n1}(t))^2 \right] \\
&= \frac{1}{nb} h(t) c_K + \frac{1}{n} \tilde{Q}_{n,11}(t),
\end{aligned}$$

siendo $\tilde{Q}_{n,11}(\cdot)$ una función que depende de $h(\cdot)$, $Q_{n1}(\cdot)$ y $Q_{n2}(\cdot)$. Entonces, el tercer término de la descomposición de $\mathbb{E} \left[(h_n(t) - h(t))^4 \right]$ es igual a

$$\begin{aligned}
& 6(\mu(t) - h(t))^2 \mathbb{E} \left[(h_n(t) - \mu(t))^2 \right] \\
&= 6 \left(\frac{1}{2} b^2 d_K h''(t) + b^4 R_n(t) \right)^2 \left(\frac{1}{nb} h(t) c_K + \frac{1}{n} \tilde{Q}_{n,11}(t) \right) \\
&= \frac{6 b^3}{4 n} d_K^2 h''(t)^2 h(t) c_K + \frac{b^4}{n} \tilde{Q}_{n,3}(t),
\end{aligned}$$

con $\tilde{Q}_{n,3}(\cdot)$ dependiendo de $h(\cdot)$, $h''(\cdot)$, $R_n(\cdot)$ y $\tilde{Q}_{n,11}(\cdot)$.

Recopilamos todas estas cotas, y obtenemos finalmente que

$$\begin{aligned}
& \mathbb{E} \left[(h_n(t) - h(t))^4 \right] \\
&= \frac{3}{(nb)^2} h^2(t) c_K^2 + \frac{3 b^3}{2 n} d_K^2 h''(t)^2 h(t) c_K + \frac{1}{16} b^8 d_K^4 h''(t)^4 + \frac{1}{n^2 b} Q_n(t),
\end{aligned}$$

donde $Q_n(\cdot)$ depende de $h(\cdot)$, $h''(\cdot)$, de $Q_{n1}(\cdot)$, $Q_{n2}(\cdot)$, $Q_{n3}(\cdot)$, $Q_{n4}(\cdot)$ y de $h^{(4)}(\cdot)$ a través de $R_n(\cdot)$. ■

Lema 2.4.9 *Sea $p(t) = \mathbb{E}(\delta|Z=t)$ la función de regresión de δ sobre Z , es decir, la probabilidad condicional de que una observación es no censurada, y sea $p_n(\cdot)$ el estimador de Nadaraya-Watson con sucesión de ventanas b_n y función núcleo $K(\cdot)$. Si $K(\cdot)$ es una función de densidad simétrica, y las funciones $h(\cdot)$ y $p(\cdot)$ son cuatro veces diferenciables con $h^{(4)}(\cdot)$ y $p^{(4)}(\cdot)$ acotadas, entonces*

$$\begin{aligned}
& \mathbb{E} \left[(p_n(t) - p(t))^4 \right] \\
&\leq \frac{8}{h^4(t)} \left(\frac{3}{(nb)^2} (h^2(t) + \psi^2(t)) c_K^2 + \frac{3 b^3}{2 n} (h''(t)^2 h(t) + \psi''(t)^2 \psi(t)) c_K d_K^2 \right. \\
&\quad \left. + \frac{1}{16} b^8 (h''(t)^4 + \psi''(t)^4) d_K^4 + \frac{1}{n^2 b} Q_n(t) \right),
\end{aligned}$$

donde $\psi(t) = p(t)h(t)$ y $Q_n(\cdot)$ depende del núcleo $K(\cdot)$, de $h(\cdot)$, $p(\cdot)$, de las segundas derivadas de $h(\cdot)$ y $p(\cdot)$ y del tamaño muestral, n , a través de b .

Demostración.

Sea $h_n(t) = n^{-1} \sum_{i=1}^n K_b(t - Z_i)$ el estimador de Parzen-Rosenblatt, con ventana b , de la función de densidad $h(\cdot)$, y sea $\psi_n(t) = n^{-1} \sum_{i=1}^n K_b(t - Z_i) \delta_i$. Debido que el estimador de Nadaraya-Watson es igual a $p_n(t) = \psi_n(t)/h_n(t)$, y por lo tanto tiene denominador aleatorio, hemos factorizado la diferencia $p_n(t) - p(t)$ de la siguiente manera:

$$\begin{aligned}
|p_n(t) - p(t)| &= h^{-1}(t) |(\psi_n(t) - \psi(t)) - p_n(t)(h_n(t) - h(t))| \\
&\leq h^{-1}(t) (|\psi_n(t) - \psi(t)| + |h_n(t) - h(t)|),
\end{aligned}$$

puesto que $p_n(\cdot) \leq 1$. Entonces, su momento de orden 4 se puede acotar por

$$\begin{aligned} \mathbb{E} \left[(p_n(t) - p(t))^4 \right] &\leq h^{-4}(t) \mathbb{E} \left[(|\psi_n(t) - \psi(t)| + |h_n(t) - h(t)|)^4 \right] \\ &\leq 8h^{-4}(t) \left(\mathbb{E} \left[(\psi_n(t) - \psi(t))^4 \right] + \mathbb{E} \left[(h_n(t) - h(t))^4 \right] \right). \end{aligned}$$

En el lema 2.4.8 se demostró que

$$\begin{aligned} &\mathbb{E} \left[(h_n(t) - h(t))^4 \right] \\ &= \frac{3}{(nb)^2} h^2(t) c_K^2 + \frac{3b^3}{2n} h''(t)^2 h(t) c_K d_K^2 + \frac{1}{16} b^8 h''(t)^4 d_K^4 + \frac{1}{n^2 b} Q_{n,h}(t). \end{aligned}$$

Mediante cálculos análogos obtenemos también que

$$\begin{aligned} &\mathbb{E} \left[(\psi_n(t) - \psi(t))^4 \right] \\ &= \frac{3}{(nb)^2} \psi^2(t) c_K^2 + \frac{3b^3}{2n} \psi''(t)^2 \psi(t) c_K d_K^2 + \frac{1}{16} b^8 \psi''(t)^4 d_K^4 + \frac{1}{n^2 b} Q_{n,\psi}(t), \end{aligned}$$

donde $Q_{nh}(\cdot)$ y $Q_{n,\psi}(\cdot)$ dependen de la función núcleo $K(\cdot)$, de $h(\cdot)$, $\psi(\cdot)$ y de sus derivadas segundas y cuartas respectivamente. Entonces,

$$\begin{aligned} &\mathbb{E} \left[(p_n(t) - p(t))^4 \right] \\ &\leq \frac{8}{h^4(t)} \left[\frac{3}{(nb)^2} (h^2(t) + \psi^2(t)) c_K^2 + \frac{3b^3}{2n} (h''(t)^2 h(t) + \psi''(t)^2 \psi(t)) c_K d_K^2 \right. \\ &\quad \left. + \frac{1}{16} b^8 (h''(t)^4 + \psi''(t)^4) d_K^4 + \frac{1}{n^2 b} Q_n(t) \right], \end{aligned}$$

como se quería demostrar. ■

2.4.2. Representación asintótica

La representación del estimador de Kaplan-Meier se basa en la representación del estimador de la razón de fallo acumulativa. Esto ocurre también con el estimador presuavizado de la función de distribución.

Los pasos que seguiremos son, en primer lugar, demostrar que el estimador presuavizado $1 - F_n^P(\cdot)$ se puede aproximar por el estimador presuavizado de la razón de fallo acumulativa $\Lambda_n^P(\cdot)$ a través de una transformación exponencial más un término de error despreciable:

$$1 - F_n^P(t) = \exp \left[-\Lambda_n^P(t) \right] + R_{n1}(t). \quad (2.8)$$

A continuación, a partir de la expresión (2.8) y mediante un desarrollo de Taylor de orden dos se deriva la siguiente igualdad:

$$\begin{aligned}
F_n^P(t) - F(t) &= \exp[-\Lambda_F(t)] - \exp[-\Lambda_n^P(t)] - R_{n1}(t) \\
&= (1 - F(t)) (\Lambda_n^P(t) - \Lambda_F(t)) \\
&\quad - \frac{1}{2} (\Lambda_n^P(t) - \Lambda_F(t))^2 \exp[-\eta_n(t)] - R_{n1}(t) \\
&= (1 - F(t)) (\Lambda_n^P(t) - \Lambda_F(t)) + R_{n2}(t) - R_{n1}(t),
\end{aligned}$$

con $R_{n2}(t) = -\frac{1}{2} (\Lambda_n^P(t) - \Lambda_F(t))^2 \exp[-\eta_n(t)]$ y $\eta_n(t)$ un valor intermedio estocástico entre $\Lambda_n^P(t)$ y $\Lambda_F(t)$. Se probará entonces la despreciabilidad casi segura y en media de este nuevo término de error $R_{n2}(t)$, teniendo en cuenta que

$$\exp[-\eta_n(t)] \leq e^0 = 1 \quad \forall t < b_H.$$

Por último, obtendremos una representación casi segura y en media del estimador presuavizado de la razón de fallo acumulativa

$$\Lambda_n^P(t) - \Lambda_F(t) = \frac{1}{n} \sum_{i=1}^n \zeta(t, Z_i, \delta_i) + S_n(t).$$

Estos resultados se demuestran a continuación como paso previo a la obtención de la representación casi segura y en media del estimador $F_n^P(\cdot)$. Para ello, será necesario suponer las siguientes condiciones sobre la función núcleo:

(K1) La función $K(\cdot)$, con soporte en un intervalo compacto el cual, sin pérdida de generalidad, supondremos $[-1, 1]$, es una función de densidad simétrica, continua, de variación acotada y dos veces continuamente diferenciable en $[-1, 1]$. Además, también satisface que $K(1) = K'(1) = K''(1) = 0$.

Fijemos un punto $t_H > 0$ tal que $t_H < b_H = \sup\{t > 0 : H(t) < 1\}$. Las condiciones sobre la función de distribución $H(\cdot)$ son:

(H1) La función $H(\cdot)$ es continuamente diferenciable en $[0, t_H]$.

(H2) La función $H(\cdot)$ es tres veces continuamente diferenciable en $[0, t_H]$.

(H3) La función $H(\cdot)$ es cinco veces continuamente diferenciable en $[0, t_H]$.

(H4) Existe $\mu > 0$ tal que $H'(t) = h(t) > \mu, \forall t \in [0, t_H]$.

La hipótesis sobre la probabilidad condicional de no censura $p(\cdot)$ es

(p1) $p(\cdot)$ es dos veces continuamente diferenciable en $[0, t_H]$.

(p2) $p(\cdot)$ es cuatro veces continuamente diferenciable en $[0, t_H]$.

Finalmente, las condiciones sobre la ventana b son:

(v1) $n^{1-\varepsilon}b \rightarrow \infty$ para algún $\varepsilon > 0$ y $\sum b^\lambda < \infty$ para algún $\lambda > 0$ cuando $n \rightarrow \infty$.

(v2) $nb^2 (\log n)^{-6} \rightarrow \infty, nb^8 (\log n)^4 \rightarrow 0$ y $b^3 (\log n)^5 \rightarrow 0$ cuando $n \rightarrow \infty$.

Las condiciones (K1), (H3) y (p2) son condiciones de regularidad estándar. El grado de diferenciable exigido en ellas se puede relajar para algunos resultados asintóticos preliminares a la representación casi segura de $F_n^P(\cdot)$ dada al final de esta sección. A su vez, la condición (v1) es necesaria para la obtención de la representación casi segura y en media de $\Lambda_n^P(\cdot) - \Lambda_F(\cdot)$ y de $F_n^P(\cdot) - F(\cdot)$, mientras que se necesita la condición (v2) para la demostración de la normalidad asintótica del estimador presuavizado $F_n^P(\cdot)$. Las hipótesis en la condición (v1) son las necesarias para la aplicación del lema 1 y teorema B de Mack y Silverman (1982).

Proposición 2.4.1 *Bajo la condición (H1) se tiene para todo $t \in [0, t_H]$*

$$1 - F_n^P(t) = \exp[-\Lambda_n^P(t)] + R_{n1}(t),$$

donde

$$\sup_{0 \leq t \leq t_H} |R_{n1}(t)| = O\left(n^{-1} (\log n)^2\right) \quad c.s. \quad \text{y} \quad \sup_{0 \leq t \leq t_H} \mathbb{E}[R_{n1}^2(t)] = O(n^{-2}).$$

Demostración.

Las funciones $1 - F_n^P(\cdot)$ y $\exp[-\Lambda_n^P(\cdot)]$ vienen dadas por las siguientes expresiones

$$1 - F_n^P(t) = \prod_{Z_{(i)} \leq t} \left(1 - \frac{p_n(Z_{(i)})}{n - i + 1}\right) \quad \text{y} \quad \exp[-\Lambda_n^P(t)] = \prod_{Z_{(i)} \leq t} \exp\left(-\frac{p_n(Z_{(i)})}{n - i + 1}\right).$$

Mediante un desarrollo de Taylor de la función $\exp(-x)$ en torno al cero, podemos escribir $\exp[-\Lambda_n^P(t)]$ de la forma

$$\exp[-\Lambda_n^P(t)] = \prod_{Z_{(i)} \leq t} \left(1 - \frac{p_n(Z_{(i)})}{n - i + 1} + \frac{1}{2} \left(\frac{p_n(Z_{(i)})}{n - i + 1}\right)^2 e^{-\xi_i}\right),$$

donde ξ_i es un valor estocástico intermedio entre 0 y $p_n(Z_{(i)}) / (n - i + 1)$.

Basándonos en la siguiente propiedad general

$$\left| \prod_{j=1}^n a_j - \prod_{j=1}^n b_j \right| \leq \sum_{j=1}^n |a_j - b_j|,$$

que se verifica siempre que $|a_j| \leq 1$ y $|b_j| \leq 1 \quad \forall j \in \{1, \dots, n\}$, escribimos

$$\begin{aligned} & |1 - F_n^P(t) - \exp[-\Lambda_n^P(t)]| \\ &= \left| \prod_{Z(i) \leq t} \left(1 - \frac{p_n(Z(i))}{n-i+1}\right) - \prod_{Z(i) \leq t} \left(1 - \frac{p_n(Z(i))}{n-i+1} + \frac{1}{2} \left(\frac{p_n(Z(i))}{n-i+1}\right)^2 e^{-\xi_i}\right) \right| \\ &\leq \sum_{Z(i) \leq t} \left| -\frac{1}{2} \left(\frac{p_n(Z(i))}{n-i+1}\right)^2 e^{-\xi_i} \right| = \sum_{Z(i) \leq t} \frac{1}{2} \left(\frac{p_n(Z(i))}{n-i+1}\right)^2 e^{-\xi_i}. \end{aligned}$$

Teniendo en cuenta además que $e^{-\xi_i} \leq e^0 = 1$ y que $p_n(Z(i)) \leq 1$, podemos entonces concluir que

$$\begin{aligned} |R_{n1}(t)| &= |1 - F_n^P(t) - \exp[-\Lambda_n^P(t)]| \\ &\leq \sum_{Z(i) \leq t} \frac{1}{2} \frac{1}{(n-i+1)^2} \leq \sum_{Z(i) \leq t} \frac{1}{2n^2 (1 - H_n(Z(i)))^2} \\ &= \sum_{Z(i) \leq t} \frac{1}{2} \frac{1}{n^2} \left(\frac{1 - H(Z(i))}{1 - H_n(Z(i))}\right)^2 \frac{1}{(1 - H(Z(i)))^2} \\ &\leq \frac{1}{2n} \left(\max_{Z(i) \leq t} \frac{1 - H(Z(i))}{1 - H_n(Z(i))}\right)^2 \left(\frac{1}{n} \sum_{i=1}^n \frac{1}{(1 - H(Z_i))^2}\right). \end{aligned}$$

Aplicando el lema 2.4.1 y la ley fuerte de los grandes números a la media muestral que aparece como último factor, por la condición (H1) se concluye que

$$\sup_{0 \leq t \leq t_H} |R_{n1}(t)| = O\left(n^{-1} (\log n)^2\right) \quad c.s.$$

Acotamos ahora el cuadrado del valor absoluto de $R_{n1}(t)$ de la siguiente manera:

$$R_{n1}^2(t) \leq \frac{1}{4n^2} \left(\max_{Z(i) \leq t} \frac{1 - H(Z(i))}{1 - H_n(Z(i))}\right)^4 \left(\frac{1}{n} \sum_{i=1}^n \frac{1}{(1 - H(Z_i))^2}\right)^2$$

de modo que, por la desigualdad de Cauchy-Schwarz, la esperanza de $R_{n1}^2(t)$ se puede acotar por

$$\mathbb{E}[R_{n1}^2(t)] \leq \frac{1}{4n^2} \left[\mathbb{E} \left(\max_{Z(i) \leq t} \frac{1 - H(Z(i))}{1 - H_n(Z(i))} \right)^8 \right]^{1/2} \left[\mathbb{E} \left(\frac{1}{n} \sum_{i=1}^n \frac{1}{(1 - H(Z_i))^2} \right)^4 \right]^{1/2}.$$

El último término está acotado, puesto que es el momento de orden 4 de la media muestral de variables aleatorias independientes idénticamente distribuidas. Basta aplicar el resultado siguiente

$$\mathbb{E} \left[\left(\frac{1}{n} \sum_{i=1}^n X_i \right)^p \right] < \infty \quad \text{si} \quad \mathbb{E}(X_1^p) < \infty \quad \text{para todo} \quad 0 < p < \infty.$$

Apelando a este resultado, al lema 2.4.2 y teniendo en cuenta (H1) resulta entonces que

$$\sup_{0 \leq t \leq t_H} \mathbb{E} [R_{n1}^2(t)] = O(n^{-2}),$$

dando por concluida la demostración. ■

A partir de la expresión (2.8) y mediante un desarrollo de Taylor de orden dos se deriva la siguiente igualdad:

$$F_n^P(t) - F(t) = (1 - F(t)) (\Lambda_n^P(t) - \Lambda_F(t)) + R_{n2}(t) - R_{n1}(t),$$

con

$$R_{n2}(t) = -\frac{1}{2} (\Lambda_n^P(t) - \Lambda_F(t))^2 \exp[-\eta_n(t)]$$

y $\eta_n(t)$ un valor intermedio estocástico entre $\Lambda_n^P(t)$ y $\Lambda_F(t)$.

El siguiente lema proporciona la despreciabilidad casi segura y en media de este nuevo término de error $R_{n2}(t)$, teniendo en cuenta que

$$\exp[-\eta_n(t)] \leq e^0 = 1 \quad \forall t < b_H.$$

Lema 2.4.10 *Bajo las condiciones (K1), (H2), (H4), (p1) y (v1), entonces*

$$\sup_{0 \leq t \leq t_H} |\Lambda_n^P(t) - \Lambda_F(t)|^2 = O \left(\left(b^2 + (nb)^{-1/2} (\log n)^{1/2} \right)^2 (\log n)^2 \right) \quad c.s.$$

y

$$\sup_{0 \leq t \leq t_H} \mathbb{E} \left[|\Lambda_n^P(t) - \Lambda_F(t)|^4 \right] = O \left(\left(b^4 + \frac{1}{nb} \right)^2 \right).$$

Demostración.

Llevamos a cabo la siguiente descomposición:

$$\begin{aligned}
& \Lambda_n^P(t) - \Lambda_F(t) \\
&= \int_0^t p_n(v) d\Lambda_n^H(v) - \int_0^t p(v) d\Lambda_H(v) = \int_0^t \frac{p_n(v) dH_n(v)}{1 - H_n(v)} - \int_0^t \frac{p(v) dH(v)}{1 - H(v)} \\
&= \int_0^t \left(\frac{p_n(v)}{1 - H_n(v)} - \frac{p(v)}{1 - H(v)} \right) dH_n(v) + \int_0^t \frac{p(v)}{1 - H(v)} d(H_n(v) - H(v)) \\
&= (I) + (II).
\end{aligned}$$

Empezaremos acotando el supremo en t del cuadrado de ambos sumandos, (I) e (II). Respecto al segundo término, (II), que es despreciable frente al primero, acotamos el valor absoluto (a) y realizamos una integración por partes (b):

$$\begin{aligned}
|(II)| &\stackrel{(a)}{\leq} \int_0^t \frac{p(v)}{1 - H(v)} d|H_n(v) - H(v)| \\
&\stackrel{(b)}{\leq} \frac{p(t)}{1 - H(t)} |H_n(t) - H(t)| + \left| \int_0^t |H_n(v) - H(v)| \left(\frac{p(v)}{1 - H(v)} \right)' dv \right| \\
&\leq \frac{p(t)}{1 - H(t)} |H_n(t) - H(t)| + \sup_{0 \leq t \leq t_H} |H_n(t) - H(t)| \int_0^t \left| \left(\frac{p(v)}{1 - H(v)} \right)' \right| dv.
\end{aligned}$$

Utilizando el siguiente resultado para la función de distribución empírica (ver, por ejemplo, Dvoretzky, Kiefer y Wolfowitz (1956)):

$$\sup_t |H_n(t) - H(t)| = O\left(n^{-1/2} (\log n)^{1/2}\right) \quad c.s. \quad (2.9)$$

y teniendo en cuenta que $p(\cdot) \leq 1$, se tiene entonces, bajo (H2) y (p1),

$$\sup_{0 \leq t \leq t_H} |(II)|^2 \leq \sup_{0 \leq t \leq t_H} |H_n(t) - H(t)|^2 \left[\frac{1}{1 - H(t_H)} + \int_0^{t_H} \left| \left(\frac{p(v)}{1 - H(v)} \right)' \right| dv \right]^2,$$

de modo que

$$\sup_{0 \leq t \leq t_H} |(II)|^2 = O(n^{-1} \log n) \quad c.s.$$

Por su parte, el valor absoluto del primer término (I) se puede acotar por

$$\begin{aligned}
|(I)| &= \left| \int_0^t \left(\frac{p_n}{1 - H_n} - \frac{p}{1 - H} \right) dH_n \right| = \left| \int_0^t \frac{p_n(1 - H) - p(1 - H_n)}{(1 - H_n)(1 - H)} dH_n \right| \\
&= \left| \int_0^t [(p_n - p)(1 - H_n) + p_n(H_n - H)] \frac{1 - H}{1 - H_n(1 - H)^2} dH_n \right| \\
&\leq \left(\sup_t |p_n - p| + \sup_t |H_n - H| \right) \left(\max_{Z_{(i)} \leq t} \frac{1 - H(Z_{(i)})}{1 - H_n(Z_{(i)})} \right) \int_0^t \frac{dH_n}{(1 - H)^2}.
\end{aligned}$$

Así, utilizando (2.5), (2.9), junto con la aplicación, bajo la hipótesis (v1), del si-

guiente resultado para la función de regresión (teorema B de Mack y Silverman (1982))

$$\sup_t |p_n(t) - p(t)| = O\left(b^2 + (nb)^{-1/2} (\log n)^{1/2}\right) \quad c.s. \quad (2.10)$$

y aplicando la ley fuerte de los grandes números a la media muestral que aparece como último factor, podemos concluir que

$$\sup_{0 \leq t \leq t_H} |(I)|^2 = O\left(\left(b^2 + (nb)^{-1/2} (\log n)^{1/2}\right)^2 (\log n)^2\right) \quad c.s.$$

En resumen,

$$\sup_{0 \leq t \leq t_H} |\Lambda_n^P(t) - \Lambda_F(t)|^2 = O\left(\left(b^2 + (nb)^{-1/2} (\log n)^{1/2}\right)^2 (\log n)^2\right) \quad c.s.$$

A continuación, acotaremos el supremo del momento de orden 4 de $|\Lambda_n^P(t) - \Lambda_F(t)|$ mediante la acotación de $\mathbb{E}|(I)^4|$ y $\mathbb{E}|(II)^4|$. El segundo término no supone ningún problema, puesto que podemos acotarlo directamente por

$$\begin{aligned} & \sup_{0 \leq t \leq t_H} \mathbb{E}|(II)^4| \\ & \leq \mathbb{E}\left(\sup_{0 \leq t \leq t_H} |H_n(t) - H(t)|^4\right) \left[\frac{1}{1 - H(t_H)} + \int_0^{t_H} \left|\left(\frac{p(v)}{1 - H(v)}\right)'\right| dv\right]^4, \end{aligned}$$

por lo que, aplicando el lema 2.4.3 y teniendo en cuenta (H2) y (p1),

$$\sup_{0 \leq t \leq t_H} \mathbb{E}|(II)^4| = O(n^{-2}).$$

Para acotar el momento de orden 4 de $|(I)|$, descomponemos previamente la integral en los siguientes sumandos:

$$\begin{aligned} (I) &= \int_0^t \frac{p_n(v)(1 - H(v)) - p(v)(1 - H_n(v))}{(1 - H_n(v))(1 - H(v))} dH_n(v) \\ &= \int_0^t (p_n(v) - p(v)) \frac{dH_n(v)}{1 - H(v)} + \int_0^t p_n(v)(H_n(v) - H(v)) \frac{1 - H(v)}{1 - H_n(v)} \frac{dH_n(v)}{(1 - H(v))^2} \end{aligned}$$

de modo que el valor absoluto de (I) se puede acotar por

$$\begin{aligned} |(I)| &\leq \int_0^t |p_n(v) - p(v)| \frac{dH_n(v)}{1 - H(v)} + \int_0^t |H_n(v) - H(v)| \frac{1 - H(v)}{1 - H_n(v)} \frac{dH_n(v)}{(1 - H(v))^2} \\ &= (I_1) + (I_2). \end{aligned}$$

Así, el momento de orden 4 de (I) es menor o igual que

$$\mathbb{E}|(I)|^4 \leq 8 \left[\mathbb{E} \left((I_1)^4 \right) + \mathbb{E} \left((I_2)^4 \right) \right].$$

La esperanza de la potencia cuarta de (I_2) se acota de la siguiente manera:

$$\begin{aligned} \mathbb{E} \left[(I_2)^4 \right] &= \mathbb{E} \left[\left(\int_0^t |H_n(v) - H(v)| \frac{1 - H(v)}{1 - H_n(v)} \frac{dH_n(v)}{(1 - H(v))^2} \right)^4 \right] \\ &\leq \mathbb{E} \left[\int_0^t |H_n(v) - H(v)|^4 \left(\frac{1 - H(v)}{1 - H_n(v)} \right)^4 \frac{dH_n(v)}{(1 - H(v))^8} \right] \\ &\leq \mathbb{E} \left[\sup_{0 \leq t \leq t_H} |H_n(v) - H(v)|^{12} \right]^{1/3} \mathbb{E} \left[\left(\max_{Z(i) \leq t} \frac{1 - H(Z(i))}{1 - H_n(Z(i))} \right)^{12} \right]^{1/3} \\ &\quad \times \mathbb{E} \left[\left(\int_0^t \frac{dH_n(v)}{(1 - H(v))^8} \right)^3 \right]^{1/3}. \end{aligned}$$

La última esperanza está acotada por

$$\mathbb{E} \left[\left(\int_0^t \frac{dH_n}{(1 - H(v))^8} \right)^3 \right]^{1/3} \leq \int_0^{t_H} \frac{dH(v)}{(1 - H(v))^8} < \frac{1}{(1 - H(t_H))^8} < \infty.$$

En virtud de esto, y junto con los resultados de los lemas 2.4.2 y 2.4.3, concluimos que

$$\sup_{0 \leq t \leq t_H} \mathbb{E} \left[(I_2)^4 \right] = O(n^{-2}).$$

Ahora descomponemos el sumando (I_1) en los siguientes términos

$$\begin{aligned} (I_1) &\leq \int_0^t \frac{|p_n(v) - p(v)|}{1 - H(v)} d(H_n(v) - H(v)) + \int_0^t |p_n(v) - p(v)| \frac{dH(v)}{1 - H(v)} \\ &= (I_{11}) + (I_{12}), \end{aligned}$$

de manera que, finalmente,

$$\mathbb{E} \left[(I_1)^4 \right] \leq 8 \left(\mathbb{E} \left[(I_{11})^4 \right] + \mathbb{E} \left[(I_{12})^4 \right] \right).$$

El primero de los sumandos, (I_{11}) , se puede acotar por

$$\begin{aligned} |(I_{11})| &\leq \sup_{0 \leq t \leq t_H} |p_n(t) - p(t)| \int_0^t \frac{d|H_n(v) - H(v)|}{1 - H(v)} \\ &= \sup_{0 \leq t \leq t_H} |p_n(t) - p(t)| \left(\frac{|H_n(t) - H(t)|}{1 - H(t)} - \int_0^t \frac{h(v) |H_n(v) - H(v)|}{(1 - H(v))^2} dv \right) \end{aligned}$$

$$\leq \frac{2}{1 - H(t_H)} \sup_{0 \leq t \leq t_H} |p_n(t) - p(t)| \sup_{0 \leq t \leq t_H} |H_n(t) - H(t)|.$$

Su momento de orden 4 es entonces menor o igual que

$$\mathbb{E} \left[(I_{11})^4 \right] \leq \frac{16}{(1 - H(t_H))^4} \mathbb{E} \left(\sup_{0 \leq t \leq t_H} |p_n(t) - p(t)|^8 \right)^{1/2} \mathbb{E} \left(\sup_{0 \leq t \leq t_H} |H_n(t) - H(t)|^8 \right)^{1/2}.$$

Los lemas 2.4.3 y 2.4.7 nos permiten concluir por tanto que $\mathbb{E} \left[(I_{11})^4 \right] = O \left((nb)^{-4} \right)$.

El momento de orden 4 de la integral (I_{12}) se puede acotar, en $[0, t_H]$, por:

$$\begin{aligned} \mathbb{E} \left[(I_{12})^4 \right] &= \mathbb{E} \left[\left(\int_0^t |p_n(v) - p(v)| \frac{h(v) dv}{1 - H(v)} \right)^4 \right] \leq \int_0^t \frac{\mathbb{E} \left[(p_n(v) - p(v))^4 \right]}{(1 - H(v))^4} h^4(v) dv \\ &\leq \sup_{0 \leq t \leq t_H} \mathbb{E} \left[(p_n(v) - p(v))^4 \right] \int_0^t \frac{h^4(v)}{(1 - H(v))^4} dv \end{aligned}$$

donde, por la aplicación del lema 2.4.9, se tiene

$$\mathbb{E} \left[(p_n(t) - p(t))^4 \right] = O \left(\left(b^4 + (nb)^{-1} \right)^2 \right).$$

Este orden deriva de la aplicación de desarrollos de Taylor de grado dos en las funciones $h(\cdot)$ y $p(\cdot)$, de modo que bajo las condiciones de regularidad impuestas por las hipótesis $(H2)$ y $(p2)$, por $(H4)$ y dado que $[0, t_H]$ es un intervalo compacto, concluimos que

$$\sup_{0 \leq t \leq t_H} \mathbb{E} \left[(I_{12})^4 \right] = O \left(\left(b^4 + (nb)^{-1} \right)^2 \right).$$

Recopilando los resultados para (I) y (II) obtenemos las cotas buscadas. ■

En resumen, el estimador presuavizado de la función de distribución se puede expresar como función de la razón de fallo acumulativa presuavizada de la siguiente manera:

$$F_n^P(t) - F(t) = (1 - F(t)) (\Lambda_n^P(t) - \Lambda_F(t)) + R_{n12}(t),$$

donde $R_{n12}(t) = R_{n1}(t) + R_{n2}(t)$ verifica que

$$\sup_{0 \leq t \leq t_H} |R_{n12}(t)| = O \left(\left(b^2 + (nb)^{-1/2} (\log n)^{1/2} \right)^2 (\log n)^2 \right) \quad c.s$$

y

$$\sup_{0 \leq t \leq t_H} \mathbb{E} \left[R_{n12}^2(t) \right] = O \left(\left(b^4 + (nb)^{-1} \right)^2 \right).$$

La representación del estimador de la función de distribución pasa ahora por obtener la representación del estimador presuavizado de la función de razón de fallo acumulativa como una suma de variables i.i.d. más un resto $S_n(\cdot)$. Ésta ha sido obtenida por Cao, López de Ullibarri, Janssen y Veraverbeke (2005) expresando el comportamiento límite del resto $S_n(\cdot)$ en probabilidad. Nosotros modificaremos esa demostración para obtener resultados de forma casi segura y en media. Para ello haremos uso de varios resultados que se demuestran, en forma de lemas, a continuación del siguiente teorema.

Teorema 2.4.1 (*Representación casi segura y en media de $\Lambda_n^P(\cdot) - \Lambda_F(\cdot)$ Bajo las condiciones (K1), (H3), (H4), (p2) y (v1) se verifica para todo $t \in [0, t_H]$ lo siguiente:*

$$\Lambda_n^P(t) - \Lambda_F(t) = \overline{\Lambda_n^P}(t) - \Lambda_F(t) + S_n(t),$$

donde

$$\overline{\Lambda_n^P}(t) = \Lambda_F(t) + \frac{1}{n} \sum_{i=1}^n (g_1(t, Z_i) - g_2(t, Z_i) + g_3(t, Z_i, \delta_i)) \quad (2.11)$$

con

$$g_1(t, Z) = \frac{p(t)}{1 - H(t)} (\mathbf{1}_{\{Z \leq t\}} - H(t)), \quad (2.12)$$

$$g_2(t, Z) = \int_0^t \frac{\mathbf{1}_{\{Z \leq v\}} - H(v)}{1 - H(v)} p'(v) dv, \quad (2.13)$$

$$g_3(t, Z, \delta) = \int_0^t K_b(v - Z) \frac{\delta - p(v)}{1 - H(v)} dv, \quad (2.14)$$

$$\sup_{0 \leq t \leq t_H} |S_n(t)| = O\left(\left(b^2 + (nb)^{-1/2} (\log n)^{1/2}\right)^2\right) \quad c.s.$$

y

$$\sup_{0 \leq t \leq t_H} \mathbb{E}[S_n^2(t)] = O\left(\left(b^4 + (nb)^{-1}\right)^2\right).$$

Demostración.

Al igual que en la demostración de Cao, López de Ullibarri, Janssen y Veraverbeke (2005), empezamos descomponiendo ahora la diferencia $\Lambda_n^P(\cdot) - \Lambda_F(\cdot)$ en tres sumandos:

$$\begin{aligned} \Lambda_n^P(t) - \Lambda_F(t) &= \int_0^t p_n(v) d\Lambda_n^H(v) - \int_0^t p(v) d\Lambda_H(v) \\ &= \int_0^t p(v) d(\Lambda_n^H(v) - \Lambda_H(v)) + \int_0^t (p_n(v) - p(v)) d\Lambda_H(v) \\ &\quad + \int_0^t (p_n(v) - p(v)) d(\Lambda_n^H(v) - \Lambda_H(v)) = (I) + (II) + (III), \end{aligned}$$

donde

$$H_n(t) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{\{Z_i \leq t\}} \quad \text{y} \quad \Lambda_n^H(t) = \int_0^t (1 - H_n(v-))^{-1} dH_n(v)$$

son los estimadores empíricos de $H(\cdot)$ y $\Lambda_H(\cdot)$ respectivamente.

Haciendo uso de la representación del estimador empírico de la razón de fallo acumulativa expresamos el primer sumando de la descomposición de la siguiente manera:

$$\Lambda_n^H(t) - \Lambda_H(t) = \frac{H_n(t) - H(t)}{1 - H(t)} + r_n(t),$$

donde

$$\sup_{0 \leq t \leq t_H} |r_n(t)| = O(n^{-1} \log n) \quad \text{c.s.} \quad \text{y} \quad \sup_{0 \leq t \leq t_H} \mathbb{E} [|r_n(t)|^2] = O(n^{-2}).$$

Las cotas del término residual uniforme y en media han sido obtenidas por Gijbels y Wang (1993) y Sánchez Selloero (2001) respectivamente para el caso más general de censura aleatoria por la derecha y truncamiento por la izquierda.

Por la aplicación de sucesivas integraciones por partes, el término (I) es igual a

$$\begin{aligned} (I) &= \int_0^t p(v) d(\Lambda_n^H(v) - \Lambda_H(v)) = \int_0^t p(v) d\left(\frac{H_n(v) - H(v)}{1 - H(v)} + r_n(v)\right) \\ &= \int_0^t \frac{p(v)}{1 - H(v)} dH_n(v) + \int_0^t \frac{p(v)h(v)}{(1 - H(v))^2} H_n(v) dv \\ &\quad - \int_0^t \frac{p(v)h(v)}{1 - H(v)} dv - \int_0^t \frac{p(v)H(v)h(v)}{(1 - H(v))^2} dv + \int_0^t p(v) dr_n(v) \\ &= \int_0^t \frac{p(v)}{1 - H(v)} dH_n(v) - \int_0^t \frac{p(v)h(v)}{(1 - H(v))^2} (1 - H_n(v)) dv + S_{n1}(t) \\ &= \frac{p(t)}{1 - H(t)} H_n(t) - \int_0^t \frac{p'(v)H_n(v)}{1 - H(v)} dv - \int_0^t \frac{p(v)h(v)}{(1 - H(v))^2} dv + S_{n1}(t) \\ &= \frac{p(t)}{1 - H(t)} (H_n(t) - H(t)) - \int_0^t \frac{p'(v)}{1 - H(v)} (H_n(v) - H(v)) dv + S_{n1}(t), \end{aligned}$$

donde

$$S_{n1}(t) = \int_0^t p(v) dr_n(v).$$

De este modo,

$$(I) = \frac{1}{n} \sum_{i=1}^n (g_1(t, Z_i) - g_2(t, Z_i)) + S_{n1}(t) \quad (2.15)$$

donde $g_1(\cdot, \cdot)$ y $g_2(\cdot, \cdot)$ son las funciones dadas en (2.12) y (2.13) respectivamente, y

$$\sup_{0 \leq t \leq t_H} |S_{n1}(t)| = O(n^{-1} \log n) \text{ c.s.} \quad (2.16)$$

y

$$\sup_{0 \leq t \leq t_H} \mathbb{E} [S_{n1}^2(t)] = O(n^{-2}).$$

Con respecto al término (III), aplicando de nuevo la descomposición de la razón de fallo acumulativa, obtenemos que:

$$\begin{aligned} (III) &= \int_0^t (p_n(v) - p(v)) d\left(\frac{H_n(v) - H(v)}{1 - H(v)}\right) + \int_0^t (p_n(v) - p(v)) dr_n(v) \\ &= \int_0^t \frac{p_n(v) - p(v)}{1 - H(v)} dH_n(v) - \int_0^t \frac{p_n(v) - p(v)}{1 - H(v)} dH(v) \\ &\quad + \int_0^t \frac{(p_n(v) - p(v))(H_n(v) - H(v))}{(1 - H(v))^2} dH(v) + \int_0^t (p_n(v) - p(v)) dr_n(v) \\ &= A_1(t) - (II) + A_2(t) + S_{n2}(t), \end{aligned}$$

donde, para

$$S_{n2}(t) = \int_0^t (p_n(v) - p(v)) dr_n(v),$$

resulta lo siguiente:

$$\sup_{0 \leq t \leq t_H} |S_{n2}(t)| = o(n^{-1} \log n) \text{ c.s.} \quad (2.17)$$

y

$$\sup_{0 \leq t \leq t_H} \mathbb{E} [S_{n2}^2(t)] = o(n^{-2}). \quad (2.18)$$

Por tanto,

$$\Lambda_n^P(t) - \Lambda_F(t) = \frac{1}{n} \sum_{i=1}^n (g_1(t, Z_i) - g_2(t, Z_i)) + A_1(t) + A_2(t) + S_{n12}(t)$$

siendo

$$S_{n12}(t) = S_{n1}(t) + S_{n2}(t).$$

Comprobaremos en primer lugar que el término $A_2(t)$ es despreciable, y a continuación descompondremos el término $A_1(t)$ en una suma de variables aleatorias i.i.d. más un término también despreciable de forma casi segura y en media.

El valor absoluto de $A_2(t)$ está acotado por

$$|A_2(t)| \leq \left\{ \sup_{0 \leq t \leq t_H} |p_n(t) - p(t)| \right\} \left\{ \sup_{0 \leq t \leq t_H} |H_n(t) - H(t)| \right\} \int_0^t \frac{h(v)}{(1-H(v))^2} dv,$$

de modo que, aplicando (2.10) y (2.9),

$$\sup_{0 \leq t \leq t_H} |A_2(t)| = O\left(n^{-1/2}b^2(\log n)^{1/2} + n^{-1}b^{-1/2}\log n\right) \quad c.s. \quad (2.19)$$

Con respecto a $A_2(t)$ operamos de la siguiente manera:

$$\begin{aligned} A_2^2(t) &= \left(\int_0^t \frac{(p_n(v) - p(v))(H_n(v) - H(v))}{(1-H(v))^2} h(v) dv \right)^2 \\ &\leq \sup_{0 \leq t \leq t_H} |H_n(t) - H(t)|^2 \int_0^t \frac{(p_n(v) - p(v))^2}{(1-H(v))^4} h^2(v) dv, \end{aligned}$$

de modo que, por la desigualdad de Cauchy-Schwarz,

$$\begin{aligned} &\mathbb{E}[A_2^2(t)] \\ &\leq \left[\mathbb{E} \left(\sup_{0 \leq t \leq t_H} |H_n(t) - H(t)|^4 \right) \right]^{1/2} \left[\mathbb{E} \left(\int_0^t \frac{(p_n(v) - p(v))^2}{(1-H(v))^4} h^2(v) dv \right)^2 \right]^{1/2} \\ &\leq \left[\mathbb{E} \left(\sup_{0 \leq t \leq t_H} |H_n(t) - H(t)|^4 \right) \right]^{1/2} \left[\int_0^t \frac{\mathbb{E}(p_n(v) - p(v))^4}{(1-H(v))^8} h^2(v) dv \right]^{1/2}. \end{aligned}$$

Bajo las condiciones (K1), (H3), (p2), la aplicación de los lemas 2.4.3 y 2.4.7 permiten concluir que

$$\sup_{0 \leq t \leq t_H} \mathbb{E}[A_2^2(t)] = O\left((nb)^{-2}\right). \quad (2.20)$$

Estudiamos ahora el término $A_1(t)$. Empezamos descomponiéndolo en dos sumandos, $A_{11}(t)$ y $A_{12}(t)$, el segundo de los cuales será despreciable:

$$\begin{aligned} A_1(t) &= \int_0^t \frac{p_n(v) - p(v)}{1-H(v)} dH_n(v) \\ &= \frac{1}{n^2} \sum_{i,j=1}^n \frac{K_b(Z_i - Z_j)(\delta_j - p(Z_i))}{(1-H(Z_i))h(Z_i)} \mathbf{1}_{\{Z_i \leq t\}} \\ &\quad - \frac{1}{n} \sum_{i=1}^n \frac{(p_n(Z_i) - p(Z_i))(h_n(Z_i) - h(Z_i))}{(1-H(Z_i))h(Z_i)} \mathbf{1}_{\{Z_i \leq t\}} = A_{11}(t) - A_{12}(t). \end{aligned}$$

El valor absoluto del término $A_{12}(t)$ se puede acotar por

$$\begin{aligned}
|A_{12}(t)| &= \left| \frac{1}{n} \sum_{i=1}^n \frac{(p_n(Z_i) - p(Z_i))(h_n(Z_i) - h(Z_i)) \mathbf{1}_{\{Z_i \leq t\}}}{(1 - H(Z_i)) h(Z_i)} \right| \\
&\leq \left\{ \sup_{0 \leq t \leq t_H} |p_n(t) - p(t)| \right\} \left\{ \sup_{0 \leq t \leq t_H} |h_n(t) - h(t)| \right\} \\
&\quad \times \left(\int_0^t \frac{dH_n(v)}{(1 - H(v)) h(v)} \right).
\end{aligned}$$

Por la hipótesis (v1) podemos aplicar el lema 1 de Mack y Silverman (1982),

$$\sup_{0 \leq t \leq t_H} |h_n(t) - h(t)| = O\left(b^2 + (nb)^{-1/2} (\log n)^{1/2}\right) \quad c.s.$$

el conocido para la función de regresión dada en (2.10), y la ley fuerte de los grandes números a la media muestral que aparece como último factor. Así, teniendo en cuenta además (H4), se concluye

$$\sup_{0 \leq t \leq t_H} |A_{12}(t)| = O\left(\left(b^2 + (nb)^{-1/2} (\log n)^{1/2}\right)^2\right) \quad c.s. \quad (2.21)$$

La acotación del momento de orden dos de $A_{12}(t)$ se puede hacer de la siguiente manera. En primer lugar, separamos la integral en dos términos, y acotaremos la esperanza del cuadrado de cada uno de ellos:

$$\begin{aligned}
A_{12}(t) &= \int_0^t \frac{(p_n(v) - p(v))(h_n(v) - h(v))}{(1 - H(v)) h(v)} dH_n(v) \\
&= \int_0^t \frac{(p_n(v) - p(v))(h_n(v) - h(v))}{(1 - H(v)) h(v)} d(H_n(v) - H(v)) \\
&\quad + \int_0^t \frac{(p_n(v) - p(v))(h_n(v) - h(v))}{(1 - H(v)) h(v)} dH(v) \\
&= A_{121}(t) + A_{122}(t).
\end{aligned}$$

La esperanza del cuadrado del primer término se puede acotar por

$$\begin{aligned}
\mathbb{E}[A_{121}^2(t)] &\leq \mathbb{E} \left[\sup_{0 \leq t \leq t_H} |p_n(t) - p(t)|^2 \sup_{0 \leq t \leq t_H} |h_n(t) - h(t)|^2 \right. \\
&\quad \left. \times \left(\int_0^t \frac{d|H_n(v) - H(v)|}{(1 - H(v)) h(v)} \right)^2 \right]
\end{aligned}$$

$$\begin{aligned} &\leq \mathbb{E} \left[\sup_{0 \leq t \leq t_H} |p_n(t) - p(t)|^4 \right]^{1/2} \mathbb{E} \left[\sup_{0 \leq t \leq t_H} |h_n(t) - h(t)|^4 \right]^{1/2} \\ &\quad \times \mathbb{E} \left[\left(\int_0^t \frac{d|H_n(v) - H(v)|}{(1 - H(v)) h(v)} \right)^4 \right]^{1/2}. \end{aligned}$$

Dado que el valor absoluto de la integral del tercer término es menor o igual que

$$\begin{aligned} &\sup_{0 \leq t \leq t_H} \left| \int_0^t \frac{d|H_n(v) - H(v)|}{(1 - H(v)) h(v)} \right| \\ &\leq \sup_{0 \leq t \leq t_H} |H_n(t) - H(t)| \left[\frac{1}{(1 - H(t_H)) \mu} + \int_0^{t_H} \left| \left(\frac{1}{(1 - H(v)) h(v)} \right)' \right| dv \right], \end{aligned}$$

en virtud de (H4) y de los lemas 2.4.3, 2.4.5 y 2.4.7, directamente se obtiene que

$$\sup_{0 \leq t \leq t_H} \mathbb{E} [A_{121}^2(t)] = O\left(\frac{1}{(nb^2)^2}\right) O\left(\frac{1}{n}\right) = O\left(\frac{1}{n^3 b^4}\right). \quad (2.22)$$

En relación al segundo término de la descomposición de $A_{12}(\cdot)$, actuamos de la siguiente manera. En el paso (a) aplicamos la desigualdad por la que $(ff)^2 \leq f^2$, y en los pasos (b) y (c) la desigualdad de Cauchy-Schwarz:

$$\begin{aligned} \mathbb{E} [A_{122}^2(t)] &= \mathbb{E} \left[\left(\int_0^t \frac{(p_n(v) - p(v))(h_n(v) - h(v))}{1 - H(v)} dv \right)^2 \right] \\ &\stackrel{(a)}{\leq} \mathbb{E} \left[\int_0^t \frac{(p_n(v) - p(v))^2 (h_n(v) - h(v))^2}{(1 - H(v))^2} dv \right] \\ &\stackrel{(b)}{\leq} \int_0^t \frac{\mathbb{E} [(p_n(v) - p(v))^4]^{1/2} \mathbb{E} [(h_n(v) - h(v))^4]^{1/2}}{(1 - H(v))^2} dv \\ &\stackrel{(c)}{\leq} \left(\int_0^t \frac{\mathbb{E} [(p_n(v) - p(v))^4]}{(1 - H(v))^4} dv \right)^{1/2} \left(\int_0^t \frac{\mathbb{E} [(h_n(v) - h(v))^4]}{(1 - H(v))^4} dv \right)^{1/2}. \end{aligned}$$

Aplicando los lemas 2.4.8 y 2.4.9, teniendo en cuenta que $[0, t_H]$ es un intervalo compacto, y bajo la condición (H4), concluimos

$$\sup_{0 \leq t \leq t_H} \mathbb{E} [A_{122}^2(t)] = O\left(\left(b^4 + (nb)^{-1}\right)^2\right). \quad (2.23)$$

Con respecto al término $A_{11}(t)$ introducimos la siguiente notación

$$\varphi(t, Z_i, Z_j, \delta_j) = \frac{K_b(Z_i - Z_j)(\delta_j - p(Z_i))}{(1 - H(Z_i)) h(Z_i)} \mathbf{1}_{\{Z_i \leq t\}},$$

por lo que

$$\begin{aligned} A_{11}(t) &= \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n \varphi(t, Z_i, Z_j, \delta_j) + \frac{1}{n^2} \sum_{i=1}^n \varphi(t, Z_i, Z_i, \delta_i) \\ &= \frac{1}{n^2} \sum_{i=1}^n \sum_{\substack{j=1 \\ i \neq j}}^n \varphi(t, Z_i, Z_j, \delta_j) + S_{n3}(t). \end{aligned}$$

El término $S_{n3}(t)$ se puede acotar de la siguiente manera:

$$\begin{aligned} |S_{n3}(t)| &= \left| \frac{1}{n^2} \sum_{i=1}^n \varphi(t, Z_i, Z_i, \delta_i) \right| \\ &= \left| \frac{1}{n^2} \frac{1}{b} K(0) \sum_{i=1}^n \frac{\delta_i - p(Z_i)}{(1 - H(Z_i)) h(Z_i)} \mathbf{1}_{\{Z_i \leq t\}} \right| \\ &\leq \frac{2}{nb} K(0) \frac{1}{n} \sum_{i=1}^n \frac{\mathbf{1}_{\{Z_i \leq t\}}}{(1 - H(Z_i)) h(Z_i)} = \frac{2}{nb} K(0) \int_0^t \frac{dH_n(v)}{(1 - H(v)) h(v)}, \end{aligned}$$

de modo que, bajo la condición (H4), entonces

$$\sup_{0 \leq t \leq t_H} |S_{n3}(t)| = O\left(\frac{1}{nb}\right) \quad c.s. \quad \text{y} \quad \sup_{0 \leq t \leq t_H} \mathbb{E}[S_{n3}^2(t)] = O\left(\frac{1}{n^2 b^2}\right). \quad (2.24)$$

Mediante la simetrización del núcleo $\varphi(\cdot, \cdot, \cdot, \cdot)$,

$$\psi(t, Z_i, \delta_i, Z_j, \delta_j) = \frac{1}{2} (\varphi(t, Z_i, Z_j, \delta_j) + \varphi(t, Z_j, Z_i, \delta_i)),$$

tenemos que el primer término de $A_{11}(t)$ se puede expresar, salvo términos despreciables, como el siguiente U -proceso:

$$U_n(t) = \binom{n}{2}^{-1} \sum_{1 \leq i < j \leq n} \psi(t, Z_i, \delta_i, Z_j, \delta_j), \quad (2.25)$$

al que calcularemos su proyección de Hajek (ver Hajek y Sidak (1967)). Para ello, definimos las siguientes funciones:

$$\begin{aligned} \theta(t) &= \mathbb{E}[\psi(t, Z_1, \delta_1, Z_2, \delta_2)], \\ \psi_1(t, Z_1, \delta_1) &= \mathbb{E}[\psi(t, Z_1, \delta_1, Z_2, \delta_2) \mid Z_1, \delta_1] - \theta(t), \\ \psi_2(t, Z_1, \delta_1, Z_2, \delta_2) &= \psi(t, Z_1, \delta_1, Z_2, \delta_2) - \mathbb{E}[\psi(t, Z_1, \delta_1, Z_2, \delta_2) \mid Z_1, \delta_1] \\ &\quad - \mathbb{E}[\psi(t, Z_1, \delta_1, Z_2, \delta_2) \mid Z_2, \delta_2] + \theta(t), \end{aligned}$$

donde

$$\begin{aligned}
\mathbb{E}[\psi(t, Z_1, \delta_1, Z_2, \delta_2) \mid Z_1, \delta_1] &= \frac{1}{2} \mathbb{E}[(\varphi(t, Z_1, Z_2, \delta_2) + \varphi(t, Z_2, Z_1, \delta_1)) \mid Z_1, \delta_1] \\
&= \frac{1}{2} \mathbb{E}[\varphi(t, Z_1, Z_2, \delta_2) \mid Z_1, \delta_1] + \frac{1}{2} \mathbb{E}[\varphi(t, Z_2, Z_1, \delta_1) \mid Z_1, \delta_1] \\
&= \frac{1}{2} G_1(t, Z_1) + \frac{1}{2} G_2(t, Z_1, \delta_1),
\end{aligned}$$

con

$$\begin{aligned}
G_1(t, Z_1) &= \frac{\mathbf{1}_{\{Z_1 \leq t\}}}{(1 - H(Z_1)) h(Z_1)} \int K_b(Z_1 - v) (p(v) - p(Z_1)) h(v) dv \quad (2.26) \\
G_2(t, Z_1, \delta_1) &= \int_0^t K_b(v - Z_1) \frac{\delta_1 - p(v)}{1 - H(v)} dv = g_3(t, Z_1, \delta_1).
\end{aligned}$$

siendo $g_3(\cdot, \cdot, \cdot)$ la función dada en (2.14). Consideremos en lo que sigue

$$Q_n(t) = \binom{n}{2}^{-1} \sum_{1 \leq i < j \leq n} \psi_2(t, Z_i, \delta_i, Z_j, \delta_j). \quad (2.27)$$

Es inmediato comprobar que

$$\begin{aligned}
\psi(t, Z_1, \delta_1, Z_2, \delta_2) &= \frac{1}{2} g_3(t, Z_1, \delta_1) + \frac{1}{2} G_1(t, Z_1) + \frac{1}{2} g_3(t, Z_2, \delta_2) \\
&\quad + \frac{1}{2} G_1(t, Z_2) - \theta(t) + \psi_2(t, Z_1, \delta_1, Z_2, \delta_2),
\end{aligned}$$

de modo que

$$U_n(t) = \frac{1}{n} \sum_{i=1}^n g_3(t, Z_i, \delta_i) + \frac{1}{n} \sum_{i=1}^n (G_1(t, Z_i) - \theta(t)) + Q_n(t).$$

Hemos obtenido así una representación del U -proceso $U_n(\cdot)$ en términos de dos sumas de variables independientes e idénticamente distribuidas, la segunda de las cuales será despreciable, y un U -proceso degenerado hasta el orden uno.

Llamemos $S_{n4}(t) = n^{-1} \sum_{i=1}^n (G_1(t, Z_i) - \theta(t)) + Q_n(t)$ al resto de la expresión del U -proceso $U_n(t)$:

$$U_n(t) = \frac{1}{n} \sum_{i=1}^n g_3(t, Z_i, \delta_i) + S_{n4}(t).$$

Para demostrar que $S_{n4}(\cdot)$ es despreciable de forma casi segura y en media, empezaremos acotando el resto $Q_n(\cdot)$.

En virtud al lema 2.4.11, que se probará al final de esta demostración, se tiene que

$$\sup_{0 \leq t \leq t_H} |Q_n(t)| = O\left(\frac{\log \log n}{nb^{1/2}}\right) \text{ c.s. y } \sup_{0 \leq t \leq t_H} \mathbb{E}[Q_n^2(t)] = O\left(\frac{(\log \log n)^2}{nb^2}\right) \quad (2.28)$$

Falta ahora acotar el otro componente de $S_{n4}(\cdot)$, el sumatorio de variables aleatorias i.i.d.:

$$\frac{1}{n} \sum_{i=1}^n (G_1(t, Z_i) - \theta(t)).$$

El lema 2.4.12, que se probará a continuación del lema 2.4.11, establece una cota casi segura para esta suma de variables aleatorias centradas, de forma que, bajo las condiciones (K1), (H2) y (p1), se tiene que

$$\sup_{0 \leq t \leq t_H} \left| \frac{1}{n} \sum_{i=1}^n (G_1(t, Z_i) - \mathbb{E}[G_1(t, Z)]) \right| = O\left(n^{-1/2} b^2 (\log n)^{1/2}\right) \text{ c.s.} \quad (2.29)$$

El lema 2.4.13 establece la acotación en supremo de la esperanza de este sumatorio, dada por

$$\sup_{0 \leq t \leq t_H} \mathbb{E} \left[\frac{1}{n} \sum_{i=1}^n G_1(t, Z_i) - \mathbb{E}[G_1(t, Z)] \right]^2 = O(n^{-1} b^4). \quad (2.30)$$

Reuniendo los resultados (2.15)-(2.30) se obtiene la representación dada en (2.11).

■

Los siguientes lemas 2.4.11-2.4.13 son los resultados usados en la demostración del teorema anterior.

Lema 2.4.11 *Sea $Q_n(\cdot)$ el U -proceso degenerado de la descomposición de Hajek del U -proceso $U_n(\cdot)$ definido en (2.25), y sea $t_H < b_H$. Entonces, bajo las condiciones (K1) y (H3),*

$$\sup_{0 \leq t \leq t_H} |Q_n(t)| = O\left(n^{-1} b^{-1/2} \log \log n\right) \text{ c.s.}$$

y

$$\sup_{0 \leq t \leq t_H} \mathbb{E}[Q_n^2(t)] = O\left(n^{-2} b^{-1} (\log \log n)^2\right).$$

Demostración.

La idea consiste en comprobar que el U -proceso verifica todas las hipótesis del teorema 5.4.1 de De la Peña y Giné (1999) para aplicar, más concretamente, la desigualdad (5.4.6) que de él se deduce. Redefinimos el U -proceso degenerado $Q_n(\cdot)$ de la siguiente manera:

$$\begin{aligned} Q_n(t) &= \binom{n}{2}^{-1} \sum_{1 \leq i < j \leq n} \psi_2(t, Z_i, \delta_i, Z_j, \delta_j) \\ &= b^{-1/2} \binom{n}{2}^{-1} \sum_{1 \leq i < j \leq n} \psi_{2,b}(t, Z_i, \delta_i, Z_j, \delta_j), \end{aligned}$$

donde $\psi_{2,b}(t, Z_i, \delta_i, Z_j, \delta_j) = b^{1/2} \psi_2(t, Z_i, \delta_i, Z_j, \delta_j)$. Consideramos ahora las familias de funciones

$$\phi_{\delta,t}(Z_1, \delta_1, Z_2, \delta_2) = b^{1/2} K_b(Z_1 - Z_2) \frac{\delta_2}{(1 - H(Z_1)) h(Z_1)} \mathbf{1}_{\{Z_1 \leq t\}}$$

y

$$\phi_{p,t}(Z_1, \delta_1, Z_2, \delta_2) = b^{1/2} K_b(Z_1 - Z_2) \frac{p(Z_1)}{(1 - H(Z_1)) h(Z_1)} \mathbf{1}_{\{Z_1 \leq t\}}.$$

Estas familias son una clase VC-subgrafo, por ser producto de una función no negativa por una indicadora de conjuntos de una clase VC. Sus envolturas son

$$\Phi_{\delta}(Z_1, \delta_1, Z_2, \delta_2) = b^{1/2} K_b(Z_1 - Z_2) \frac{\delta_2}{(1 - H(Z_1)) h(Z_1)}$$

y

$$\Phi_p(Z_1, \delta_1, Z_2, \delta_2) = b^{1/2} K_b(Z_1 - Z_2) \frac{p(Z_1)}{(1 - H(Z_1)) h(Z_1)}$$

respectivamente, cuyos momentos de orden 2 son

$$\begin{aligned} \mathbb{E} [\Phi_{\delta}^2(Z_1, \delta_1, Z_2, \delta_2)] &= b \int_0^{t_H} \int_0^{t_H} K_b^2(u - v) \frac{p(v) h(v)}{(1 - H(u))^2 h(u)} dudv \\ &= \int_0^{t_H} \int_0^{t_H} K^2(z) \frac{p(u - bz) h(u - bz)}{(1 - H(u))^2 h(u)} dudz \leq C_1 \end{aligned}$$

y

$$\begin{aligned} \mathbb{E} [\Phi_p^2(Z_1, \delta_1, Z_2, \delta_2)] &= b \int_0^{t_H} \int_0^{t_H} K_b^2(u - v) \frac{p(u) h(v)}{(1 - H(u))^2 h(u)} dudv \\ &= \int_0^{t_H} \int_0^{t_H} K^2(z) \frac{p(u) h(u - bz)}{(1 - H(u))^2 h(u)} dudz \leq C_2 \end{aligned}$$

para todo $n \geq n_0$ suficientemente grande.

El núcleo $\psi_{2,b}(t, Z_1, \delta_1, Z_2, \delta_2)$ del U -proceso $Q_n(t)$ es, entonces, la siguiente combinación lineal de VC-subgrafos:

$$\frac{1}{2} (\phi_{\delta,t}(Z_1, \delta_1, Z_2, \delta_2) - \phi_{p,t}(Z_1, \delta_1, Z_2, \delta_2)) + \phi_{\delta,t}(Z_2, \delta_2, Z_1, \delta_1) - \phi_{p,t}(Z_2, \delta_2, Z_1, \delta_1)]$$

con envoltura de cuadrado integrable:

$$\begin{aligned}\Phi(Z_1, \delta_1, Z_2, \delta_2) &= \frac{1}{2} (\Phi_\delta(Z_1, \delta_1, Z_2, \delta_2) + \Phi_p(Z_1, \delta_1, Z_2, \delta_2) \\ &\quad + \Phi_\delta(Z_2, \delta_2, Z_1, \delta_1) + \Phi_p(Z_2, \delta_2, Z_1, \delta_1)).\end{aligned}$$

Por tanto, en aplicación de la ley del logaritmo iterado para U -procesos (teorema 5.4.1 de De la Peña y Giné (1999) y más concretamente su consecuencia (5.4.6)), y teniendo en cuenta que $\mathbb{E}[\Phi^2(Z_1, \delta_1, Z_2, \delta_2)] = O(1)$, obtenemos que

$$\sup_{0 \leq t \leq t_H} \left| \sum_{1 \leq i < j \leq n} \psi_{2,b}(t, Z_i, \delta_i, Z_j, \delta_j) \right| = O(n \log \log n) \quad c.s.$$

y, por tanto,

$$\begin{aligned}\sup_{0 \leq t \leq t_H} |Q_n(t)| &= b^{-1/2} \binom{n}{2}^{-1} \sup_{0 \leq t \leq t_H} \left| \sum_{1 \leq i < j \leq n} \psi_{2,b}(t, Z_i, \delta_i, Z_j, \delta_j) \right| \\ &= O\left(n^{-1} b^{-1/2} \log \log n\right) \quad c.s.\end{aligned}$$

Para acotar uniformemente el momento de orden 2 de $Q_n(t)$, $\mathbb{E}[Q_n^2(t)]$, utilizaremos de nuevo el teorema 5.4.1 de De la Peña y Giné (1999):

$$\begin{aligned}\sup_{0 \leq t \leq t_H} \mathbb{E}[Q_n^2(t)] &= b^{-1} \binom{n}{2}^{-2} (2n \log \log n)^2 \\ &\quad \times \sup_{0 \leq t \leq t_H} \mathbb{E} \left[\left(\frac{1}{2n \log \log n} \sum_{1 \leq i < j \leq n} \psi_{2,b}(t, Z_i, \delta_i, Z_j, \delta_j) \right)^2 \right] \\ &\leq b^{-1} \binom{n}{2}^{-2} (2n \log \log n)^2 \\ &\quad \times \sup_{0 \leq t \leq t_H} \mathbb{E} \left[\sup_{n \in \mathbb{N}} \left(\frac{1}{2n \log \log n} \sum_{1 \leq i < j \leq n} \psi_{2,b}(t, Z_i, \delta_i, Z_j, \delta_j) \right)^2 \right] \\ &\leq b^{-1} \binom{n}{2}^{-2} (2n \log \log n)^2 \sup_{0 \leq t \leq t_H} C \mathbb{E}[\Phi^2(Z_1, \delta_1, Z_2, \delta_2)]\end{aligned}$$

para cierta constante $C > 0$, de modo que

$$\sup_{0 \leq t \leq t_H} \mathbb{E}[Q_n^2(t)] = O\left(n^{-2} b^{-1} (\log \log n)^2\right).$$

■

Lema 2.4.12 Sea $G_1(\cdot, \cdot)$ la función dada en (2.26); bajo las condiciones (K1), (H1), (H4) y (p1), se tiene, para $t_H < b_H$, que

$$\sup_{0 \leq t \leq t_H} \left| \frac{1}{n} \sum_{i=1}^n (G_1(t, Z_i) - \mathbb{E}[G_1(t, Z)]) \right| = O\left(n^{-1/2} b^2 (\log n)^{1/2}\right) \quad c.s.$$

Demostración.

La función $G_1(\cdot, \cdot)$ es igual a

$$G_1(t, Z) = \frac{\mathbf{1}_{\{Z \leq t\}}}{(1 - H(Z)) h(Z)} \int K_b(Z - u) (p(u) - p(Z)) h(u) du.$$

Si aplicamos el cambio de variable $s = Z - vb$ (paso (a)) y un desarrollo de Taylor en las funciones $p(\cdot)$ y $h(\cdot)$ (paso (b)), bajo las hipótesis (H1) y (p1) podemos expresar la función $G_1(\cdot, \cdot)$ de la siguiente manera:

$$\begin{aligned} G_1(t, Z) &= \frac{\mathbf{1}_{\{Z \leq t\}}}{(1 - H(Z)) h(Z)} \int K_b(Z - u) (p(u) - p(Z)) h(u) du \\ &\stackrel{(a)}{=} \frac{\mathbf{1}_{\{Z \leq t\}}}{(1 - H(Z)) h(Z)} \int K(v) (p(Z - vb) - p(Z)) h(Z - vb) dv \\ &\stackrel{(b)}{=} \frac{\mathbf{1}_{\{Z \leq t\}}}{(1 - H(Z)) h(Z)} \\ &\quad \times \int K(v) \left(-vb p'(Z) + \frac{1}{2} v^2 b^2 p''(\zeta_{Z,t}) \right) (h(Z) - vb h'(\xi_{Z,t})) dv \\ &= \frac{\mathbf{1}_{\{Z \leq t\}}}{(1 - H(Z)) h(Z)} \int K(v) \left(\frac{1}{2} v^2 b^2 p''(\zeta_{Z,t}) h(Z) \right. \\ &\quad \left. + v^2 b^2 p'(Z) h'(\xi_{Z,t}) - \frac{1}{2} v^3 b^3 p''(\zeta_{Z,t}) h'(\xi_{Z,t}) \right) dv, \end{aligned}$$

donde $\zeta_{Z,t}$ y $\xi_{Z,t}$ son dos puntos intermedios entre Z y $Z - vb$. Bajo (H4),

$$\begin{aligned} \sup_{0 \leq z \leq t \leq t_H} |G_1(t, z)| &\leq \frac{1}{(1 - H(t_H))} \sup_{0 \leq t \leq t_H} \left| \frac{1}{h(t)} \right| \\ &\quad \times \left(\frac{1}{2} b^2 (\|p''\|_\infty \|h\|_\infty + \|p'\|_\infty \|h'\|_\infty) d_K \right. \\ &\quad \left. + \frac{1}{2} b^3 \|p''\|_\infty \|h'\|_\infty \int |v|^3 K(v) dv \right) = C b^2 \end{aligned}$$

para cierta constante $C > 0$, para todo $Z_i, i = 1, \dots, n$ y todo $t \in [0, t_H]$, siendo $\|h\|_\infty = \sup_{0 \leq t \leq t_H} |h(t)|$.

Dado que el intervalo $D = [0, t_H]$ es compacto, podemos hacer una partición en L_n intervalos tales que

$$\bigcup_{k=1}^{L_n} (D \cap I_k) = D \quad \text{y que } I_k \cap I_j = \emptyset \quad \text{si } k \neq j.$$

Entonces,

$$\begin{aligned}
\sup_{u \in D} \left| \frac{1}{n} \sum_{i=1}^n G_1(u, Z_i) - \mathbb{E}[G_1(u, Z)] \right| &= \max_{1 \leq k \leq L_n} \sup_{u \in I_k \cap D} \left| \frac{1}{n} \sum_{i=1}^n G_1(u, Z_i) - \mathbb{E}[G_1(u, Z)] \right| \\
&\leq \max_{1 \leq k \leq L_n} \sup_{u \in I_k \cap D} \left| \frac{1}{n} \sum_{i=1}^n G_1(u, Z_i) - \frac{1}{n} \sum_{i=1}^n G_1(u_k, Z_i) \right| \\
&\quad + \max_{1 \leq k \leq L_n} \left| \frac{1}{n} \sum_{i=1}^n G_1(u_k, Z_i) - \mathbb{E}[G_1(u_k, Z)] \right| \\
&\quad + \max_{1 \leq k \leq L_n} \sup_{u \in I_k \cap D} |\mathbb{E}[G_1(u_k, Z)] - \mathbb{E}[G_1(u, Z)]| \\
&= Q_1 + Q_2 + Q_3
\end{aligned}$$

donde u_k es el centro de los intervalos de la partición. Analizamos a continuación cada uno de los términos Q_i .

El sumando Q_1 se puede acotar de la siguiente manera:

$$\begin{aligned}
Q_1 &= \max_{1 \leq k \leq L_n} \sup_{u \in I_k \cap D} \left| \frac{1}{n} \sum_{i=1}^n G_1(u, Z_i) - \frac{1}{n} \sum_{i=1}^n G_1(u_k, Z_i) \right| \\
&= \max_{1 \leq k \leq L_n} \sup_{u \in I_k \cap D} \left| \int G_1(u, v) dH_n(v) - \int G_1(u_k, v) dH_n(v) \right| \\
&\leq \sup_{0 \leq Z \leq t \leq t_H} |G_1(t, Z)| \max_{1 \leq k \leq L_n} \sup_{u \in I_k \cap D} \left| \int_{u_k}^u dH_n(v) \right| \\
&\leq Cb^2 \max_{1 \leq k \leq L_n} \sup_{u \in I_k \cap D} |H_n(u) - H_n(u_k)|.
\end{aligned}$$

La idea es tomar la partición lo suficientemente fina para que la última cota pueda ser controlada a través de $\max_{1 \leq k \leq L_n} \sup_{u \in I_k \cap D} |u - u_k|$. Sea L_n el número de intervalos y sea $l_n \rightarrow 0$ la longitud de cada uno de ellos. Si tomamos l_n de orden exacto $(n^{-1/2} (\log n)^{1/2})$ entonces

$$Q_1 = O\left(n^{-1/2} b^2 (\log n)^{1/2}\right) \text{ c.s.}$$

Esta misma idea se aplicará al término Q_3 :

$$\begin{aligned}
Q_3 &= \max_{1 \leq k \leq L_n} \sup_{u \in I_k \cap D} |\mathbb{E}[G_1(u_k, Z)] - \mathbb{E}[G_1(u, Z)]| \\
&= \max_{1 \leq k \leq L_n} \sup_{u \in I_k \cap D} \left| \mathbb{E} \left(\frac{\mathbf{1}_{\{Z \leq u\}} - \mathbf{1}_{\{Z \leq u_k\}}}{(1 - H(Z)) h(Z)} \int K_b(Z - s) (p(s) - p(Z)) h(s) ds \right) \right| \\
&= \max_{1 \leq k \leq L_n} \sup_{u \in I_k \cap D} \left| \int_{u_k}^u \frac{1}{1 - H(z)} \int K_b(z - s) (p(s) - p(z)) h(s) ds dz \right|
\end{aligned}$$

$$\leq \sup_{0 \leq z \leq t \leq t_H} |G_1(t, z)| \sup_{0 \leq t \leq t_H} |h(t)| \max_{1 \leq k \leq L_n} \sup_{u \in I_k \cap D} |u - u_k|$$

para todo $t \in [0, t_H]$. Si tomamos l_n del mismo orden que para Q_1 , es decir, $l_n = O(n^{-1/2} (\log n)^{1/2})$, entonces

$$Q_3 = O(n^{-1/2} b^2 (\log n)^{1/2}).$$

Y por último, para el sumando Q_2 se tiene lo siguiente,

$$Q_2 = \max_{1 \leq k \leq L_n} \left| \frac{1}{n} \sum_{i=1}^n G_1(u_k, Z_i) - \mathbb{E}[G_1(u_k, Z)] \right|.$$

Para demostrar que $Q_2 = O(a_n)$ c.s. con $a_n \rightarrow 0$, veremos que

$$P(|Q_2| > a_n \text{ infinitas veces}) = 0.$$

La idea consiste en acotar la probabilidad $P(|Q_2| > a_n)$ por una sucesión b_n de modo que la serie $\sum_{n=1}^{\infty} b_n$ sea convergente, y aplicar posteriormente el lema de Borel-Cantelli. Empezamos aplicando la desigualdad de Bonferroni a la siguiente probabilidad:

$$\begin{aligned} P(|Q_2| > \varepsilon) &= P\left(\max_{1 \leq k \leq L_n} \left| \frac{1}{n} \sum_{i=1}^n G_1(u_k, Z_i) - \mathbb{E}[G_1(u_k, Z)] \right| > \varepsilon\right) \\ &\leq \sum_{1 \leq k \leq L_n} P\left(\left| \frac{1}{n} \sum_{i=1}^n G_1(u_k, Z_i) - \mathbb{E}[G_1(u_k, Z)] \right| > \varepsilon\right) \\ &\leq L_n \max_{1 \leq k \leq L_n} P\left(\left| \frac{1}{n} \sum_{i=1}^n G_1(u_k, Z_i) - \mathbb{E}[G_1(u_k, Z)] \right| > \varepsilon\right) \\ &= L_n \max_{1 \leq k \leq L_n} P\left(\left| \sum_{i=1}^n (G_1(u_k, Z_i) - \mathbb{E}[G_1(u_k, Z)]) \right| > n\varepsilon\right). \end{aligned}$$

Trataremos de aplicar a la probabilidad

$$P\left(\left| \sum_{i=1}^n (G_1(u_k, Z_i) - \mathbb{E}[G_1(u_k, Z)]) \right| > n\varepsilon\right)$$

la desigualdad de Bernstein, por la cual si W_1, W_2, \dots, W_n son n variables aleatorias independientes verificando que

$$P(|W_i - \mathbb{E}[W_i]| \leq m) = 1 \text{ para cada } i = 1, 2, \dots, n, \text{ donde } m < \infty$$

entonces, para cada $t > 0$ y para todo $n = 1, 2, \dots$ se tiene

$$P \left(\left| \sum_i W_i - \sum_i \mathbb{E}[W_i] \right| \geq nt \right) \leq 2 \exp \left\{ - \frac{(nt)^2}{2 \sum_i \text{Var}[W_i] + \frac{2}{3}mnt} \right\}$$

Puesto que las variables $G_1(\cdot, Z)$ verifican que

$$\sup_{0 \leq Z_i \leq t \leq t_H} |G_1(t, Z_i)| \leq Cb^2 \text{ para todo } Z_i \text{ y todo } t \in [0, t_H]$$

entonces

$$P(|G_1(t, Z_i) - \mathbb{E}[G_1(t, Z_i)]| \leq m) = 1 \text{ para cada } i = 1, 2, \dots, n$$

con $m = 2Cb^2$. Además,

$$\text{Var}(G_1(t, Z)) \leq 2\mathbb{E} \left(\sup_{0 \leq t \leq t_H} |G_1(t, Z)| \right)^2 \leq 2C^2b^4.$$

Entonces, aplicando la desigualdad de Bernstein y la independiencia de las variables $G_1(\cdot, Z)$ (paso (a)), que $\text{Var}(G_1(\cdot, Z)) \leq 2C^2b^4$ y que $m = 2Cb^2$ (paso (b)), y finalmente simplificando (paso (c)), obtenemos las siguientes desigualdades:

$$\begin{aligned} P \left(\left| \sum_{i=1}^n (G_1(u_k, Z_i) - \mathbb{E}[G_1(u_k, Z_i)]) \right| > n\varepsilon \right) &\stackrel{(a)}{\leq} 2 \exp \left\{ - \frac{(n\varepsilon)^2}{2n\text{Var}[G_1(t, Z)] + \frac{2}{3}mn\varepsilon} \right\} \\ &\stackrel{(b)}{\leq} 2 \exp \left\{ - \frac{n^2\varepsilon^2}{4nC^2b^4 + \frac{2}{3}2Cb^2n\varepsilon} \right\} \stackrel{(c)}{\leq} 2 \exp \left\{ - \frac{n\varepsilon^2}{4C^2b^4 + \frac{4}{3}Cb^2\varepsilon} \right\}. \end{aligned}$$

Si tomamos $\varepsilon = C_\varepsilon n^{-1/2} b^2 (\log n)^{1/2}$, entonces

$$\begin{aligned} P \left(\left| \sum_{i=1}^n (G_1(u_k, Z_i) - \mathbb{E}[G_1(u_k, Z_i)]) \right| > n\varepsilon \right) &\leq 2 \exp \left\{ - \frac{n\varepsilon^2}{4C^2b^4 + \frac{4}{3}Cb^2\varepsilon} \right\} \\ &= 2 \exp \left\{ - \frac{C_\varepsilon^2 b^4 \log n}{4C^2b^4 + \frac{4}{3}CC_\varepsilon b^4 n^{-1/2} (\log n)^{1/2}} \right\} \\ &= 2 \exp \left\{ - \frac{C_\varepsilon^2 \log n}{4C^2 + \frac{4}{3}CC_\varepsilon n^{-1/2} (\log n)^{1/2}} \right\}. \end{aligned}$$

Para un valor de n suficientemente grande se tiene que el denominador del exponente se puede acotar por una constante C_0 :

$$4C^2 + \frac{4}{3}CC_\varepsilon n^{-1/2} (\log n)^{1/2} \leq C_0 \text{ para todo } n \geq n_{C_0}.$$

Por tanto,

$$\begin{aligned} P\left(\left|\sum_{i=1}^n (G_1(u_k, Z_i) - \mathbb{E}[G_1(u_k, Z_i)])\right| > n\varepsilon\right) \\ \leq 2 \exp\left\{-\frac{C_\varepsilon^2 \log n}{4C^2 + \frac{4}{3}CC_\varepsilon n^{-1/2} (\log n)^{1/2}}\right\} \\ \leq 2 \exp\left\{-\frac{C_\varepsilon^2 \log n}{C_0}\right\} = 2 \exp\left(\log n^{-C_\varepsilon^2/C_0}\right) = 2n^{-C_\varepsilon^2/C_0}. \end{aligned}$$

Ahora bien, volviendo a la acotación inicial para la probabilidad $P(|Q_2| > \varepsilon)$, tenemos lo siguiente:

$$P(|Q_2| > \varepsilon) \leq L_n \max_{1 \leq k \leq L_n} P\left(\left|\sum_{i=1}^n (G_1(u_k, Z_i) - \mathbb{E}[G_1(u_k, Z)])\right| > n\varepsilon\right) \leq 2L_n n^{-C_\varepsilon^2/C_0},$$

donde L_n es el número de intervalos de la partición. Si, en consonancia con el orden fijado para l_n , tomamos L_n de forma que

$$L_n \leq C_L n^{1/2} (\log n)^{-1/2},$$

entonces basta tomar una cota C_ε suficientemente grande para que, fijado C_0 ,

$$\begin{aligned} \sum_{n=1}^{\infty} L_n n^{-C_\varepsilon^2/C_0} &\leq \sum_{n=1}^{\infty} C_L n^{1/2} (\log n)^{-1/2} n^{-C_\varepsilon^2/C_0} \\ &= C_L \sum_{n=1}^{\infty} n^{1/2-C_\varepsilon^2/C_0} (\log n)^{-1/2} < \infty. \end{aligned}$$

En ese caso, puesto que habíamos tomado $\varepsilon = C_\varepsilon n^{-1/2} b^2 (\log n)^{1/2}$, podemos concluir que

$$P(|Q_2| > \varepsilon) = P\left(|Q_2| > \varepsilon = C_\varepsilon n^{-1/2} b^2 (\log n)^{1/2} \quad \text{infinitas veces}\right) = 0,$$

de modo que

$$Q_2 = O\left(n^{-1/2} b^2 (\log n)^{1/2}\right) \quad c.s.$$

En resumen,

$$\sup_{0 \leq t \leq t_H} \left| \frac{1}{n} \sum_{i=1}^n G_1(t, Z_i) - \mathbb{E}[G_1(t, Z)] \right| = O\left(n^{-1/2} b^2 (\log n)^{1/2}\right) \quad c.s.$$

finalizando así la demostración. ■

Lema 2.4.13 Sea $G_1(\cdot, \cdot)$ la función dada en (2.26). Bajo las condiciones (K1), (H2) y (p1) se tiene, para un n suficientemente grande y para todo $t_H < b_H$,

$$\sup_{0 \leq t \leq t_H} \mathbb{E} \left[\frac{1}{n} \sum_{i=1}^n G_1(t, Z_i) - \mathbb{E}[G_1(t, Z)] \right]^2 = O(n^{-1}b^4).$$

Demostración.

El momento de orden dos de este término es igual a

$$\mathbb{E} \left[\frac{1}{n} \sum_{i=1}^n G_1(t, Z_i) - \mathbb{E}[G_1(t, Z)] \right]^2 = \frac{1}{n} \mathbb{E}[G_1^2(t, Z)] - \frac{1}{n} \theta^2(t).$$

Calculamos ahora el momento de orden dos de la variable $G_1(t, Z)$:

$$\begin{aligned} \mathbb{E}[G_1^2(t, Z)] &= \int_0^t \frac{1}{(1-H(u))^2 h(u)} \iint K_b(u-v_1) K_b(u-v_2) \\ &\quad \times (p(v_1) - p(u))(p(v_2) - p(u)) h(v_1) h(v_2) dv_1 dv_2 du \\ &= \int_0^t \frac{1}{(1-H(u))^2 h(u)} \iint K(z_1) K(z_2) (p(u-bz_1) - p(u)) \\ &\quad \times (p(u-bz_2) - p(u)) h(u-bz_1) h(u-bz_2) dz_1 dz_2 du \\ &= \int_0^t \frac{1}{(1-H(u))^2 h(u)} \iint K(z_1) K(z_2) (-bz_1 p'(u) + b^2 z_1^2 p''(\zeta_{u,b,z_1})) \\ &\quad \times (-bz_2 p'(u) + b^2 z_2^2 p''(\zeta_{u,b,z_2})) (h(u) - bz_1 h'(u) + \frac{1}{2} b^2 z_1^2 h''(\xi_{b,u,z_1})) \\ &\quad \times (h(u) - bz_2 h'(u) + \frac{1}{2} b^2 z_2^2 h''(\xi_{b,u,z_2})) dz_1 dz_2 du, \end{aligned}$$

donde $\zeta_{u,b,z_1}, \xi_{u,b,z_1}, \zeta_{u,b,z_2}$ y ξ_{u,b,z_2} son valores intermedios entre u y $u - bz_1, u - bz_2$ respectivamente. Bajo las hipótesis (K1), (H2) y (p1) se tiene que

$$\sup_{0 \leq t \leq t_H} \mathbb{E}[G_1^2(t, Z)] = O(b^4).$$

Como además se tenía que, bajo estas mismas condiciones, $\sup_{0 \leq t \leq t_H} |\theta(t)| = O(b^2)$, podemos concluir que

$$\sup_{0 \leq t \leq t_H} \mathbb{E} \left[\frac{1}{n} \sum_{i=1}^n G_1(t, Z_i) - \mathbb{E}[G_1(t, Z)] \right]^2 = O(n^{-1}b^4).$$

■

Finalmente, la representación asintótica de $\Lambda_n^P(\cdot) - \Lambda_F(\cdot)$ dada en el teorema 2.4.1, el lema 2.4.10 y la relación

$$F_n^P(t) - F(t) = (1 - F(t)) (\Lambda_n^P(t) - \Lambda_F(t)) + R_{n12}(t)$$

demuestran el siguiente teorema.

Teorema 2.4.2 (*Representación casi segura y en media de $F_n^P(\cdot) - F(\cdot)$*) Bajo las condiciones del teorema 2.4.1 se tiene, para todo $t \in [0, t_H]$ con $t_H < b_H$,

$$F_n^P(t) - F(t) = \overline{F_n^P}(t) - F(t) + R_n(t), \quad (2.31)$$

donde

$$\overline{F_n^P}(t) = F(t) + \frac{1}{n} (1 - F(t)) \sum_{i=1}^n (g_1(t, Z_i) - g_2(t, Z_i) + g_3(t, Z_i, \delta_i))$$

con

$$\sup_{0 \leq t \leq t_H} |R_n(t)| = O\left(\left(b^2 + (nb)^{-1/2} (\log n)^{1/2}\right)^2 (\log n)^2\right) \quad c.s. \quad (2.32)$$

y

$$\sup_{0 \leq t \leq t_H} \mathbb{E}[R_n^2(t)] = O\left(\left(b^4 + (nb)^{-1}\right)^2\right). \quad (2.33)$$

Este resultado es muy importante, pues nos basaremos en él para obtener, en el siguiente capítulo, una descomposición del estimador presuavizado de la función de densidad $f_n^P(\cdot)$ en términos de la función de densidad teórica $f(\cdot)$, un término $\beta_n(\cdot)$ que representa el sesgo, otro término $\sigma_n(\cdot)$ que representa la varianza, y $e_n(\cdot)$ que sería el error cometido con dicha aproximación.

La importancia de esta representación radica en que permite obtener propiedades asintóticas del estimador presuavizado de la función de densidad trabajando únicamente con una suma de variables i.i.d., y no con la complicada estructura del estimador original $f_n^P(\cdot)$.

Teorema 2.4.3 (*Consistencia casi segura de $F_n^P(\cdot) - F(\cdot)$*) Bajo las condiciones del teorema 2.4.1 se verifica que

$$\sup_{0 \leq t \leq t_H} |F_n^P(t) - F(t)| = O\left(\left(b^2 + (nb)^{-1/2} (\log n)^{1/2}\right) \log n\right) \quad c.s.$$

Demostración.

El estimador presuavizado de la función de distribución se relaciona con el estimador de la razón de fallo acumulativa de la siguiente manera:

$$F_n^P(t) - F(t) = (1 - F(t)) (\Lambda_n^P(t) - \Lambda_F(t)) + R_{n12}(t),$$

donde

$$\sup_{0 \leq t \leq t_H} |R_{n12}(t)| = O\left(\left(b^2 + (nb)^{-1/2} (\log n)^{1/2}\right)^2 (\log n)^2\right) \quad c.s.$$

Por otro lado, en el lema 2.4.10 se ha probado que

$$\sup_{0 \leq t \leq t_H} |\Lambda_n^P(t) - \Lambda_F(t)|^2 = O\left(\left(b^2 + (nb)^{-1/2} (\log n)^{1/2}\right)^2 (\log n)^2\right) \quad c.s.$$

Teniendo en cuenta que la función de supervivencia $1 - F(\cdot)$ está acotada, se obtiene directamente el resultado deseado. ■

Basándose en la representación asintótica del estimador $1 - F_n^P(\cdot)$ dada en (2.31), en el mismo trabajo de Cao, López de Ullibarri, Janssen y Veraberveke (2005) se dan las expresiones asintóticas para el sesgo y la varianza:

$$\mathbb{E}\left[\overline{F_n^P}(t) - F(t)\right] = (1 - F(t)) \alpha(t) d_K b^2 + o(b^2), \quad (2.34)$$

$$Var\left(\overline{F_n^P}(t) - F(t)\right) = n^{-1} (1 - F(t))^2 (\gamma(t) - 2q_1(t)e_K b + O(b^2)), \quad (2.35)$$

con la siguiente notación:

$$d_K = \int v^2 K(v) dv, \quad (2.36)$$

$$e_K = \int v K(v) \mathbb{K}(v) dv \quad \text{donde} \quad \mathbb{K}(v) = \int_{-1}^v K(u) du, \quad (2.37)$$

$$\alpha(t) = \int_0^t \frac{\frac{1}{2} p''(v) h(v) + p'(v) h'(v)}{1 - H(v)} dv, \quad (2.38)$$

$$\gamma(t) = \int_0^t \frac{dH^1(v)}{(1 - H(v))^2} = \int_0^t \frac{p(v) h(v)}{(1 - H(v))^2} dv = \int_0^t q_2(v) dv, \quad (2.39)$$

con

$$q_1(t) = \frac{p(t)(1 - p(t))h(t)}{(1 - H(t))^2} \quad \text{y} \quad q_2(t) = \frac{p(t)h(t)}{(1 - H(t))^2}. \quad (2.40)$$

A partir del sesgo y varianza dados en (2.34) y (2.35), los autores derivan la

expresión para el error cuadrático medio de $\overline{F}_n^P(\cdot)$:

$$MSE\left(\overline{F}_n^P(t)\right) = AMSE\left(\overline{F}_n^P(t)\right) + o(b^4) + O(n^{-1}b^2),$$

donde

$$AMSE\left(\overline{F}_n^P(t)\right) = (1 - F(t))^2 \alpha^2(t) d_K^2 b^4 + \frac{1}{n} (1 - F(t))^2 \gamma(t) - 2 \frac{b}{n} (1 - F(t))^2 q_1(t) e_K.$$

El efecto beneficioso de la presuavización en el caso de la función de supervivencia es apuntado también por Cao, López de Ullibarri, Janssen y Veraverbeke (2005), quienes muestran que la parte dominante del estimador presuavizado $\overline{F}_n^P(\cdot)$ tiene un error cuadrático medio menor que el estimador de Kaplan-Meier clásico $F_n^{KM}(\cdot)$, aunque esta eficiencia será de segundo orden.

Así, la ventana puntual asintóticamente óptima, $b_{OPT}(t)$, que minimiza $AMSE\left(\overline{F}_n^P(t)\right)$, es

$$b_{OPT}(t) = \left(\frac{q_1(t) e_K}{2\alpha^2(t) d_K^2} \right)^{1/3} n^{-1/3}.$$

De este modo, la expresión de $AMSE\left(\overline{F}_n^P(t)\right)$ para esta ventana es

$$AMSE_{b_{OPT}(t)}\left(\overline{F}_n^P(t)\right) = \frac{1}{n} (1 - F(t))^2 \gamma(t) - \frac{3}{2^{4/3}} (1 - F(t))^2 \left(\frac{q_1^2(t) e_K^2}{\alpha(t) d_K} \right)^{2/3} n^{-4/3},$$

mientras que, en el caso del estimador de Kaplan-Meier, el error cuadrático medio con la respectiva ventana puntual $AMSE$ es

$$AMSE_{b_{KM}(t)}\left(F_n^{KM}(t)\right) = \frac{1}{n} (1 - F(t))^2 \gamma(t) + O\left(n^{-3/2}\right).$$

Los términos dominantes de $AMSE_{b_{OPT}(t)}\left(\overline{F}_n^P(t)\right)$ y $AMSE_{b_{KM}(t)}\left(F_n^{KM}(t)\right)$ son iguales, pero el término de segundo orden de $AMSE_{b_{OPT}(t)}$ del estimador presuavizado es de signo negativo y de mayor orden frente al del estimador de Kaplan-Meier, de modo que $F_n^P(\cdot)$ presenta una eficiencia de segundo orden con respecto a $F_n^{KM}(\cdot)$.

2.4.3. Normalidad asintótica

A partir de las expresiones (2.34) y (2.35), y de la representación del estimador presuavizado $F_n^P(\cdot)$ dada en el teorema 2.4.2, es fácil demostrar la normalidad asintótica puntual del proceso $\sqrt{n}\left(F_n^P(\cdot) - F(\cdot)\right)$. El resultado viene dado por el siguiente teorema.

Teorema 2.4.4 (Normalidad asintótica de $F_n^P(\cdot) - F(\cdot)$) Bajo las condiciones del teorema 2.4.1, y si la ventana b verifica la condición (v2), se tiene, para cada $t < b_H$,

a) Si $nb^4 \rightarrow 0$ entonces

$$\sqrt{n} (F_n^P(t) - F(t)) \xrightarrow{d} N(0, \sigma^2(t))$$

donde

$$\sigma^2(t) = (1 - F(t))^2 \gamma(t) \quad (2.41)$$

y $\gamma(\cdot)$ es la función dada en (2.39).

b) Si $nb^4 \rightarrow C^4$, entonces

$$\sqrt{n} (F_n^P(t) - F(t)) \xrightarrow{d} N(b(t), \sigma^2(t))$$

donde

$$b(t) = C^2 (1 - F(t)) \alpha(t) d_K,$$

y $\alpha(\cdot)$ y d_K dados en (2.38) y (2.36).

Demostración.

Por el teorema 2.4.2 se tiene que, para todo $t < b_H$,

$$F_n^P(t) - F(t) = \frac{1}{n} (1 - F(t)) \sum_{i=1}^n (g_1(t, Z_i) - g_2(t, Z_i) + g_3(t, Z_i, \delta_i)) + R_n(t),$$

donde

$$\sup_{0 \leq t \leq t_H} |R_n(t)| = O\left(\left(b^2 + (nb)^{-1/2} (\log n)^{1/2}\right)^2 (\log n)^2\right) \text{ c.s.}$$

Como, bajo la hipótesis (v2), se tiene que

$$\sqrt{n} \left(b^2 + (nb)^{-1/2} (\log n)^{1/2}\right)^2 (\log n)^2 \rightarrow 0,$$

entonces la distribución asintótica de $\sqrt{n} (F_n^P(t) - F(t))$ coincide con la de

$$\begin{aligned} \sqrt{n} \left(\overline{F_n^P}(t) - F(t)\right) &= \sqrt{n} (1 - F(t)) \frac{1}{n} \sum_{i=1}^n \xi(t, Z_i, \delta_i) \\ &= \sqrt{n} (1 - F(t)) \frac{1}{n} \sum_{i=1}^n (\xi(t, Z_i, \delta_i) - \mathbb{E}[\xi(t, Z_i, \delta_i)]) \\ &\quad + \sqrt{n} (1 - F(t)) \frac{1}{n} \sum_{i=1}^n \mathbb{E}[\xi(t, Z_i, \delta_i)] \\ &= (I) + (II) \end{aligned}$$

siendo

$$\xi(t, Z, \delta) = g_1(t, Z) - g_2(t, Z) + g_3(t, Z, \delta). \quad (2.42)$$

El sesgo de $\sqrt{n}(F_n^P(t) - F(t))$ viene del término (II), para el que se tiene

$$\begin{aligned} (II) &= \sqrt{n}(1 - F(t)) \frac{1}{n} \sum_{i=1}^n \mathbb{E}[\xi(t, Z_i, \delta_i)] = \sqrt{n}(1 - F(t)) \mathbb{E}[\xi(t, Z_1, \delta_1)] \\ &= \sqrt{n}(1 - F(t)) \mathbb{E}[g_3(t, Z_1, \delta_1)] = \sqrt{n}(1 - F(t)) (d_K \alpha(t) b^2 + o(b^2)) \\ &= (1 - F(t)) d_K \alpha(t) \sqrt{nb^4} + o\left((nb^4)^{1/2}\right), \end{aligned}$$

puesto que $\mathbb{E}[g_1(t, Z)] = \mathbb{E}[g_2(t, Z)] = 0$. Bajo la hipótesis $nb^4 \rightarrow 0$ se tiene que (II) = $o(1)$, y por tanto la distribución asintótica de $\sqrt{n}(F_n^P(t) - F(t))$ coincide únicamente con la de (I) y el sesgo será despreciable. En caso de que $b = Cn^{-1/4} + o(n^{-1/4})$, entonces

$$\mathbb{E}[\sqrt{n}(F_n^P(t) - F(t))] = C^2(1 - F(t)) \alpha(t) d_K + o(1).$$

Con respecto al término (I), éste puede escribirse como una suma de variables aleatorias i.i.d. de media cero:

$$(I) = \sum_{i=1}^n \eta_{i,n}(t)$$

siendo

$$\eta_{i,n}(t) = \sqrt{n}(1 - F(t)) \frac{1}{n} (\xi(t, Z_i, \delta_i) - \mathbb{E}[\xi(t, Z_i, \delta_i)]).$$

La idea consiste en aplicar el Teorema Central del Límite para disposiciones triangulares (teorema 7.2 de Billingsley (1968), pag. 42) para tener

$$\frac{\sum_{i=1}^n \eta_{i,n}(t)}{\sum_{i=1}^n \text{Var}[\eta_{i,n}(t)]} \xrightarrow{d} N(0, 1).$$

Para ello será necesario demostrar que $\text{Var}[\eta_{i,n}(t)] = \sigma_{i,n}^2 < \infty$ para cada $i = 1, \dots, n$, que $\sigma_n^2(t) = \sum_{i=1}^n \text{Var}[\eta_{i,n}(t)]$ es positiva y finita, y además que se verifica la condición de Lindeberg.

La primera condición se cumple, puesto que la varianza de $\eta_{i,n}(t)$ es

$$\text{Var}[\eta_{i,n}(t)] = \frac{1}{n} (1 - F(t))^2 \text{Var}[\xi(t, Z, \delta)] = \frac{1}{n} (1 - F(t))^2 (\gamma(t) + O(b)).$$

Además, este resultado lleva a que $\sigma_n^2(t)$ es finita:

$$\sigma_n^2(t) = \sum_{i=1}^n \text{Var} [\eta_{i,n}(t)] = (1 - F(t))^2 \gamma(t) + O(b).$$

Por último, demostraremos que se verifica la condición de Lindeberg:

$$\frac{1}{\sigma_n^2(t)} \sum_{i=1}^n \int_{\{|\eta_{i,n}(t)| > \varepsilon \sigma_n(t)\}} \eta_{i,n}^2(t) dP \rightarrow 0 \quad \forall \varepsilon > 0. \quad (2.43)$$

Si definimos las siguientes cantidades

$$I_{i,n} = \mathbf{1}_{\{|\eta_{i,n}(t)| > \varepsilon \sigma_n(t)\}} \quad \text{y} \quad \eta_n = \sum_{i=1}^n \eta_{i,n}^2 I_{i,n},$$

entonces la condición de Lindeberg dada en (2.43) puede escribirse como

$$\frac{1}{\sigma_n^2(t)} \mathbb{E}(\eta_n) \rightarrow 0 \quad \forall \varepsilon > 0.$$

Usando que $n \rightarrow \infty$, y que la función $\xi(t, Z, \delta)$ definida en (2.42) es acotada para todo $t \in [0, t_H]$, entonces:

$$\begin{aligned} \exists n_0 &\in \mathbb{N} : \forall n \geq n_0 \Rightarrow I_{i,n}(\omega) = \mathbf{1}_{\{|\eta_{i,n}(\omega)| > \varepsilon \sigma_n(\omega)\}} = 0 \quad \forall \omega, \forall i = 1, \dots, n \\ &\Rightarrow \exists n_0 \in \mathbb{N} : \forall n \geq n_0 \Rightarrow \eta_n(\omega) = 0 \quad \forall \omega \\ &\Rightarrow \exists n_0 \in \mathbb{N} : \forall n \geq n_0 \Rightarrow \mathbb{E}(\eta_n) = 0 \end{aligned}$$

de modo que

$$\lim_{n \rightarrow \infty} \frac{1}{\sigma_n^2(t)} \mathbb{E}(\eta_n) = 0 \quad \forall \varepsilon > 0,$$

quedando así demostrada la condición de Lindeberg y, finalmente, lo establecido en el teorema. ■

Observación 2.4.2

Breslow y Cowley (1974) demostraron que el proceso $\sqrt{n}(F_n^{KM}(\cdot) - F(\cdot))$ converge puntualmente a una distribución normal centrada y con varianza $\sigma^2(\cdot)$ cuya expresión viene dada en (2.41), es decir, los procesos $\sqrt{n}(F_n^{KM}(\cdot) - F(\cdot))$ y $\sqrt{n}(F_n^P(\cdot) - F(\cdot))$ tienen la misma distribución asintótica puntual.

2.5. Estudio de simulación

En esta sección se presentan los resultados de un estudio de simulación realizado para comparar el comportamiento en la práctica del estimador de Kaplan-Meier y el estimador presuavizado de la función de distribución. Para ello hemos trabajado con 4 modelos distintos, con el fin de estudiar la posible influencia de la forma de la función $p(\cdot)$. En el modelo 1 la función $p(\cdot)$ es casi constante, en el segundo modelo es una función unimodal, mientras que en el modelo 3 es una función decreciente en todos los puntos, y finalmente para el modelo 4 se ha tomado una función $p(\cdot)$ que presenta tanto un mínimo como un máximo local.

En los modelos 1, 2 y 4 la variable observable sigue la distribución $Z \stackrel{d}{=} U[0, 1]$, mientras que en el modelo 3 se tiene que $Z \stackrel{d}{=} W(1.75, 5)$, donde $W(a, b)$ denota la distribución de Weibull con parámetro de escala a y parámetro de forma b , cuya densidad es $f_{W(a,b)}(t) = a^b b t^{b-1} \exp[-(at)^b]$.

Para favorecer la comparación, en todos los modelos la probabilidad incondicional de no censura es semejante (ver tabla 2.1).

En la tabla 2.2 se muestran las funciones de distribución teóricas a estimar, $F(\cdot)$, y las probabilidades condicionales de no censura $p(\cdot)$ para cada uno de los modelos. En la figura 2.1 se representan la función de distribución $F(\cdot)$ de la variable de interés, la función de densidad de la variable observable $h(\cdot)$, y la función de probabilidad condicional de no censura $p(\cdot)$ para cada uno de los modelos.

En todo el estudio de simulación se ha empleado el núcleo de Epanechnikov. Se han simulado $m = 500$ muestras, para cada modelo, de tamaños $n = 50, 100, 200$ y 1000 , para aproximar por Montecarlo el valor del error cuadrático medio integrado ($MISE$) de ambos estimadores de la función de distribución, el estimador presuavizado $F_n^P(\cdot)$ y el estimador clásico de Kaplan-Meier $F_n^{KM}(\cdot)$. El $MISE$ de un estimador genérico $\widehat{F}_n(\cdot)$ viene dado por

$$MISE[\widehat{F}_n(\cdot)] = \mathbb{E} \left[\int \left(\widehat{F}_n(v) - F(v) \right)^2 \omega(v) dv \right]$$

siendo $\omega(\cdot)$ una función peso convenientemente elegida para evitar problemas en la frontera de los intervalos de integración. En estas simulaciones se han tomado las funciones de ponderación $\omega(\cdot)$ como funciones indicadoras de los intervalos dados por las líneas verticales en la figura 2.1, es decir, $\omega(t) = \mathbf{1}_{\{0 \leq t \leq 0.9\}}$ para el modelo 1, $\omega(t) = \mathbf{1}_{\{0 \leq t \leq 0.95\}}$ para el modelo 2, $\omega(t) = \mathbf{1}_{\{0.2 \leq t \leq 0.75\}}$ para el modelo 3 y $\omega(t) = \mathbf{1}_{\{0 \leq t \leq 0.95\}}$ para el modelo 4.

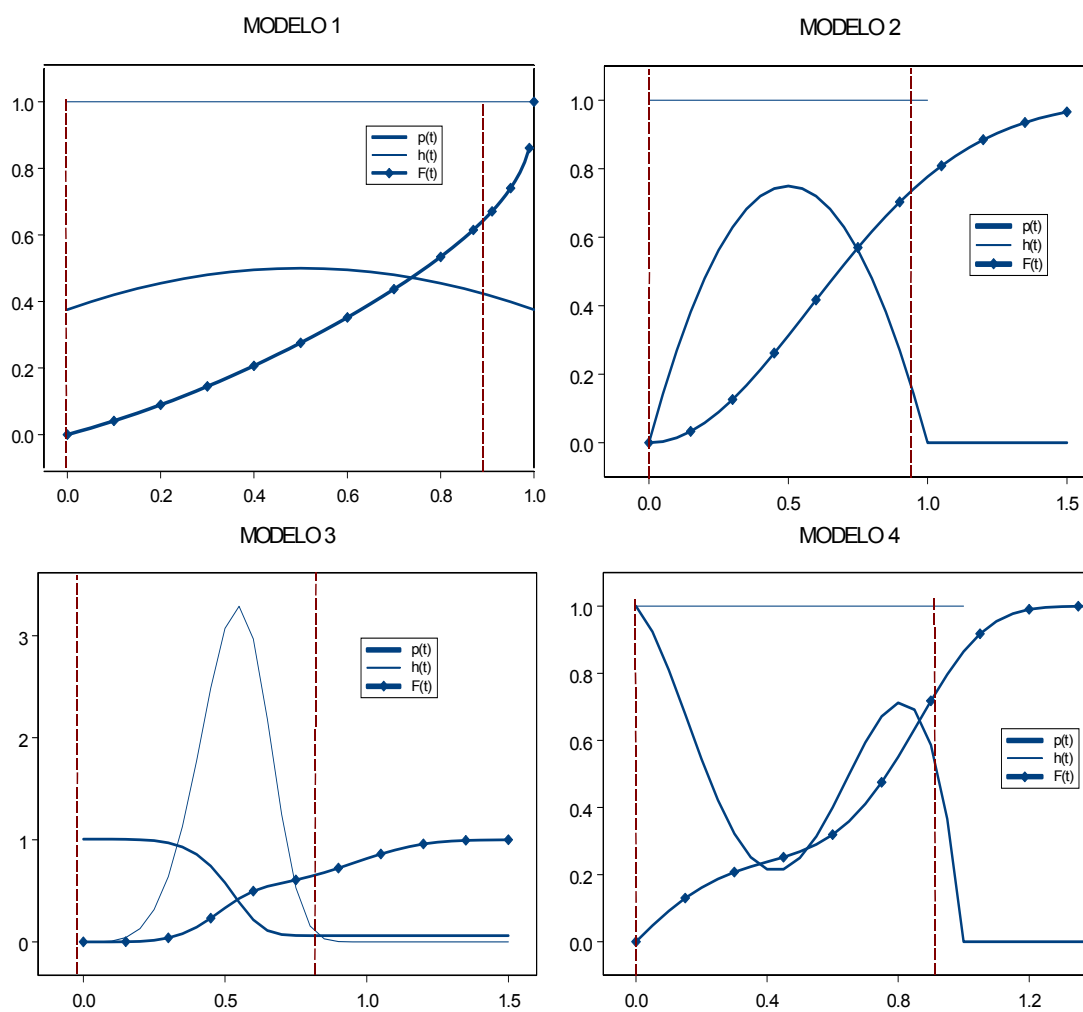


Figura 2.1. Funciones $p(\cdot)$, $h(\cdot)$ y $F(\cdot)$ para los modelos 1-4. La función peso $\omega(\cdot)$ es la función indicadora que toma el valor 1 en el intervalo marcado por las líneas verticales.

Tabla 2.1. Probabilidad incondicional de no censura en los modelos 1-4.

Modelo	1	2	3	4
γ	0.4583	0.5	0.5355	0.5

Tabla 2.2. Funciones de distribución $F(\cdot)$ y probabilidad condicional de no censura $p(\cdot)$ en los modelos 1-4 para $t \geq 0$.

Modelo	Función de distribución	Función $p(t)$
1	$F(t) = 1 - (1 - t)^{3/8} e^{-\frac{1}{4}t^2}$	$p(t) = \frac{1}{2} \left[1 - \left(t - \frac{1}{2} \right)^2 \right] \mathbf{1}_{\{0 \leq t \leq 1\}}$
2	$F(t) = 1 - e^{-\frac{3}{2}t^2}$	$p(t) = \left[\frac{3}{4} - 3 \left(t - \frac{1}{2} \right)^2 \right] \mathbf{1}_{\{0 \leq t \leq 1\}}$
3	$F(t) = 1 - \frac{1}{2} \left(e^{-t^5} + e^{-32t^5} \right)$	$p(t) = \frac{\frac{5}{2}e^{-t^5} + 80e^{-32t^5}}{41.0327 \left(e^{-t^5} + e^{-32t^5} \right)}$
4	$F(t) = 1 - \exp \left(-5t^4 + 4t^3 - t \right)$	$p(t) = \left(20t^3 - 12t^2 + 1 \right) (1 - t)$

Además, se ha calculado la eficiencia relativa del estimador de Kaplan-Meier $F_n^{KM}(\cdot)$ con respecto al estimador presuavizado $F_n^P(\cdot)$ como el cociente entre $MISE[F_n^P(\cdot)]$ y $MISE[F_n^{KM}(\cdot)]$ en función del parámetro de presuavizado b :

$$RE^{KM,P}(b) = \frac{MISE[F_n^P(\cdot)]}{MISE[F_n^{KM}(\cdot)]}.$$

Valores de $RE^{KM,P}(\cdot)$ menores que 1 indican una mayor eficiencia del estimador presuavizado $F_n^P(\cdot)$ con respecto al estimador de Kaplan-Meier $F_n^{KM}(\cdot)$ para la correspondiente ventana de presuavizado b .

Los resultados se representan en las figuras 2.2-2.5. Se puede ver que, para todos los modelos, existe un amplio rango de valores para la ventana de presuavizado b (en escala logarítmica neperiana) con los que se obtiene una estimación de $F(\cdot)$ mediante la presuavización más eficiente que mediante el estimador de Kaplan-Meier.

En las tablas 2.3-2.6 se muestran, para cada modelo, el valor mínimo que alcanza $RE^{KM,P}(\cdot)$, así como la ventana b_{MISE} en la que se alcanza dicho mínimo y la ventana b_{AMISE} que minimiza la expresión dominante de $MISE[F_n^P(\cdot)]$. En general, el valor mínimo de $RE^{KM,P}(\cdot)$ para cada uno de los modelos se acerca a uno a medida que aumenta el tamaño muestral, debido a que la eficiencia de $F_n^P(\cdot)$ con respecto a $F_n^{KM}(\cdot)$ es de segundo orden. Esto hace particularmente aconsejable la estimación presuavizada de la función de distribución cuando se dispone de muestras de tamaño relativamente reducido.

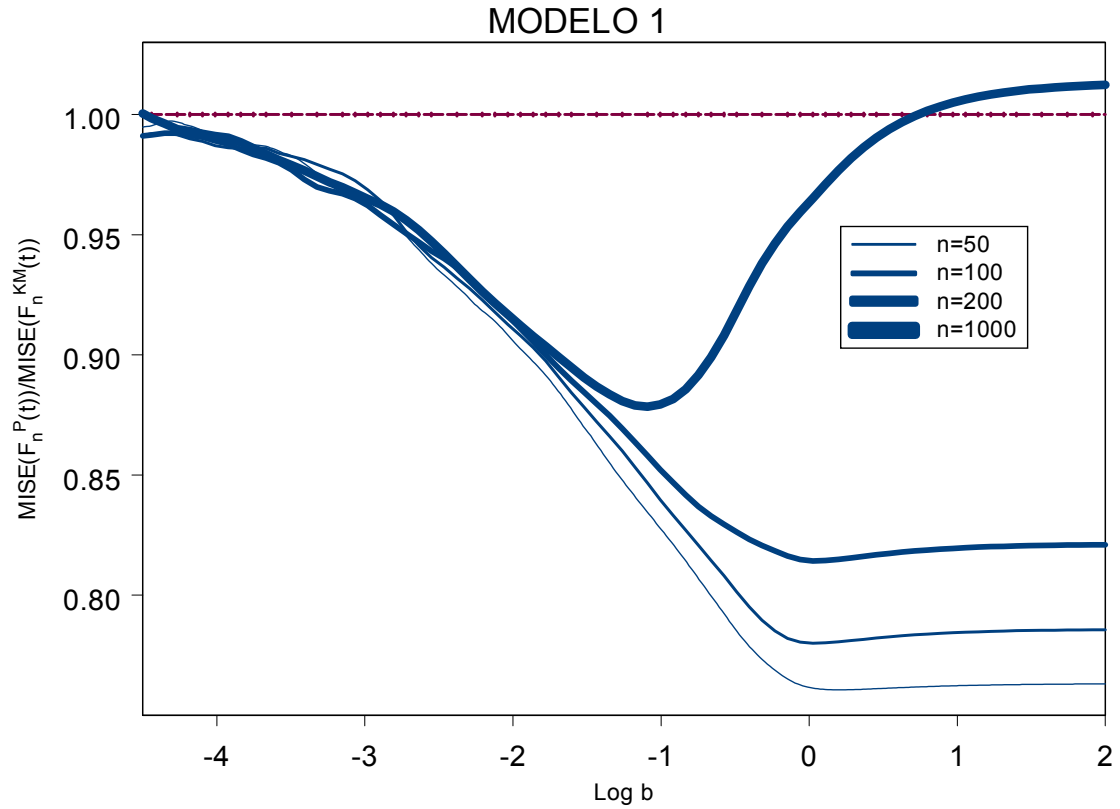


Figura 2.2. Eficiencia relativa, $RE^{KM,P}(b)$, de $F_n^{KM}(\cdot)$ respecto a $F_n^P(\cdot)$ para el modelo 1 con los tamaños muestrales $n = 50, 100, 200$ y 1000 .

Tabla 2.3. Ventanas b_{AMISE} y b_{MISE} que minimizan la expresión exacta del $AMISE(F_n^P(\cdot))$ y, por simulación, del $MISE(F_n^P(\cdot))$, y valor mínimo de la eficiencia relativa $RE^{KM,P}(b)$ de $F_n^{KM}(\cdot)$ respecto a $F_n^P(\cdot)$ para el modelo 1.

n	b_{AMISE}	b_{MISE}	mín $RE^{KM,P}(b)$ $b>0$
50	0.803	1.214	0.761
100	0.637	1.025	0.780
200	0.506	1.020	0.814
1000	0.373	0.336	0.878

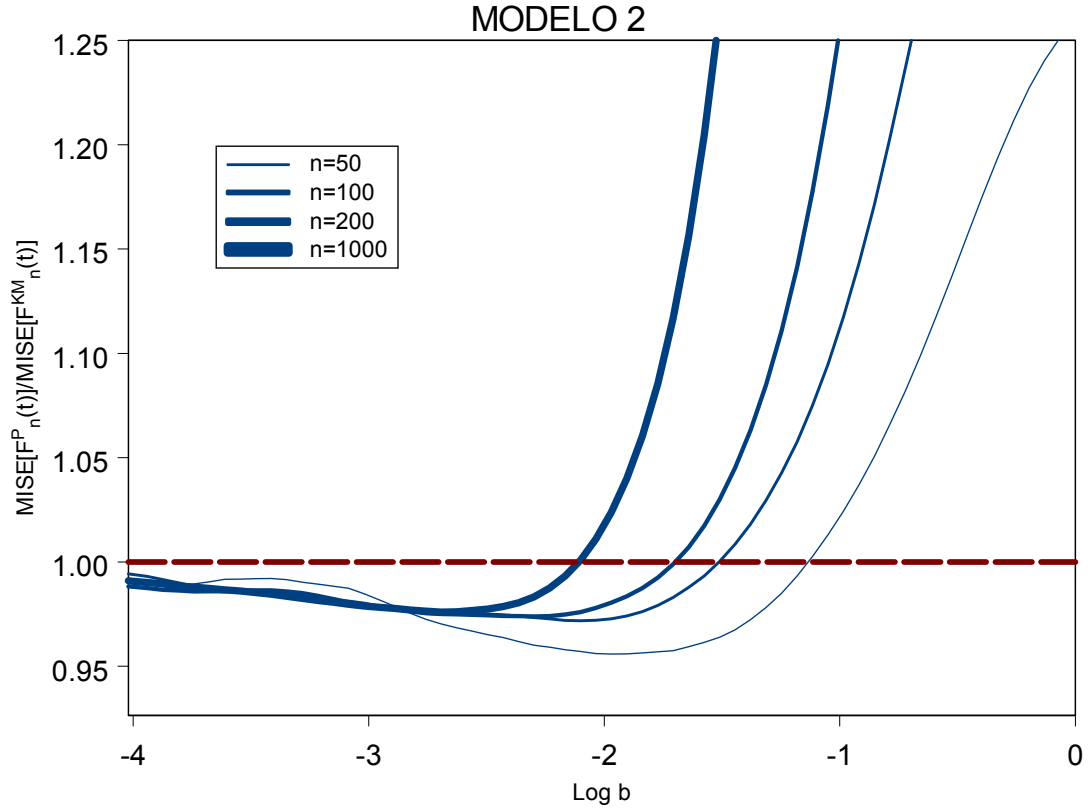


Figura 2.3. Eficiencia relativa, $RE^{KM,P}(b)$, de $F_n^{KM}(\cdot)$ respecto a $F_n^P(\cdot)$ para el modelo 2 con los tamaños muestrales $n = 50, 100, 200$ y 1000 .

Tabla 2.4. Ventanas b_{AMISE} y b_{MISE} que minimizan la expresión exacta del $AMISE(F_n^P(\cdot))$ y, por simulación, del $MISE(F_n^P(\cdot))$, y valor mínimo de la eficiencia relativa $RE^{KM,P}(b)$ de $F_n^{KM}(\cdot)$ respecto a $F_n^P(\cdot)$ para el modelo 2.

n	b_{AMISE}	b_{MISE}	$\min_{b>0} RE^{KM,P}(b)$
50	0.233	0.139	0.956
100	0.176	0.123	0.972
200	0.140	0.102	0.974
1000	0.082	0.067	0.976

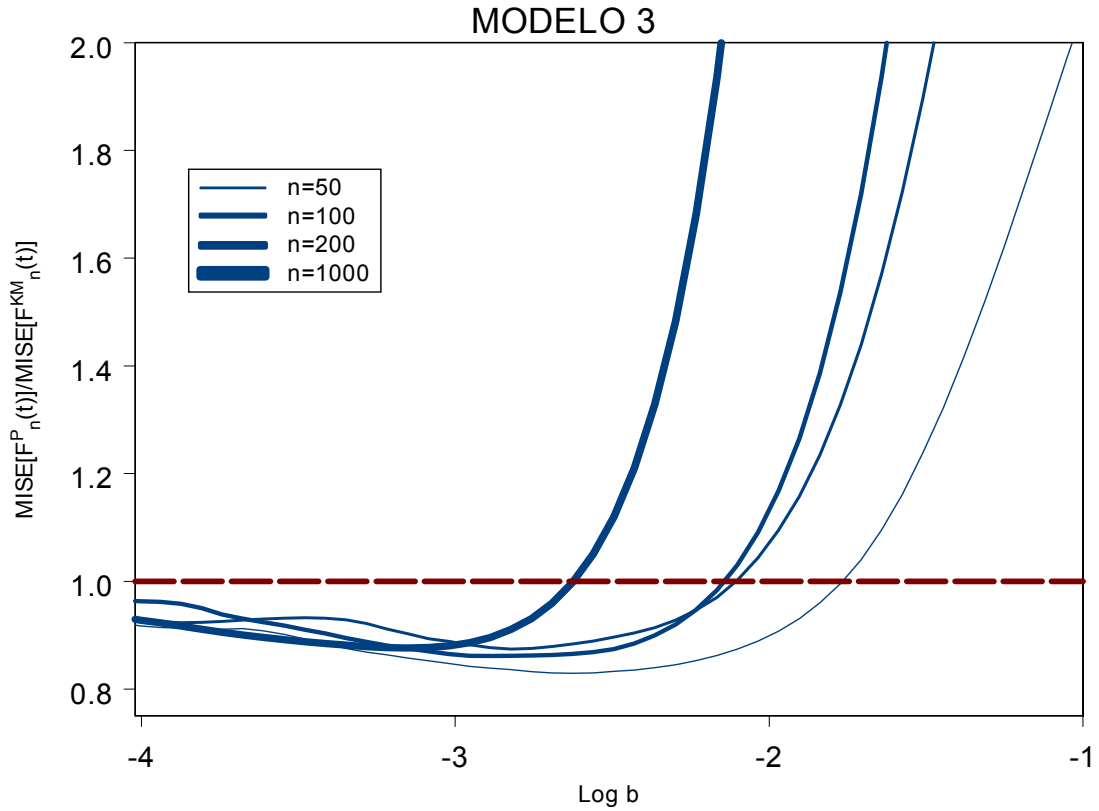


Figura 2.4. Eficiencia relativa, $RE^{KM,P}(b)$, de $F_n^{KM}(\cdot)$ respecto a $F_n^P(\cdot)$ para el modelo 3 con los tamaños muestrales $n = 50, 100, 200$ y 1000 .

Tabla 2.5. Ventanas b_{AMISE} y b_{MISE} que minimizan la expresión exacta del $AMISE(F_n^P(\cdot))$ y, por simulación, del $MISE(F_n^P(\cdot))$, y valor mínimo de la eficiencia relativa $RE^{KM,P}(b)$ de $F_n^{KM}(\cdot)$ respecto a $F_n^P(\cdot)$ para el modelo 3.

n	b_{AMISE}	b_{MISE}	$\min_{b>0} RE^{KM,P}(b)$
50	0.140	0.072	0.829
100	0.110	0.060	0.874
200	0.088	0.063	0.862
1000	0.052	0.043	0.876

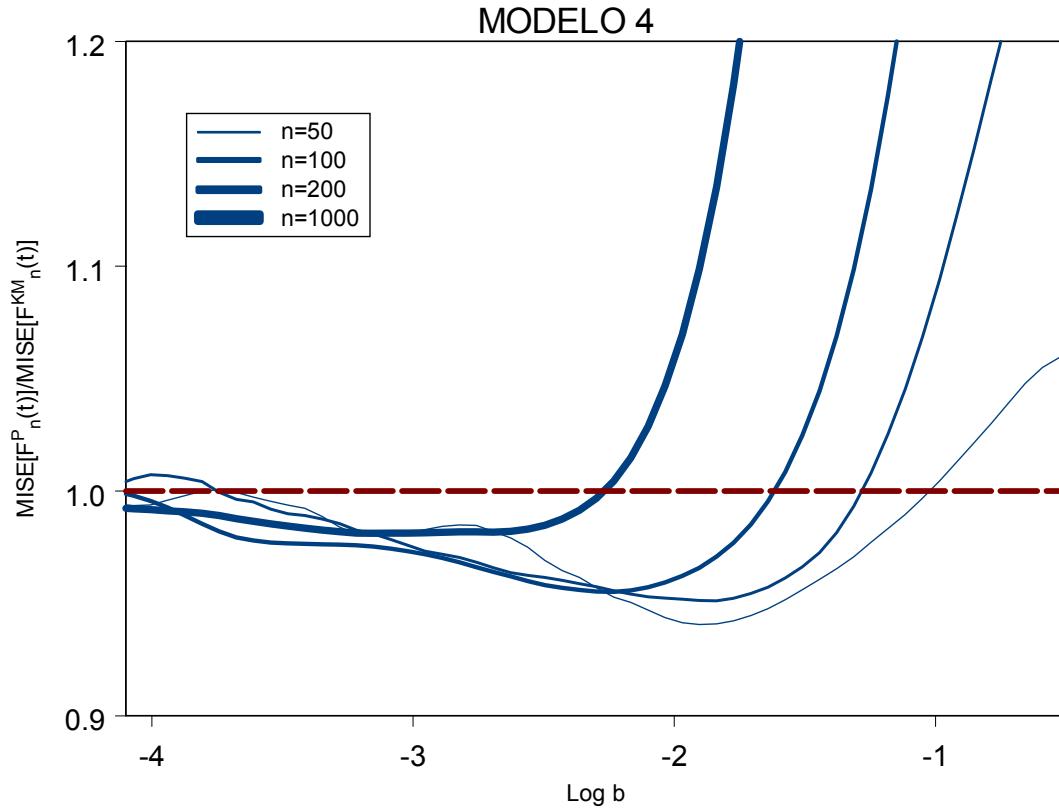


Figura 2.5. Eficiencia relativa, $RE^{KM,P}(b)$, de $F_n^{KM}(\cdot)$ respecto a $F_n^P(\cdot)$ para el modelo 4 con los tamaños muestrales $n = 50, 100, 200$ y 1000 .

Tabla 2.6. Ventanas b_{AMISE} y b_{MISE} que minimizan la expresión exacta del $AMISE(F_n^P(\cdot))$ y, por simulación, del $MISE(F_n^P(\cdot))$, y valor mínimo de la eficiencia relativa $RE^{KM,P}(b)$ de $F_n^{KM}(\cdot)$ respecto a $F_n^P(\cdot)$ para el modelo 4.

n	b_{AMISE}	b_{MISE}	$\min_{b>0} RE^{KM,P}(b)$
50	0.216	0.149	0.941
100	0.171	0.159	0.951
200	0.136	0.108	0.955
1000	0.101	0.077	0.972

Estos resultados son totalmente concordantes con los obtenidos por Cao *et al.* (2005) para el caso del estimador presuavizado de la función de fallo acumulativa, estimador que se propone como alternativa al estimador clásico de Nelson-Aalen.

El modelo 1 es, en cierto modo, especial, puesto que el estimador presuavizado $F_n^P(\cdot)$ es más eficiente que el estimador de Kaplan-Meier $F_n^{KM}(\cdot)$ prácticamente para cualquier ventana $b > 0$. Esto se debe a la forma aproximadamente plana de la función $p(\cdot)$ en el intervalo de integración $[0, 0.95]$ de modo que se puede considerar que prácticamente se verifica el modelo de Koziol-Green. Bajo dicho modelo, el estimador ACL de la función de distribución $F_n^{ACL}(\cdot)$ (ver Abdushukurov (1987) y Chen y Lin (1987)), al que tiende el estimador $F_n^P(\cdot)$ cuando el parámetro de presuavización b se hace infinitamente grande, es un estimador de la función de distribución más eficiente que el estimador de Kaplan-Meier $F_n^{KM}(\cdot)$.

Capítulo 3

Estimación presuavizada de la función de densidad

3.1. Definición del estimador. Propiedades

El estimador presuavizado de la función de densidad se obtiene sin más que sustituir en la expresión general del estimador tipo núcleo de la densidad

$$\hat{f}_h(t) = h^{-1} \int K\left(\frac{t-v}{h}\right) d\hat{F}(v) = \int K_h(t-v) d\hat{F}(v) = (K_h * \hat{F})(t),$$

el estimador genérico $\hat{F}(\cdot)$ de la función de distribución por el estimador presuavizado $F_n^P(\cdot)$, de forma que:

$$f_n^P(t) = s^{-1} \int K\left(\frac{t-v}{s}\right) dF_n^P(v) = \int K_s(t-v) dF_n^P(v) = (K_s * F_n^P)(t), \quad (3.1)$$

donde $K_s(\cdot) = s^{-1}K(\cdot/s)$ es la función núcleo reescalada y $s > 0$ es el parámetro de suavización.

Obsérvese que (3.1) tiene la estructura de los estimadores tipo núcleo, pero con pesos del estimador presuavizado de la función de distribución:

$$f_n^P(t) = \sum_{i=1}^n K_s(t - Z_i) (F_n^P(Z_i) - F_n^P(Z_i^-)) = \sum_{i=1}^n K_s(t - Z_i) W_i^P,$$

y que serán necesarios, por tanto, dos parámetros ventana, el parámetro de presuavización, b , para estimar los pesos W_i^P dados en (2.4), y el parámetro de suavización, s , para, a partir de esos pesos, estimar mediante el método núcleo la función de densidad.

Propiedad 3.1.1

Cuando se toma una ventana de presuavizado b muy pequeña ($b \rightarrow 0^+$), el estimador $f_n^P(\cdot)$ de la función de densidad coincide con el clásico estimador con pesos Kaplan-Meier $f_n^{KM}(\cdot)$.

Como se vio en el capítulo anterior, el límite del estimador $p_n(Z_i)$ cuando la ventana de presuavizado, b , es muy pequeña es precisamente δ_i , de modo que el estimador presuavizado $F_n^P(\cdot)$ se reduce al estimador de Kaplan-Meier $F_n^{KM}(\cdot)$. Es inmediato concluir entonces que

$$f_n^P(t) = \sum_{i=1}^n K_s(t - Z_i) W_i^P \xrightarrow{b \rightarrow 0^+} \sum_{i=1}^n K_s(t - Z_i) W_i^{KM} = f_n^{KM}(t).$$

Propiedad 3.1.2

En el caso de no censura, el estimador presuavizado $f_n^P(\cdot)$ se reduce al estimador clásico tipo núcleo de la densidad $\hat{f}_n(\cdot)$.

El estimador presuavizado de la función de distribución $F_n^P(\cdot)$ se reduce, en ausencia de censura, al estimador empírico $F_n(t) = \sum_{i=1}^n \mathbf{1}_{\{Z_i \leq t\}}$, de modo que

$$f_n^P(t) = \sum_{i=1}^n K_s(t - Z_i) W_i^P = \sum_{i=1}^n K_s(t - Z_i) \frac{1}{n} = \hat{f}_n(t).$$

3.2. Propiedades asintóticas del estimador

Desde el punto de vista computacional, este estimador es análogo al estimador clásico tipo núcleo, sin embargo, las propiedades estadísticas son más difíciles de obtener debido a la complicada estructura de los pesos aleatorios W_i^P . No obstante, derivaremos la representación casi segura y otras propiedades asintóticas a partir de las propiedades del estimador presuavizado de la función de distribución expuestas en el capítulo anterior.

3.2.1. Representación asintótica del estimador

Una de las propiedades más importantes del estimador presuavizado de la densidad $f_n^P(\cdot)$ es su representación asintótica como suma de variables independientes e idénticamente distribuidas, más un término de orden despreciable. Esta representación es clave en la obtención de otras propiedades del estimador, tales como su consistencia o su normalidad asintótica.

El comportamiento límite del término del resto en la representación ha sido obtenido de forma casi segura así como en media, puesto que se requerirá en el estudio de la selección de las ventanas s y b que se tratará en el siguiente capítulo.

Además de las condiciones $(K1)$, $(H1) - (H4)$, $(p1)$, $(p2)$ y $(v1)$ enunciadas en el capítulo anterior, en éste se necesitarán las siguientes hipótesis adicionales sobre la función de densidad $f(\cdot)$:

(f1) La función de densidad $f(\cdot)$ es dos veces continuamente diferenciable en el intervalo $[0, t_H]$.

(f2) La función de densidad $f(\cdot)$ es cuatro veces continuamente diferenciable en el intervalo $[0, t_H]$.

y sobre las ventanas s y b :

(v3) $nb^2s(\log n)^{-6} \rightarrow \infty$, $b^3s^{-1}(\log n)^5 \rightarrow 0$ y $nb^8s^{-1}(\log n)^4 \rightarrow 0$ cuando $n \rightarrow \infty$.

(v4) $nb^2s \rightarrow \infty$, $b^3s^{-1} \rightarrow 0$, $nb^8s^{-1} \rightarrow 0$, $nb^4s^2 \rightarrow 0$, $nb^6 \rightarrow 0$ y $b^{-1}s^2 \rightarrow 0$ cuando $n \rightarrow \infty$.

Las condiciones $(K1)$, $(H3)$, $(p2)$ y $(f2)$ son condiciones de regularidad estándar. El grado de diferenciable exigido en ellas se puede relajar para algunos resultados preliminares al cálculo del sesgo y varianza de $f_n^P(\cdot)$.

Con respecto a las condiciones sobre las ventanas, $(v1)$ es necesaria para la obtención de la representación casi segura y en media del estimador presuavizado $f_n^P(\cdot)$, puesto que se requería para la representación casi segura y en media de $F_n^P(\cdot) - F(\cdot)$, y de dicha representación se deriva la del estimador presuavizado de la densidad. La condición $(v3)$ es necesaria para la demostración de la normalidad asintótica del estimador $f_n^P(\cdot)$. Por último, se pedirá la condición $(v4)$ para la obtención de la representación asintótica del $MISE$ de $f_n^P(\cdot)$.

Estas condiciones no son mutuamente excluyentes. Por ejemplo, la condición $(v3)$ implica las condiciones $nb^2s \rightarrow \infty$, $nb^8s^{-1} \rightarrow 0$ y $b^3s^{-1} \rightarrow 0$ en $(v4)$.

El siguiente resultado proporciona la representación asintótica del estimador como una suma de variables aleatorias independientes e idénticamente distribuidas más un término de error despreciable.

Teorema 3.2.1 (*Representación casi segura y en media de $f_n^P(\cdot) - f(\cdot)$*) Bajo las hipótesis (K1), (H3), (H4), (p2) y (v1), entonces para todo $t < b_H$, se verifica que el estimador presuavizado de la función de densidad se puede expresar como

$$f_n^P(t) = f(t) + \beta_n(t) + \sigma_n(t) + e_n(t)$$

con

$$\beta_n(t) = \int f(t - sv) K(v) dv - f(t) \quad (3.2)$$

$$\sigma_n(t) = \frac{1}{ns} \sum_{i=1}^n \int \varepsilon_i(t - sv) K'(v) dv \quad (3.3)$$

$$\varepsilon_1(t) = (1 - F(t)) (g_1(t, Z_1) - g_2(t, Z_1) + g_3(t, Z_1, \delta_1)) \quad (3.4)$$

donde las funciones $g_1(\cdot, \cdot)$, $g_2(\cdot, \cdot)$ y $g_3(\cdot, \cdot, \cdot)$ vienen dadas en (2.12), (2.13) y (2.14), y el término despreciable en la representación, $e_n(\cdot)$, verifica

$$\sup_{t \in [0, t_H]} |e_n(t)| = O\left(s^{-1} \left(b^2 + (nb)^{-1/2} (\log n)^{1/2}\right)^2 (\log n)^2\right) \quad c.s.$$

y

$$\sup_{t \in [0, t_H]} \mathbb{E}[e_n^2(t)] = O\left(\left(b^4 s^{-1} + (nbs)^{-1}\right)^2\right).$$

Demostración.

Para descomponer el estimador presuavizado de la densidad $f_n^P(\cdot)$, en el paso (a) realizamos una integración por partes, en el paso (b) un cambio de variable, y por último, en el paso (c) sustituimos el estimador presuavizado $F_n^P(\cdot)$ por su representación asintótica dada en (2.31):

$$\begin{aligned} f_n^P(t) &= \frac{1}{s} \int K\left(\frac{t-v}{s}\right) dF_n^P(v) \stackrel{(a)}{=} \frac{1}{s^2} \int F_n^P(v) K'\left(\frac{t-v}{s}\right) dv \\ &\stackrel{(b)}{=} \frac{1}{s} \int F_n^P(t - sv) K'(v) dv \\ &\stackrel{(c)}{=} \frac{1}{s} \int \left(F(t - sv) + \frac{1}{n} \sum_{i=1}^n \varepsilon_i(t - sv) + R_n(t - sv) \right) K'(v) dv. \end{aligned}$$

El primer término da lugar a $f(t) + \beta_n(t)$:

$$\frac{1}{s} \int F(t - sv) K'(v) dv = \int f(t - sv) K(v) dv = f(t) + \beta_n(t)$$

mientras que el segundo sumando es $\sigma_n(t)$. Con respecto al resto, éste viene dado por

$$e_n(t) = \frac{1}{s} \int R_n(t - vs) K'(v) dv.$$

Entonces,

$$\sup_{t \in [0, t_H]} |e_n(t)| = \sup_{t \in [0, t_H]} \left| \frac{1}{s} \int R_n(t - vs) K'(v) dv \right| \leq \frac{1}{s} \sup_{t \in [0, t_H]} |R_n(t)| \int |K'(v)| dv.$$

Bajo la hipótesis (K1) y en virtud de (2.32) podemos concluir que

$$\sup_{t \in [0, t_H]} |e_n(t)| = O \left(s^{-1} \left(b^2 + (nb)^{-1/2} (\log n)^{1/2} \right)^2 (\log n)^2 \right) \text{ c.s.}$$

Respecto a la tasa en media cuadrática del resto, en el paso (a) aplicamos el teorema de Fubini, en el paso (b) usamos la desigualdad de Cauchy-Schwarz y, por último, extraemos el supremo en el paso (c),

$$\begin{aligned} \mathbb{E} [e_n^2(t)] &= \mathbb{E} \left[\left(\frac{1}{s} \int R_n(t - vs) K'(v) dv \right)^2 \right] \\ &\stackrel{(a)}{\leq} \frac{1}{s^2} \iint \mathbb{E} |R_n(t - v_1s) R_n(t - v_2s)| |K'(v_1)| |K'(v_2)| dv_1 dv_2 \\ &\stackrel{(b)}{\leq} \frac{1}{s^2} \left(\int \mathbb{E} [R_n^2(t - v_1s)]^{1/2} |K'(v_1)| dv_1 \right) \\ &\quad \times \left(\int \mathbb{E} [R_n^2(t - v_2s)]^{1/2} |K'(v_2)| dv_2 \right) \\ &\stackrel{(c)}{\leq} \frac{1}{s^2} \sup_{0 \leq t \leq t_H} \mathbb{E} [R_n^2(t)] \left(\int |K'(v)| dv \right)^2 \end{aligned}$$

de modo que, bajo la hipótesis (K1) y teniendo en cuenta (2.33), se concluye

$$\sup_{t \in [0, t_H]} \mathbb{E} [e_n^2(t)] = O \left(\left(b^4 s^{-1} + (nbs)^{-1} \right)^2 \right).$$

■

En la representación del teorema anterior, el término $\beta_n(\cdot)$ representa el sesgo, $\sigma_n(\cdot)$ representa el término de variabilidad, y $e_n(\cdot)$ es el error cometido con dicha aproximación. Cabe destacar que la parte del sesgo $\beta_n(\cdot)$ no es aleatoria, y que el término de la variabilidad $\sigma_n(\cdot)$ es una suma de variables i.i.d. Este resultado resulta muy útil, puesto que permite obtener propiedades asintóticas del estimador presuavizado de la función de densidad trabajando únicamente con una suma de variables i.i.d., y no con la complicada estructura del estimador original $f_n^P(\cdot)$.

Considérese $\overline{f_n^P}(\cdot) = f(\cdot) + \beta_n(\cdot) + \sigma_n(\cdot)$ la parte dominante del estimador presuavizado $f_n^P(\cdot)$. Calcularemos a continuación las expresiones asintóticas del sesgo y varianza de $\overline{f_n^P}(\cdot)$.

El sesgo de $\overline{f_n^P}(\cdot)$ se calcula fácilmente, ya que es igual a

$$\mathbb{E} \left[\overline{f_n^P}(t) - f(t) \right] = \beta_n(t) + \mathbb{E}[\sigma_n(t)],$$

con $\beta_n(\cdot)$ y $\sigma_n(\cdot)$ dados en (3.2) y (3.3).

Con respecto a la varianza de $\overline{f_n^P}(\cdot)$, ésta coincide con la varianza de $\sigma_n(\cdot)$ dado que las funciones $f(\cdot)$ y $\beta_n(\cdot)$ no son aleatorias. Para calcularla, reescribimos la función $\sigma_n(\cdot)$ de la siguiente manera:

$$\begin{aligned} \sigma_n(t) &= \frac{1}{ns} \sum_{i=1}^n \int \varepsilon_i(t - vs) K'(v) dv = \frac{1}{n} \sum_{i=1}^n \int K_s(t - v) d\varepsilon_i(v) \\ &= -\frac{1}{n} \sum_{i=1}^n \omega_1(t, Z_i) + \frac{1}{n} \sum_{i=1}^n \omega_2(t, Z_i) - \frac{1}{n} \sum_{i=1}^n \omega_3(t, Z_i, \delta_i) \\ &\quad + \frac{1}{n} \sum_{i=1}^n \rho_1(t, Z_i) - \frac{1}{n} \sum_{i=1}^n \rho_2(t, Z_i) + \frac{1}{n} \sum_{i=1}^n \rho_3(t, Z_i, \delta_i), \end{aligned}$$

donde

$$\omega_j(t, Z_i) = \int K_s(t - v) f(v) g_j(v, Z_i) dv, \quad j = 1, 2 \quad (3.5)$$

$$\omega_3(t, Z_i, \delta_i) = \int K_s(t - v) f(v) g_3(v, Z_i, \delta_i) dv, \quad (3.6)$$

y

$$\rho_j(t, Z_i) = \int K_s(t - v) (1 - F(v)) dg_j(v, Z_i), \quad j = 1, 2 \quad (3.7)$$

$$\rho_3(t, Z_i, \delta_i) = \int K_s(t - v) (1 - F(v)) dg_3(v, Z_i, \delta_i). \quad (3.8)$$

De esta manera, la varianza de $\sigma_n(\cdot)$ es igual a

$$\begin{aligned} Var(\sigma_n(t)) &= n^{-1} [Var(\omega_1(t, Z)) + Var(\omega_2(t, Z)) + Var(\omega_3(t, Z, \delta)) \\ &\quad + Var(\rho_1(t, Z)) + Var(\rho_2(t, Z)) + Var(\rho_3(t, Z, \delta)) \\ &\quad - 2Cov(\omega_1(t, Z), \omega_2(t, Z)) + 2Cov(\omega_1(t, Z), \omega_3(t, Z, \delta)) \\ &\quad - 2Cov(\omega_1(t, Z), \rho_1(t, Z)) + 2Cov(\omega_1(t, Z), \rho_2(t, Z)) \\ &\quad - 2Cov(\omega_1(t, Z), \rho_3(t, Z, \delta)) - 2Cov(\omega_2(t, Z), \omega_3(t, Z, \delta)) \end{aligned} \quad (3.9)$$

$$\begin{aligned}
& +2Cov(\omega_2(t, Z), \rho_1(t, Z)) - 2Cov(\omega_2(t, Z), \rho_2(t, Z)) \\
& +2Cov(\omega_2(t, Z), \rho_3(t, Z, \delta)) - 2Cov(\omega_3(t, Z, \delta), \rho_1(t, Z)) \\
& +2Cov(\omega_3(t, Z, \delta), \rho_2(t, Z)) - 2Cov(\omega_3(t, Z, \delta), \rho_3(t, Z, \delta)) \\
& -2Cov(\rho_1(t, Z), \rho_2(t, Z)) + 2Cov(\rho_1(t, Z), \rho_3(t, Z, \delta)) \\
& -2Cov(\rho_2(t, Z), \rho_3(t, Z, \delta))].
\end{aligned}$$

Puesto que $\mathbb{E}[g_1(t, Z)] = \mathbb{E}[g_2(t, Z)] = 0$ y que

$$\mathbb{E}[g_3(t, Z, \delta)] = d_K \alpha(t) b^2 + o(b^2),$$

fácilmente se puede comprobar que

$$\begin{aligned}
\mathbb{E}[\omega_1(t, Z)] &= \mathbb{E}[\omega_2(t, Z)] = \mathbb{E}[\rho_1(t, Z)] = \mathbb{E}[\rho_2(t, Z)] = 0, \\
\mathbb{E}[\omega_3(t, Z, \delta)] &= b^2 f(t) \alpha(t) d_K + o(b^2), \\
\mathbb{E}[\rho_3(t, Z, \delta)] &= b^2 (1 - F(t)) \alpha'(t) d_K + o(b^2).
\end{aligned}$$

Las varianzas y covarianzas del desarrollo de $Var(\sigma_n(\cdot))$ se calcularán en los lemas siguientes. En las demostraciones de dichos lemas se hará uso de ciertas igualdades, que se enumeran en forma de propiedad a continuación, que involucran integrales de funciones genéricas $m(\cdot)$ y $n(\cdot)$, y de las funciones $K(\cdot)$, $p(\cdot)$, $h(\cdot)$, $H(\cdot)$ y $\gamma(\cdot)$, donde $\gamma(\cdot)$ viene dada por (2.39).

Propiedad 3.2.1

$$\int_0^t \frac{p'(v)}{1-H(v)} dv = \frac{p(t)}{1-H(t)} - \gamma(t). \quad (3.10)$$

Demostración.

Inmediato, usando integración por partes y que $\gamma(\cdot)$ es la función dada en (2.39).

■

Propiedad 3.2.2

$$\int_0^t \frac{p'(v) H(v)}{1-H(v)} dv = \frac{p(t) H(t)}{1-H(t)} - \gamma(t). \quad (3.11)$$

Demostración.

Inmediato al usar de nuevo integración por partes. ■

Propiedad 3.2.3

$$\int_0^t h(v) \left(\frac{p(v)}{1-H(v)} - \gamma(v) \right) dv = \gamma(t) (1-H(t)). \quad (3.12)$$

Demostración.

En el paso (a) hacemos un cambio de variable, en el paso (b) utilizamos (3.10), en el paso (c) se usa (3.11), y por último, en el paso (d) simplificamos la expresión final.

$$\begin{aligned}
& \int_0^t h(v) \left(\frac{p(v)}{1-H(v)} - \gamma(v) \right) dv \\
& \stackrel{(a)}{=} \left[\left(\frac{p(v)}{1-H(v)} - \gamma(v) \right) H(v) \right]_0^t - \int_0^t H(v) \left(\frac{p(v)}{1-H(v)} - \gamma(v) \right)' dv \\
& \stackrel{(b)}{=} \frac{p(t)H(t)}{1-H(t)} - \gamma(t)H(t) - \int_0^t \frac{p'(v)H(v)}{1-H(v)} dv \\
& \stackrel{(c)}{=} \frac{p(t)H(t)}{1-H(t)} - \gamma(t)H(t) - \left(\frac{p(t)H(t)}{1-H(t)} - \gamma(t) \right) \stackrel{(d)}{=} \gamma(t)(1-H(t)).
\end{aligned}$$

■

Propiedad 3.2.4

$$\int_0^t h(v) \left(\frac{p(v)}{1-H(v)} - \gamma(v) \right)^2 dv = \int_0^t \frac{p^2(v)h(v)}{(1-H(v))^2} dv - \gamma^2(t)(1-H(t)). \quad (3.13)$$

Demostración.

En el paso (a) hacemos un cambio de variable y utilizamos (3.12), en el paso (b) usamos (3.10), en el paso (c) hacemos integración por partes, y finalmente en el paso (d) sustituimos la derivada de $\gamma(\cdot)$ por su expresión en función de $p(\cdot)$, $h(\cdot)$ y $H(\cdot)$:

$$\begin{aligned}
& \int_0^t h(v) \left(\frac{p(v)}{1-H(v)} - \gamma(v) \right)^2 dv \\
& = \int_0^t h(v) \left(\frac{p(v)}{1-H(v)} - \gamma(v) \right) \left(\frac{p(v)}{1-H(v)} - \gamma(v) \right) dv \\
& \stackrel{(a)}{=} \left[\left(\frac{p(v)}{1-H(v)} - \gamma(v) \right) \gamma(v) (1-H(v)) \right]_0^t \\
& \quad - \int_0^t \left(\frac{p(v)}{1-H(v)} - \gamma(v) \right)' \gamma(v) (1-H(v)) dv \\
& \stackrel{(b)}{=} \left(\frac{p(t)}{1-H(t)} - \gamma(t) \right) \gamma(t) (1-H(t)) - \int_0^t p'(v) \gamma(v) dv \\
& \stackrel{(c)}{=} p(t) \gamma(t) - \gamma^2(t) (1-H(t)) - \left(\gamma(t) p(t) - \int_0^t p(v) \gamma'(v) dv \right) \\
& \stackrel{(d)}{=} \int_0^t \frac{p^2(v)h(v)}{(1-H(v))^2} dv - \gamma^2(t) (1-H(t)).
\end{aligned}$$

■

Con respecto al núcleo $K(\cdot)$, necesitaremos los resultados que se exponen a continuación. En ellos consideramos dos funciones genéricas $m(\cdot)$ y $n(\cdot)$, junto a la función de distribución $H(\cdot)$.

Propiedad 3.2.5

$$\int K(u_1) \int_{-1}^{u_1} K(u_2) du_2 du_1 = \frac{1}{2}.$$

Demostración.

Inmediata, usando integración por partes. ■

Propiedad 3.2.6

$$\int K(u_1) \int_{-1}^{u_1} u_2 K(u_2) du_2 du_1 = e_K.$$

Demostración.

Se realiza una integración por partes, teniendo en cuenta que e_K es la constante dada en (2.37) en el capítulo anterior. ■

Propiedad 3.2.7 *Si el núcleo $K(\cdot)$ es una función de densidad simétrica y $m(\cdot)$ es dos veces continuamente diferenciable, entonces*

$$\int K_s(t-u) m(u) du = m(t) + O(s^2). \quad (3.14)$$

Demostración.

Se realiza en primer lugar un cambio de variable, y a continuación un desarrollo de Taylor de orden dos con resto en forma integral en la función $m(\cdot)$:

$$\begin{aligned} \int K_s(t-u) m(u) du &= \int K(v) m(t-vs) dv \\ &= \int K(v) \left[m(t) - vsm'(t) + v^2s^2 \int_0^1 (1-x) m''(t-vsx) dx \right] dv. \end{aligned}$$

Puesto que $K(\cdot)$ es una función de densidad simétrica, y el valor absoluto del tercer sumando se puede acotar por $\frac{1}{2}s^2 d_K \|m''(\cdot)\|$, donde d_K viene dada por (2.7), y siendo $\|g(\cdot)\| = \sup_x |g(x)|$ de una función arbitraria $g(\cdot)$, se obtiene (3.14). ■

Propiedad 3.2.8 *Si el núcleo $K(\cdot)$ es una función de densidad simétrica y $m(\cdot)$ es cuatro veces continuamente diferenciable, entonces*

$$\int K_s^2(t-u) m(u) du = \frac{1}{s} m(t) c_K + \frac{1}{2} sm''(t) m_K + O(s^3). \quad (3.15)$$

donde $m_K = \int v^2 K^2(v) dv$, y c_K dado en (2.7) en el capítulo anterior.

Demostración.

Se realiza en primer lugar un cambio de variable, y a continuación un desarrollo de Taylor de orden cuatro con resto en forma integral en la función $m(\cdot)$. Puesto que $K(\cdot)$ es una función de densidad simétrica, entonces

$$\begin{aligned} \int K_s^2(t-u) m(u) du &= \frac{1}{s} \int K^2(v) m(t-vs) dv \\ &= \frac{1}{s} m(t) c_K + \frac{1}{2} s m''(t) m_K + \frac{1}{6} s^3 \int v^4 K^2(v) \int_0^1 (1-x)^3 m^{(4)}(t-vsx) dx dv. \end{aligned}$$

De la acotación del valor absoluto del último término por $\frac{1}{24} s^3 \int v^4 K^2(v) dv \|m^{(4)}(\cdot)\|$ se obtiene (3.15). ■

Propiedad 3.2.9 *Sea el núcleo $K(\cdot)$ una función de densidad simétrica y las funciones $m(\cdot)$ y $n(\cdot)$ dos veces continuamente diferenciables. Entonces*

$$\begin{aligned} \iint K_s(t-u) K_s(t-v) m(u) n(v) \mathbf{1}_{\{u \leq v\}} du dv \\ = \frac{1}{2} m(t) n(t) - s(m'(t) n(t) - m(t) n'(t)) e_K + O(s^2), \end{aligned} \quad (3.16)$$

donde e_K viene dada en (2.37).

Demostración.

En primer lugar, en el paso (a), realizamos un cambio de variable; a continuación, paso (b), llevamos a cabo un desarrollo de Taylor en las funciones $m(\cdot)$ y $n(\cdot)$, donde ζ_{t,s,u_1} y ξ_{t,s,v_1} son dos puntos intermedios entre t y $t-su_1$, y t y $t-sv_1$ respectivamente, en el paso (c) desarrollamos el producto, y finalmente en el paso (d) utilizamos la definición de e_K dada en (2.37).

$$\begin{aligned} \iint K_s(t-u) K_s(t-v) m(u) n(v) \mathbf{1}_{\{u \leq v\}} du dv \\ \stackrel{(a)}{=} \iint K(u_1) K(v_1) m(t-su_1) n(t-sv_1) \mathbf{1}_{\{u_1 \geq v_1\}} du_1 dv_1 \\ \stackrel{(b)}{=} \iint K(u_1) K(v_1) \left(m(t) - su_1 m'(t) + \frac{1}{2} s^2 u_1^2 m''(\zeta_{t,s,u_1}) \right) \\ \times \left(n(t) - sv_1 n'(t) + \frac{1}{2} s^2 v_1^2 n''(\xi_{t,s,v_1}) \right) \mathbf{1}_{\{u_1 \geq v_1\}} du_1 dv_1 \\ \stackrel{(c)}{=} m(t) n(t) \iint K(u_1) K(v_1) \mathbf{1}_{\{u_1 \geq v_1\}} du_1 dv_1 \end{aligned}$$

$$\begin{aligned}
& -sm'(t)n(t) \iint u_1 K(u_1) K(v_1) \mathbf{1}_{\{u_1 \geq v_1\}} du_1 dv_1 \\
& -sm(t)n'(t) \iint v_1 K(u_1) K(v_1) \mathbf{1}_{\{u_1 \geq v_1\}} du_1 dv_1 + O(s^2) \\
& \stackrel{(d)}{=} \frac{1}{2}m(t)n(t) - s(m'(t)n(t) - m(t)n'(t))e_K + O(s^2)
\end{aligned}$$

■

Propiedad 3.2.10 *Sea el núcleo $K(\cdot)$ una función de densidad simétrica, y sean las funciones $m(\cdot)$ y $n(\cdot)$ dos veces continuamente diferenciables. Entonces*

$$\begin{aligned}
& \iint K_s(t-v_1) K_s(t-v_2) m(v_1) n(v_2) H(v_1 \wedge v_2) dv_1 dv_2 \\
& = m(t)n(t)H(t) - 2sm(t)n(t)h(t)e_K + O(s^2).
\end{aligned} \tag{3.17}$$

Demostración.

Este resultado se obtiene sin más que aplicar desarrollos de Taylor de orden dos a las funciones $m(\cdot)$ y $n(\cdot)$, junto con (3.10), (3.11) y la definición de e_K dada en (2.37):

$$\begin{aligned}
& \iint K_s(t-v_1) K_s(t-v_2) m(v_1) n(v_2) H(v_1 \wedge v_2) dv_1 dv_2 \\
& = \iint K(u_1) K(u_2) m(t-u_1s) n(t-u_2s) H(t-s(u_1 \vee u_2)) du_1 du_2 \\
& = \int K(u_1) m(t-u_1s) H(t-u_1s) \int_{-1}^{u_1} K(u_2) n(t-u_2s) du_2 du_1 \\
& \quad + \int K(u_2) n(t-u_2s) H(t-u_2s) \int_{-1}^{u_2} K(u_1) m(t-u_1s) du_1 du_2 \\
& = m(t)H(t)n(t) \int K(u) \mathbb{K}(u) du - s[m(t)H(t)]'n(t) \int uK(u) \mathbb{K}(u) du \\
& \quad - sm(t)H(t)n'(t) \int K(u) \int_{-1}^u vK(v) dv du \\
& \quad + n(t)H(t)m(t) \int K(u) \mathbb{K}(u) du - s[n(t)H(t)]'m(t) \int uK(u) \mathbb{K}(u) du \\
& \quad - sn(t)H(t)m'(t) \int K(u) \int_{-1}^u vK(v) dv du + O(s^2) \\
& = m(t)n(t)H(t) - 2sm(t)n(t)h(t)e_K + O(s^2).
\end{aligned}$$

■

Una vez establecidas estas propiedades, volvemos a la tarea de calcular la varianza de $\sigma_n(\cdot)$ mediante el cálculo de todas las varianzas y covarianzas en las que se descompone $Var(\sigma_n(\cdot))$ dadas en el desarrollo (3.9). De todas ellas, empezamos calculando las que no involucran la ventana de presuavizado b , es decir, las que no incluyen a $\omega_3(\cdot, Z, \delta)$ ni a $\rho_3(\cdot, Z, \delta)$.

Lema 3.2.1 *Bajo las condiciones (K1), (H2), (p1) y (f1) se tiene*

$$Var(\omega_1(t, Z)) = f^2(t) \frac{p^2(t) H(t)}{1 - H(t)} - 2sf^2(t) \frac{p^2(t) h(t)}{(1 - H(t))^2} e_K + O(s^2).$$

Demostración.

Considerando (3.5) y (2.12), la función $\omega_1(\cdot, \cdot)$ es igual a

$$\omega_1(t, Z) = \int K_s(t - v) \frac{f(v)p(v)}{1 - H(v)} (\mathbf{1}_{\{Z \leq v\}} - H(v)) dv. \quad (3.18)$$

Dado que la esperanza de $\omega_1(\cdot, Z)$ es cero, la varianza de $\omega_1(\cdot, Z)$ es igual a

$$\begin{aligned} Var(\omega_1(t, Z)) &= \mathbb{E}[\omega_1^2(t, Z)] \\ &= \mathbb{E} \left(\iint K_s(t - v_1) K_s(t - v_2) \frac{f(v_1)p(v_1)}{1 - H(v_1)} \frac{f(v_2)p(v_2)}{1 - H(v_2)} \right. \\ &\quad \times (\mathbf{1}_{\{Z \leq v_1\}} - H(v_1)) (\mathbf{1}_{\{Z \leq v_2\}} - H(v_2)) dv_1 dv_2 \Big) \\ &= \iint K_s(t - v_1) K_s(t - v_2) \frac{f(v_1)p(v_1)}{1 - H(v_1)} \frac{f(v_2)p(v_2)}{1 - H(v_2)} \\ &\quad \times (H(v_1 \wedge v_2) - H(v_1)H(v_2)) dv_1 dv_2. \end{aligned}$$

Finalmente, aplicamos los resultados (3.14) y (3.17) tomando las funciones $m(t) = n(t) = f(t)p(t)(1 - H(t))^{-1}$ y bajo (H2), (p1) y (f1):

$$\begin{aligned} Var(\omega_1(t, Z)) &= \iint K_s(t - v_1) K_s(t - v_2) m(v_1) m(v_2) H(v_1 \wedge v_2) dv_1 dv_2 \\ &\quad - \left[\int K_s(t - v) m(v) H(v) dv \right]^2 \\ &= m^2(t) H(t) - m^2(t) H^2(t) - 2sm^2(t) h(t) e_K + O(s^2) \\ &= f^2(t) \frac{p^2(t) H(t)}{1 - H(t)} - 2sf^2(t) \frac{p^2(t) h(t)}{(1 - H(t))^2} e_K + O(s^2). \end{aligned}$$

■

Lema 3.2.2 *Bajo las condiciones (K1), (H2), (p1) y (f1) se tiene*

$$\begin{aligned}
& \text{Var}(\rho_1(t, Z)) \\
&= \frac{1}{s} \left(\frac{1-F(t)}{1-H(t)} \right)^2 p^2(t) h(t) c_K + \frac{(1-F(t))^2}{1-H(t)} \left(p'(t) H(t) + \frac{p(t) h(t)}{1-H(t)} \right) p'(t) \\
&+ \frac{1}{2} s \left[\left(\frac{1-F(t)}{1-H(t)} \right)^2 p^2(t) h(t) \right]'' m_K - 2s(1-F(t))^2 \left[\frac{h(t) p'(t)^2}{(1-H(t))^2} \right. \\
&\left. + p'(t) \left(\frac{h(t) p(t)}{(1-H(t))^2} \right)' - \frac{p(t) h(t)}{(1-H(t))^2} p''(t) \right] e_K + O(s^2).
\end{aligned}$$

Demostración.

Considerando (3.7) y (2.12), la función $\rho_1(\cdot, \cdot)$ es igual a

$$\begin{aligned}
\rho_1(t, Z) &= \int K_s(t-v) m(v) \mathbf{1}_{\{Z \leq v\}} dv + K_s(t-Z) \frac{1-F(Z)}{1-H(Z)} p(Z) \quad (3.19) \\
&\quad - \int K_s(t-v) n(v) dv,
\end{aligned}$$

donde las funciones $m(\cdot)$ y $n(\cdot)$ son

$$m(t) = (1-F(t)) \left(\frac{p(t)}{1-H(t)} \right)' \quad \text{y} \quad n(t) = (1-F(t)) \left(\frac{p(t) H(t)}{1-H(t)} \right)'.$$

Dado que, de nuevo, la esperanza de $\rho_1(\cdot, Z)$ es cero, la varianza de $\rho_1(\cdot, Z)$ es igual a

$$\begin{aligned}
& \text{Var}(\rho_1(t, Z)) = \mathbb{E}[\rho_1^2(t, Z)] \\
&= \mathbb{E} \left[\iint K_s(t-v_1) K_s(t-v_2) m(v_1) m(v_2) \mathbf{1}_{\{Z \leq v_1 \wedge v_2\}} dv_1 dv_2 \right. \\
&\quad + K_s^2(t-Z) \left(\frac{1-F(Z)}{1-H(Z)} \right)^2 p^2(Z) + \left(\int K_s(t-v) n(v) dv \right)^2 \\
&\quad + 2K_s(t-Z) \frac{1-F(Z)}{1-H(Z)} p(Z) \int K_s(t-v) m(v) \mathbf{1}_{\{Z \leq v\}} dv \\
&\quad - 2 \int K_s(t-v) m(v) \mathbf{1}_{\{Z \leq v\}} dv \times \int K_s(t-v) n(v) dv \\
&\quad \left. - 2K_s(t-Z) \frac{1-F(Z)}{1-H(Z)} p(Z) \int K_s(t-v) n(v) dv \right] \\
&= \mathbb{E}[I_1 + I_2 + I_3 + I_4 - I_5 - I_6]
\end{aligned}$$

Calculamos ahora la esperanza de cada uno de los seis sumandos en los que se descompone la varianza $Var(\rho_1(t, Z))$. La primera de las esperanzas es, aplicando (3.17) y en virtud de (H2), (p1) y (f1),

$$\begin{aligned}\mathbb{E}[I_1] &= \iiint K_s(t-v_1) K_s(t-v_2) m(v_1) m(v_2) H(v_1 \wedge v_2) dv_1 dv_2 \\ &= m^2(t) H(t) - 2sm^2(t) h(t) e_K + O(s^2),\end{aligned}$$

mientras que, aplicando (3.15), la segunda esperanza es igual a

$$\begin{aligned}\mathbb{E}[I_2] &= \int K_s^2(t-u) \left(\frac{1-F(u)}{1-H(u)} \right)^2 p^2(u) h(u) du \\ &= \frac{1}{s} \left(\frac{1-F(t)}{1-H(t)} \right)^2 p^2(t) h(t) c_K + \frac{1}{2}s \left[\left(\frac{1-F(t)}{1-H(t)} \right)^2 p^2(t) h(t) \right]'' m_K + O(s^3).\end{aligned}$$

La tercera esperanza es

$$\mathbb{E}[I_3] = \left(\int K_s(t-v) n(v) dv \right)^2 = n^2(t) + O(s^2),$$

sin más que usar (3.14). Para calcular la cuarta esperanza aplicamos (3.16):

$$\begin{aligned}\mathbb{E}[I_4] &= 2 \int K_s(t-u) \frac{1-F(u)}{1-H(u)} p(u) h(u) \int K_s(t-v) m(v) \mathbf{1}_{\{u \leq v\}} dv du \\ &= \frac{1-F(t)}{1-H(t)} p(t) h(t) m(t) \\ &\quad - 2s \left[\left(\frac{1-F(t)}{1-H(t)} p(t) h(t) \right)' m(t) - \frac{1-F(t)}{1-H(t)} p(t) h(t) m'(t) \right] e_K + O(s^2).\end{aligned}$$

La quinta esperanza es igual a

$$\mathbb{E}[I_5] = 2 \int K_s(t-v) m(v) H(v) dv \times \int K_s(t-v) n(v) dv = 2m(t) H(t) n(t) + O(s^2),$$

mientras que la última es igual a

$$\begin{aligned}\mathbb{E}[I_6] &= 2 \int K_s(t-u) \frac{1-F(u)}{1-H(u)} p(u) h(u) \int K_s(t-v) n(v) dv du \\ &= 2n(t) \frac{1-F(t)}{1-H(t)} p(t) h(t) + O(s^2),\end{aligned}$$

por la aplicación de (3.14). La suma de todas estas esperanzas demuestra el lema. ■

Lema 3.2.3 *Bajo las condiciones (K1), (H2), (p1) y (f1) se tiene*

$$\text{Var}(\omega_2(t, Z)) = 2f^2(t) \int_0^t (p(t) - p(v)) \frac{H(v)p'(v)}{1-H(v)} dv + O(s^2).$$

Demostración.

De las expresiones (3.5) y (2.13) se deduce que

$$\begin{aligned} \omega_2(t, Z) &= \int K_s(t-v_1) f(v_1) \int_0^{v_1} \frac{p'(v_2)}{1-H(v_2)} \mathbf{1}_{\{Z \leq v_2\}} dv_2 dv_1 \\ &\quad - \int K_s(t-v) f(v) \left(\frac{p(v)H(v)}{1-H(v)} - \gamma(v) \right) dv, \end{aligned} \quad (3.20)$$

Puesto que $\mathbb{E}[g_2(\cdot, Z)] = 0$, de nuevo la esperanza de $\omega_2(\cdot, Z)$ es cero, de modo que su varianza es igual a

$$\begin{aligned} \text{Var}(\omega_2(t, Z)) &= \mathbb{E}[\omega_2^2(t, Z)] \\ &= \mathbb{E} \left[\iint K_s(t-v_1) K_s(t-w_1) f(v_1) f(w_1) \right. \\ &\quad \times \left. \int_0^{v_1} \frac{p'(v_2)}{1-H(v_2)} \int_0^{w_1} \frac{p'(w_2)}{1-H(w_2)} \mathbf{1}_{\{Z \leq v_2\}} \mathbf{1}_{\{Z \leq w_2\}} dv_2 dw_2 \right. \\ &\quad \left. - 2 \int K_s(t-v) f(v) \left(\frac{p(v)H(v)}{1-H(v)} - \gamma(v) \right) dv \right. \\ &\quad \times \mathbb{E} \left(\int K_s(t-v_1) f(v_1) \int_0^{v_1} \frac{p'(v_2)}{1-H(v_2)} \mathbf{1}_{\{Z \leq v_2\}} dv_2 dv_1 \right) \\ &\quad \left. + \left[\int K_s(t-v) f(v) \left(\frac{p(v)H(v)}{1-H(v)} - \gamma(v) \right) dv \right]^2 \right] \\ &= \mathbb{E}[I_1] - 2\mathbb{E}[I_2] + I_3. \end{aligned}$$

La primera de las esperanzas en las que se descompone esta varianza, aplicando (3.10) en el paso (a), es igual a

$$\begin{aligned} \mathbb{E}[I_1] &= \mathbb{E} \left[\iint K_s(t-v_1) K_s(t-w_1) f(v_1) f(w_1) \int_Z^{v_1} \frac{p'(v_2)}{1-H(v_2)} \mathbf{1}_{\{Z \leq v_1\}} dv_2 \right. \\ &\quad \times \left. \left(\int_Z^{w_1} \frac{p'(w_2)}{1-H(w_2)} \mathbf{1}_{\{Z \leq w_1\}} dw_2 \right) dv_1 dw_1 \right] \\ &\stackrel{(a)}{=} \mathbb{E} \left[\iint K_s(t-v_1) K_s(t-w_1) f(v_1) f(w_1) \right. \end{aligned}$$

$$\begin{aligned} & \times \left(\frac{p(v_1)}{1-H(v_1)} - \gamma(v_1) - \frac{p(Z)}{1-H(Z)} + \gamma(Z) \right) \\ & \times \left(\frac{p(w_1)}{1-H(w_1)} - \gamma(w_1) - \frac{p(Z)}{1-H(Z)} + \gamma(Z) \right) \mathbf{1}_{\{Z \leq v_1 \wedge w_1\}} dv_1 dw_1 \Big] \end{aligned}$$

Con el fin de aplicar la expresión (3.17) definimos la función $m(\cdot)$ como:

$$m(t) = \frac{p(t)}{1-H(t)} - \gamma(t).$$

Entonces la esperanza anterior es igual a

$$\begin{aligned} \mathbb{E}[I_1] &= \iint K_s(t-v) K_s(t-w) f(v) m(v) f(w) m(w) H(v \wedge w) dv dw \\ &\quad - 2 \iiint K_s(t-v) K_s(t-w) f(v) f(w) m(w) m(u) h(u) \mathbf{1}_{\{u \leq v \wedge w\}} du dv dw \\ &\quad + \iiint K_s(t-v) K_s(t-w) f(v) f(w) h(u) m^2(u) \mathbf{1}_{\{u \leq v \wedge w\}} du dv dw. \end{aligned}$$

Para aplicar las propiedades (3.12) y (3.13), definimos ahora las funciones $m_1(\cdot)$ y $m_2(\cdot)$ de la siguiente manera:

$$\int_0^t h(u) m(u) du = \gamma(t) (1-H(t)) = m_1(t)$$

y

$$\int_0^t h(u) m^2(u) du = \int_0^t \frac{p^2(u) h(u)}{(1-H(u))^2} du - \gamma^2(t) (1-H(t)) = m_2(t).$$

Entonces, aplicando (3.17) en el paso (a), y sustituyendo las funciones $m(\cdot)$, $m_1(\cdot)$ y $m_2(\cdot)$ por sus expresiones (paso (b)), podemos simplificar la expresión la primera de las integrales en las que se descompone $Var(\omega_2(\cdot, Z))$ de la siguiente manera:

$$\begin{aligned} \mathbb{E}[I_1] &= \iint K_s(t-v) K_s(t-w) f(v) m(v) f(w) m(w) H(v \wedge w) dv dw \\ &\quad - 2 \iint K_s(t-v) K_s(t-w) f(v) f(w) m(w) m_1(v \wedge w) dv dw \\ &\quad + \iint K_s(t-v) K_s(t-w) f(v) f(w) m_2(v \wedge w) dv dw \end{aligned}$$

$$\begin{aligned}
& \stackrel{(a)}{=} f^2(t) m^2(t) H(t) - 2s f^2(t) m^2(t) h(t) e_K \\
& - 2f^2(t) m(t) m_1(t) + 4s f^2(t) m(t) m_1'(t) e_K \\
& + f^2(t) m_2(t) - 2s f^2(t) m_2'(t) e_K + O(s^2) \\
& \stackrel{(b)}{=} f^2(t) \left[\left(\frac{p(t)}{1-H(t)} - \gamma(t) \right)^2 H(t) - 2\gamma(t) (1-H(t)) \left(\frac{p(t)}{1-H(t)} - \gamma(t) \right) \right. \\
& \left. + \int_0^t \frac{p^2(u) h(u)}{(1-H(u))^2} du - \gamma^2(t) (1-H(t)) \right] + O(s^2) \\
& = f^2(t) \left[\left(\frac{p(t)}{1-H(t)} \right)^2 H(t) - 2\frac{\gamma(t)p(t)}{1-H(t)} + \gamma^2(t) + \int_0^t \frac{p^2(u) h(u)}{(1-H(u))^2} du \right] + O(s^2).
\end{aligned}$$

La esperanza de la segunda integral de la descomposición de $Var(\omega_2(\cdot, Z))$ es, por la aplicación de (3.11) y (3.14),

$$\begin{aligned}
\mathbb{E}[I_2] &= \int K_s(t-v) f(v) \left(\frac{p(v)H(v)}{1-H(v)} - \gamma(v) \right) dv \\
&\quad \times \int K_s(t-v_1) f(v_1) \int_0^{v_1} \frac{p'(v_2)H(v_2)}{1-H(v_2)} dv_2 dv_1 \\
&= f^2(t) \left(\frac{p(t)H(t)}{1-H(t)} - \gamma(t) \right)^2 + O(s^2).
\end{aligned}$$

Para terminar, el tercer término de la descomposición es, por (3.14),

$$\begin{aligned}
I_3 &= \left[\int K_s(t-v) f(v) \left(\frac{p(v)H(v)}{1-H(v)} - \gamma(v) \right) dv \right]^2 \\
&= f^2(t) \left(\frac{p(t)H(t)}{1-H(t)} - \gamma(t) \right)^2 + O(s^2).
\end{aligned}$$

Todos estos cálculos permiten concluir que

$$Var(\omega_2(t, Z)) = f^2(t) \left(\frac{p^2(t)H(t)}{1-H(t)} - 2\gamma(t)p(t) + \int_0^t \frac{p^2(u)h(u)}{(1-H(u))^2} du \right) + O(s^2).$$

Esta varianza se puede expresar también de la siguiente manera:

$$\begin{aligned}
Var(\omega_2(t, Z)) &= 2f^2(t) \left(\frac{p^2(t)H(t)}{1-H(t)} - \gamma(t)p(t) \right. \\
&\quad \left. + \frac{1}{2} \int_0^t \frac{p^2(u)h(u)}{(1-H(u))^2} du - \frac{1}{2} \frac{p^2(t)H(t)}{1-H(t)} \right) + O(s^2).
\end{aligned}$$

La diferencia entre los dos primeros términos, por la propiedad 3.2.2, es igual a

$$p(t) \int_0^t \frac{p'(v) H(v)}{1-H(v)} dv,$$

mientras que los dos últimos términos son el resultado de aplicar integración por partes a la integral

$$\int_0^t \frac{p(v) p'(v) H(v)}{1-H(v)} dv.$$

De esto se deduce que

$$\text{Var}(\omega_2(t, Z)) = 2f^2(t) \int_0^t (p(t) - p(v)) \frac{H(v) p'(v)}{1-H(v)} dv + O(s^2).$$

■

Lema 3.2.4 *Bajo las condiciones (K1), (H2), (p1) y (f1) se tiene*

$$\text{Var}(\rho_2(t, Z)) = \frac{(1-F(t))^2}{1-H(t)} p'(t)^2 H(t) - 2s \left(\frac{1-F(t)}{1-H(t)} \right)^2 p'(t)^2 h(t) e_K + O(s^2).$$

Demostración.

De las expresiones (3.7) y (2.13) se obtiene que

$$\rho_2(t, Z) = \int K_s(t-v) \frac{1-F(v)}{1-H(v)} p'(v) (\mathbf{1}_{\{Z \leq v\}} - H(v)) dv. \quad (3.21)$$

La esperanza de $\rho_2(\cdot, Z)$ es cero, dado que $\mathbb{E}[g_2(\cdot, Z)] = 0$, y por tanto su varianza es igual a

$$\begin{aligned} \text{Var}(\rho_2(t, Z)) &= \mathbb{E}[\rho_2^2(t, Z)] \\ &= \mathbb{E} \left(\iint K_s(t-v_1) K_s(t-v_2) \frac{1-F(v_1)}{1-H(v_1)} p'(v_1) \frac{1-F(v_2)}{1-H(v_2)} p'(v_2) \right. \\ &\quad \times \mathbf{1}_{\{Z \leq v_1\}} \mathbf{1}_{\{Z \leq v_2\}} dv_1 dv_2 \\ &\quad - 2\mathbb{E} \left(\iint K_s(t-v_1) K_s(t-v_2) \frac{1-F(v_1)}{1-H(v_1)} p'(v_1) \frac{1-F(v_2)}{1-H(v_2)} H(v_2) p'(v_2) \right. \\ &\quad \times \mathbf{1}_{\{Z \leq v_1\}} dv_1 dv_2 \left. \right) + \left[\int K_s(t-v) \frac{1-F(v)}{1-H(v)} H(v) p'(v) dv \right]^2 = \mathbb{E}[I_1] - 2\mathbb{E}[I_2] + I_3. \end{aligned}$$

Por (3.17), la esperanza de la primera integral es igual a

$$\begin{aligned}
& \mathbb{E}[I_1] \\
&= \iint K_s(t-v_1) K_s(t-v_2) \frac{1-F(v_1)}{1-H(v_1)} p'(v_1) \frac{1-F(v_2)}{1-H(v_2)} p'(v_2) H(v_1 \wedge v_2) dv_1 dv_2 \\
&= \left(\frac{1-F(t)}{1-H(t)} \right)^2 p'(t)^2 H(t) - 2s \left(\frac{1-F(t)}{1-H(t)} \right)^2 p'(t)^2 h(t) e_K + O(s^2).
\end{aligned}$$

Las esperanzas de la segunda y tercera integral son, aplicando (3.14), igual a

$$\begin{aligned}
\mathbb{E}[I_2] &= \iint K_s(t-v_1) K_s(t-v_2) \frac{1-F(v_1)}{1-H(v_1)} p'(v_1) H(v_1) \\
&\quad \times \frac{1-F(v_2)}{1-H(v_2)} H(v_2) p'(v_2) dv_1 dv_2 = \left(\frac{1-F(t)}{1-H(t)} \right)^2 p'(t)^2 H^2(t) + O(s^2)
\end{aligned}$$

y

$$I_3 = \left[\int K_s(t-v) \frac{1-F(v)}{1-H(v)} H(v) p'(v) dv \right]^2 = \left(\frac{1-F(t)}{1-H(t)} \right)^2 p'(t)^2 H^2(t) + O(s^2)$$

respectivamente. Con la suma de estas tres integrales queda demostrado el lema. ■

Pasamos ahora a calcular las covarianzas que no involucren $\rho_3(\cdot, Z)$ ni $\omega_3(\cdot, Z, \delta)$. Empezaremos por la covarianza entre $\rho_1(\cdot, Z)$ y $\omega_1(\cdot, Z)$.

Lema 3.2.5 *Bajo las condiciones (K1), (H2), (p1) y (f1) se tiene*

$$\begin{aligned}
& Cov(\rho_1(t, Z), \omega_1(t, Z)) \\
&= \frac{1}{2} \frac{1-F(t)}{(1-H(t))^2} p^2(t) h(t) f(t) + \frac{1-F(t)}{1-H(t)} p(t) p'(t) f(t) H(t) \\
&\quad + s \left[f^2(t) \frac{p^2(t) h(t)}{(1-H(t))^2} - f(t) (1-F(t)) \left(\frac{p^2(t) h(t)}{(1-H(t))^2} \right)' \right. \\
&\quad \left. + f'(t) (1-F(t)) \frac{p^2(t) h(t)}{(1-H(t))^2} \right] e_K + O(s^2).
\end{aligned}$$

Demostración.

Dado que $\mathbb{E}[\rho_1(\cdot, Z)] = \mathbb{E}[\omega_1(\cdot, Z)] = 0$, la covarianza de ambas funciones es la esperanza de su producto. Si definimos las siguientes funciones,

$$\begin{aligned}
m(t) &= \frac{1-F(t)}{1-H(t)} \left(p'(t) + \frac{p(t) h(t)}{1-H(t)} \right), \\
n(t) &= \frac{1-F(t)}{1-H(t)} \left(p'(t) H(t) + \frac{p(t) h(t)}{1-H(t)} \right),
\end{aligned}$$

el producto entre $\rho_1(\cdot, Z)$ y $\omega_1(\cdot, Z)$ es igual a

$$\begin{aligned}
& \rho_1(t, Z) \omega_1(t, Z) \\
&= \iint K_s(t - v_1) K_s(t - v_2) \frac{f(v_1)p(v_1)}{1 - H(v_1)} m(v_2) \mathbf{1}_{\{Z \leq v_1\}} \mathbf{1}_{\{Z \leq v_2\}} dv_1 dv_2 \\
&\quad - \int K_s(t - v_1) \frac{f(v_1)p(v_1)}{1 - H(v_1)} H(v_1) dv_1 \int K_s(t - v_2) m(v_2) \mathbf{1}_{\{Z \leq v_2\}} dv_2 \\
&\quad + K_s(t - Z) \frac{1 - F(Z)}{1 - H(Z)} p(Z) \int K_s(t - v_1) \frac{f(v_1)p(v_1)}{1 - H(v_1)} \mathbf{1}_{\{Z \leq v_1\}} dv_1 \\
&\quad - K_s(t - Z) \frac{1 - F(Z)}{1 - H(Z)} p(Z) \int K_s(t - v_1) \frac{f(v_1)p(v_1)}{1 - H(v_1)} H(v_1) dv_1 \\
&\quad - \iint K_s(t - v_1) K_s(t - v_2) \frac{f(v_1)p(v_1)}{1 - H(v_1)} \mathbf{1}_{\{Z \leq v_1\}} n(v_2) dv_1 dv_2 \\
&\quad + \iint K_s(t - v_1) K_s(t - v_2) \frac{f(v_1)p(v_1)}{1 - H(v_1)} H(v_1) n(v_2) dv_1 dv_2 \\
&= I_1 - I_2 + I_3 - I_4 - I_5 + I_6.
\end{aligned}$$

Calculamos ahora la esperanza de cada uno de estos sumandos, teniendo en cuenta que la esperanza de la resta de las dos últimas integrales es cero, es decir, $\mathbb{E}[I_5] = \mathbb{E}[I_6]$. Empezamos calculando la esperanza del primero de los sumandos, para lo cual aplicamos la propiedad (3.17):

$$\begin{aligned}
\mathbb{E}[I_1] &= \iint K_s(t - v_1) K_s(t - v_2) \frac{f(v_1)p(v_1)}{1 - H(v_1)} m(v_2) H(v_1 \wedge v_2) dv_1 dv_2 \\
&= \frac{f(t)p(t)}{1 - H(t)} m(t) H(t) - 2s \frac{f(t)p(t)}{1 - H(t)} m(t) h(t) e_K + O(s^2).
\end{aligned}$$

Por la propiedad (3.14), la segunda esperanza es igual a

$$\begin{aligned}
\mathbb{E}[I_2] &= \int K_s(t - v_1) \frac{f(v_1)p(v_1)}{1 - H(v_1)} H(v_1) dv_1 \int K_s(t - v_2) m(v_2) H(v_2) dv_2 \\
&= \frac{f(t)p(t)}{1 - H(t)} H^2(t) m(t) + O(s^2),
\end{aligned}$$

mientras que la tercera, en virtud a la propiedad (3.16), es

$$\mathbb{E}[I_3] = \iint K_s(t - v) K_s(t - u) \frac{f(v)p(v)}{1 - H(v)} \frac{1 - F(u)}{1 - H(u)} p(u) h(u) \mathbf{1}_{\{u \leq v\}} dv du$$

$$\begin{aligned}
&= \frac{1}{2} \frac{1-F(t)}{(1-H(t))^2} f(t) p^2(t) h(t) - s \left[\left(\frac{1-F(t)}{1-H(t)} p(t) h(t) \right)' \frac{f(t) p(t)}{1-H(t)} \right. \\
&\quad \left. - \frac{1-F(t)}{1-H(t)} p(t) h(t) \left(\frac{f(t) p(t)}{1-H(t)} \right)' \right] e_K + O(s^2).
\end{aligned}$$

Por último, aplicando de nuevo (3.14), la cuarta integral se puede escribir como

$$\begin{aligned}
\mathbb{E}[I_4] &= \int K_s(t-u) \frac{1-F(u)}{1-H(u)} p(u) h(u) du \times \int K_s(t-v_1) \frac{f(v_1) p(v_1)}{1-H(v_1)} H(v_1) dv_1 \\
&= \frac{1-F(t)}{(1-H(t))^2} f(t) p^2(t) h(t) H(t) + O(s^2).
\end{aligned}$$

La suma de todos estos términos da como resultado la expresión del enunciado del lema. ■

Lema 3.2.6 *Bajo las condiciones (K1), (H2), (p1) y (f1) se tiene*

$$\text{Cov}(\omega_1(t, Z) \omega_2(t, Z)) = f^2(t) p(t) \left(\frac{H(t) p(t)}{1-H(t)} - \gamma(t) \right) + O(s^2).$$

Demostración.

De nuevo se tiene que $\mathbb{E}[\omega_1(\cdot, Z)] = \mathbb{E}[\omega_2(\cdot, Z)] = 0$, de modo que la covarianza entre $\omega_1(\cdot, Z)$ y $\omega_2(\cdot, Z)$ es igual a la esperanza de

$$\begin{aligned}
\omega_1(t, Z) \omega_2(t, Z) &= \iint K_s(t-v_1) K_s(t-v_2) \frac{f(v_1) p(v_1)}{1-H(v_1)} f(v_2) \\
&\quad \times \left(\int_0^{v_2} \frac{p'(v_3)}{1-H(v_3)} \mathbf{1}_{\{Z \leq v_3\}} dv_3 \right) \mathbf{1}_{\{Z \leq v_1\}} dv_1 dv_2 \\
&\quad - \iint K_s(t-v_1) K_s(t-v_2) \frac{f(v_1) p(v_1)}{1-H(v_1)} H(v_1) f(v_2) \\
&\quad \times \left(\int_0^{v_2} \frac{p'(v_3)}{1-H(v_3)} \mathbf{1}_{\{Z \leq v_3\}} dv_3 \right) dv_1 dv_2 \\
&\quad - \iint K_s(t-v_1) K_s(t-v_2) \frac{f(v_1) p(v_1)}{1-H(v_1)} (\mathbf{1}_{\{Z \leq v_1\}} - H(v_1)) \\
&\quad \times f(v_2) \left(\frac{p(v_2) H(v_2)}{1-H(v_2)} - \gamma(v_2) \right) dv_1 dv_2 \\
&= I_1 - I_2 - I_3.
\end{aligned}$$

La esperanza del último sumando es cero, $\mathbb{E}[I_3] = 0$, mientras que, por la aplicación, en primer lugar, de la propiedad 3.2.1, y posteriormente la propiedad 3.2.3 (paso (a)), la propiedad 3.2.10 (paso (b)), y la propiedad 3.2.3 (paso (c)), la esperanza de la primera integral es

$$\begin{aligned}
\mathbb{E}[I_1] &= \mathbb{E} \left[\iint K_s(t-v_1) K_s(t-v_2) \frac{f(v_1)p(v_1)}{1-H(v_1)} f(v_2) \right. \\
&\quad \left. \times \left(\frac{p(v_2)}{1-H(v_2)} - \gamma(v_2) - \frac{p(Z)}{1-H(Z)} + \gamma(Z) \right) \mathbf{1}_{\{Z \leq v_1 \wedge v_2\}} dv_1 dv_2 \right] \\
&\stackrel{(a)}{=} \iint K_s(t-v_1) K_s(t-v_2) \frac{f(v_1)p(v_1)}{1-H(v_1)} \\
&\quad \times f(v_2) \left(\frac{p(v_2)}{1-H(v_2)} - \gamma(v_2) \right) H(v_1 \wedge v_2) dv_1 dv_2 \\
&\quad - \iint K_s(t-v_1) K_s(t-v_2) \frac{f(v_1)p(v_1)}{1-H(v_1)} f(v_2) \gamma(v_1 \wedge v_2) (1-H(v_1 \wedge v_2)) dv_1 dv_2 \\
&\stackrel{(b)}{=} \frac{f^2(t)p(t)}{1-H(t)} \left(\frac{p(t)}{1-H(t)} - \gamma(t) \right) H(t) - 2s \frac{f^2(t)p(t)}{1-H(t)} \left(\frac{p(t)}{1-H(t)} - \gamma(t) \right) h(t) e_K \\
&\quad - \frac{f^2(t)p(t)}{1-H(t)} \gamma(t) (1-H(t)) + 2s \frac{f^2(t)p(t)}{1-H(t)} [\gamma(t) (1-H(t))] e_K + O(s^2) \\
&\stackrel{(c)}{=} \frac{f^2(t)p(t)}{1-H(t)} \left(\frac{p(t)H(t)}{1-H(t)} - \gamma(t) \right) + O(s^2).
\end{aligned}$$

La esperanza $\mathbb{E}[I_2]$ es, aplicando las propiedades 3.2.2 (paso (a)) y 3.2.7 (paso (b)),

$$\begin{aligned}
&\mathbb{E}[I_2] \\
&= \iint K_s(t-v_1) K_s(t-v_2) \frac{f(v_1)p(v_1)}{1-H(v_1)} H(v_1) f(v_2) \int_0^{v_2} \frac{p'(v_3)H(v_3)}{1-H(v_3)} dv_3 dv_1 dv_2 \\
&\stackrel{(a)}{=} \iint K_s(t-v_1) K_s(t-v_2) \frac{f(v_1)p(v_1)}{1-H(v_1)} H(v_1) f(v_2) \left(\frac{p(v_2)H(v_2)}{1-H(v_2)} - \gamma(v_2) \right) dv_1 dv_2 \\
&\stackrel{(b)}{=} \frac{f^2(t)p(t)}{1-H(t)} H(t) \left(\frac{p(t)H(t)}{1-H(t)} - \gamma(t) \right) + O(s^2).
\end{aligned}$$

Con la diferencia entre ambas esperanzas se da por concluida la demostración. ■

Lema 3.2.7 *Bajo las condiciones (K1), (H2), (p1) y (f1) se tiene*

$$\begin{aligned}
&\text{Cov}(\omega_1(t, Z), \rho_2(t, Z)) \\
&= \frac{1-F(t)}{1-H(t)} p(t) p'(t) H(t) f(t) - 2s \frac{1-F(t)}{(1-H(t))^2} p(t) p'(t) h(t) f(t) e_K + O(s).
\end{aligned}$$

Demostración.

Puesto que $\mathbb{E}[\omega_1(\cdot, Z)] = \mathbb{E}[\rho_2(\cdot, Z)] = 0$, la covarianza entre $\omega_1(\cdot, Z)$ y $\rho_2(\cdot, Z)$ es igual a $\mathbb{E}[\omega_1(\cdot, Z)\rho_2(\cdot, Z)]$. De las expresiones de $\omega_1(\cdot, Z)$ y $\rho_2(\cdot, Z)$ dadas en (3.18) y (3.21) respectivamente, se obtiene que el producto de ambas funciones es igual a

$$\begin{aligned} & \omega_1(t, Z), \rho_2(t, Z) = \\ & \iint K_s(t - v_1) K_s(t - v_2) \frac{f(v_1)p(v_1)}{1 - H(v_1)} \frac{1 - F(v_2)}{1 - H(v_2)} p'(v_2) \mathbf{1}_{\{Z \leq v_1 \wedge v_2\}} dv_1 dv_2 \\ & - \iint K_s(t - v_1) K_s(t - v_2) \frac{f(v_1)p(v_1)}{1 - H(v_1)} H(v_1) \frac{1 - F(v_2)}{1 - H(v_2)} p'(v_2) \mathbf{1}_{\{Z \leq v_2\}} dv_1 dv_2 \\ & - \iint K_s(t - v_1) K_s(t - v_2) \frac{f(v_1)p(v_1)}{1 - H(v_1)} \frac{1 - F(v_2)}{1 - H(v_2)} H(v_2) p'(v_2) \mathbf{1}_{\{Z \leq v_1\}} dv_1 dv_2 \\ & + \iint K_s(t - v_1) K_s(t - v_2) \frac{f(v_1)p(v_1)}{1 - H(v_1)} H(v_1) \frac{1 - F(v_2)}{1 - H(v_2)} H(v_2) p'(v_2) dv_1 dv_2 \\ & = I_1 - I_2 - I_3 + I_4. \end{aligned}$$

La media del tercer sumando en la expresión anterior coincide con la media del cuarto sumando, es decir $\mathbb{E}[I_3] = \mathbb{E}[I_4]$, de modo que la covarianza $Cov(\omega_1(\cdot, Z), \rho_2(\cdot, Z))$ es igual a la esperanza del primer sumando menos la del segundo. La esperanza del primer sumando es, aplicando (3.17),

$$\begin{aligned} & \mathbb{E}[I_1] \\ & = \iint K_s(t - v_1) K_s(t - v_2) \frac{f(v_1)p(v_1)}{1 - H(v_1)} \frac{1 - F(v_2)}{1 - H(v_2)} p'(v_2) H(v_1 \wedge v_2) dv_1 dv_2 \\ & = \frac{1 - F(t)}{(1 - H(t))^2} p(t) p'(t) H(t) f(t) - 2s \frac{1 - F(t)}{(1 - H(t))^2} p(t) p'(t) h(t) f(t) e_K + O(s^2), \end{aligned}$$

mientras que, por la aplicación de (3.14), la esperanza del segundo sumando es igual a

$$\begin{aligned} \mathbb{E}[I_2] & = \iint K_s(t - v_1) K_s(t - v_2) \frac{f(v_1)p(v_1)}{1 - H(v_1)} H(v_1) \frac{1 - F(v_2)}{1 - H(v_2)} p'(v_2) H(v_2) dv_1 dv_2 \\ & = \frac{1 - F(t)}{(1 - H(t))^2} p(t) p'(t) H^2(t) f(t) + O(s^2). \end{aligned}$$

El resultado del presente lema se obtiene con la diferencia entre estas dos esperanzas. ■

Lema 3.2.8 *Bajo las condiciones (K1), (H2), (p1) y (f1) se tiene*

$$\begin{aligned} \text{Cov}(\rho_1(t, Z), \omega_2(t, Z)) &= \frac{1 - F(t)}{1 - H(t)} f(t) [H(t)p(t) - \gamma(t)(1 - H(t))] p'(t) \\ &\quad + 2s \frac{1 - F(t)}{(1 - H(t))^2} p(t) p'(t) h(t) f(t) e_K + O(s^2). \end{aligned}$$

Demostración.

Las esperanzas de $\omega_2(\cdot, Z)$ y $\rho_1(\cdot, Z)$ son cero, de modo que para calcular su covarianza basta obtener la esperanza del producto entre ambas funciones, dadas en (3.20) y (3.19) respectivamente. Si definimos previamente las funciones $m(\cdot)$ y $n(\cdot)$ como

$$\begin{aligned} m(t) &= \frac{1 - F(t)}{1 - H(t)} \left(p'(t) + \frac{p(t)h(t)}{1 - H(t)} \right), \\ n(t) &= \frac{1 - F(t)}{1 - H(t)} \left(p'(t)H(t) + \frac{p(t)h(t)}{1 - H(t)} \right), \end{aligned} \quad (3.22)$$

este producto se puede escribir de la forma

$$\begin{aligned} &\rho_1(t, Z), \omega_2(t, Z) \\ &= \int K_s(t - v_1) m(v_1) \mathbf{1}_{\{Z \leq v_1\}} dv_1 \int K_s(t - v_2) f(v_2) \int_0^{v_2} \frac{p'(v_3)}{1 - H(v_3)} \mathbf{1}_{\{Z \leq v_3\}} dv_2 dv_3 \\ &\quad + K_s(t - Z) \frac{1 - F(Z)}{1 - H(Z)} p(Z) \int K_s(t - v_1) f(v_1) \int_0^{v_1} \frac{p'(v_2)}{1 - H(v_2)} \mathbf{1}_{\{Z \leq v_2\}} dv_1 dv_2 \\ &\quad - \int K_s(t - v_1) n(v_1) dv_1 \int K_s(t - v_2) f(v_2) \int_0^{v_2} \frac{p'(v_3)}{1 - H(v_3)} \mathbf{1}_{\{Z \leq v_3\}} dv_2 dv_3 \\ &\quad - \int K_s(t - v_1) m(v_1) \mathbf{1}_{\{Z \leq v_1\}} dv_1 \int K_s(t - v_2) f(v_2) \left(\frac{p(v_2)H(v_2)}{1 - H(v_2)} - \gamma(v_2) \right) dv_2 \\ &\quad - K_s(t - Z) \frac{1 - F(Z)}{1 - H(Z)} p(Z) \int K_s(t - v) f(v) \left(\frac{p(v)H(v)}{1 - H(v)} - \gamma(v) \right) dv \\ &\quad + \int K_s(t - v_1) n(v_1) dv_1 \int K_s(t - v_2) f(v_2) \left(\frac{p(v_2)H(v_2)}{1 - H(v_2)} - \gamma(v_2) \right) dv_2 \\ &= I_1 + I_2 - I_3 - I_4 - I_5 + I_6. \end{aligned}$$

Para calcular la esperanza del primer término del producto entre $\omega_2(\cdot, Z)$ y $\rho_1(\cdot, Z)$ se usa en primer lugar la propiedad 3.2.1 y se toman esperanzas (paso (a)), se aplican las propiedades 3.2.10 y 3.2.3 (paso (b)) y, finalmente, se sustituye la función $m(\cdot)$ por su expresión dada en (3.22):

$$\begin{aligned}
\mathbb{E}[I_1] &= \mathbb{E} \left[\iint K_s(t-v_1) K_s(t-v_2) m(v_1) f(v_2) \left(\frac{p(v_2)}{1-H(v_2)} - \gamma(v_2) \right. \right. \\
&\quad \left. \left. - \frac{p(Z)}{1-H(Z)} + \gamma(Z) \right) \mathbf{1}_{\{Z \leq v_1 \wedge v_2\}} dv_3 dv_1 \right] \\
&\stackrel{(a)}{=} \iint K_s(t-v_1) K_s(t-v_2) m(v_1) f(v_2) \left(\frac{p(v_2)}{1-H(v_2)} - \gamma(v_2) \right) H(v_1 \wedge v_2) dv_2 dv_1 \\
&\quad - \iint K_s(t-v_1) K_s(t-v_2) m(v_1) f(v_2) \\
&\quad \times \int_0^{v_1 \wedge v_2} \left(\frac{p(u)}{1-H(u)} - \gamma(u) \right) h(u) du dv_2 dv_1 \\
&\stackrel{(b)}{=} m(t) f(t) \left(\frac{p(t)}{1-H(t)} - \gamma(t) \right) H(t) - 2sm(t) f(t) \left(\frac{p(t)}{1-H(t)} - \gamma(t) \right) h(t) e_K \\
&\quad - m(t) f(t) \gamma(t) (1-H(t)) + 2sm(t) f(t) \left(\frac{p(t)}{1-H(t)} - \gamma(t) \right) h(t) e_K + O(s^2) \\
&\stackrel{(c)}{=} f(t) \frac{1-F(t)}{1-H(t)} \left(p'(t) + \frac{p(t)h(t)}{1-H(t)} \right) \left(\frac{p(t)H(t)}{1-H(t)} - \gamma(t) \right) + O(s^2).
\end{aligned}$$

Con respecto al segundo término, en el cálculo de su esperanza se aplicaron la propiedad 3.2.1 (paso (a)), la propiedad 3.2.9 (paso (b)) y de nuevo la propiedad 3.2.1 (paso (c)). Se tiene por tanto

$$\begin{aligned}
\mathbb{E}[I_2] &\stackrel{(a)}{=} \mathbb{E} \left[\int K_s(t-v) K_s(t-Z) \frac{1-F(Z)}{1-H(Z)} p(Z) f(v) \left(\frac{p(v)}{1-H(v)} - \gamma(v) \right) \mathbf{1}_{\{Z \leq v\}} dv \right] \\
&\quad - \mathbb{E} \left[\int K_s(t-v) K_s(t-Z) \frac{1-F(Z)}{1-H(Z)} p(Z) f(v) \left(\frac{p(Z)}{1-H(Z)} - \gamma(Z) \right) \mathbf{1}_{\{Z \leq v\}} dv \right] \\
&\stackrel{(b)}{=} 2s \frac{1-F(t)}{1-H(t)} p(t) h(t) f(t) \left(\frac{p(t)}{1-H(t)} - \gamma(t) \right)' e_K + O(s^2) \\
&\stackrel{(c)}{=} 2s \frac{1-F(t)}{(1-H(t))^2} p(t) p'(t) h(t) f(t) e_K + O(s^2)
\end{aligned}$$

La esperanza de la tercera integral es, por la propiedad 3.2.1 (paso (a)), las propiedades 3.2.3 y 3.2.7 y por la definición de la función $n(\cdot)$ (paso (b)), igual a

$$\begin{aligned}
\mathbb{E}[I_3] &\stackrel{(a)}{=} \int K_s(t-v_1) n(v_1) dv_1 \times \int K_s(t-v_2) f(v_2) \left(\frac{p(v_2)}{1-h(v_2)} - \gamma(v_2) \right) dv_2 \\
&\quad - \int K_s(t-v_1) n(v_1) dv_1 \int K_s(t-v_2) f(v_2) \mathbb{E} \left[\left(\frac{p(Z)}{1-h(Z)} - \gamma(Z) \right) \mathbf{1}_{\{Z \leq v_2\}} \right] dv_2
\end{aligned}$$

$$\stackrel{(b)}{=} \frac{1-F(t)}{1-H(t)} f(t) \left(p'(t) H(t) + \frac{p(t) h(t)}{1-H(t)} \right) \gamma(t) (1-H(t)) + O(s^2),$$

y la de la cuarta integral, tomando esperanzas (paso (a)), aplicando la propiedad 3.2.7 y teniendo en cuenta la definición de la función $m(\cdot)$ (paso (b)), es igual a

$$\begin{aligned} \mathbb{E}[I_4] &\stackrel{(a)}{=} \int K_s(t-v_1) m(v_1) H(v_1) dv_1 \\ &\quad \times \int K_s(t-v_2) f(v_2) \left(\frac{p(v_2) H(v_2)}{1-H(v_2)} - \gamma(v_2) \right) dv_2 \\ &\stackrel{(b)}{=} \frac{1-F(t)}{1-H(t)} H(t) f(t) \left(p'(t) + \frac{p(t) h(t)}{1-H(t)} \right) \left(\frac{p(t) H(t)}{1-H(t)} - \gamma(t) \right) + O(s^2). \end{aligned}$$

La esperanza del quinto término es, por la propiedad 3.2.7,

$$\begin{aligned} \mathbb{E}[I_5] &= \mathbb{E} \left[K_s(t-Z) \frac{1-F(Z)}{1-H(Z)} p(Z) \int K_s(t-v) f(v) \left(\frac{p(v) H(v)}{1-H(v)} - \gamma(v) \right) dv \right] \\ &= \frac{1-F(t)}{1-H(t)} p(t) h(t) f(t) \left(\frac{p(t) H(t)}{1-H(t)} - \gamma(t) \right) + O(s^2). \end{aligned}$$

Finalmente, el último sumando es, aplicando la propiedad 3.2.7, igual a

$$\begin{aligned} \mathbb{E}[I_6] &= \int K_s(t-v_1) n(v_1) dv_1 \int K_s(t-v_2) f(v_2) \left(\frac{p(v_2) H(v_2)}{1-H(v_2)} - \gamma(v_2) \right) dv_2 \\ &= \frac{1-F(t)}{1-H(t)} \left(p'(t) H(t) + \frac{p(t) h(t)}{1-H(t)} \right) f(t) \left(\frac{p(t) H(t)}{1-H(t)} - \gamma(t) \right) + O(s^2). \end{aligned}$$

Con la suma de todas estas esperanzas se obtiene la expresión de $Cov(\rho_1(\cdot, Z), \omega_2(\cdot, Z))$ que se establece en este lema. ■

Lema 3.2.9 *Bajo las condiciones (K1), (H2), (p1) y (f1) se tiene*

$$\begin{aligned} Cov(\rho_1(t, Z), \rho_2(t, Z)) &= \left(\frac{1-F(t)}{1-H(t)} \right)^2 p'(t) \left[p'(t) H(t) (1-H(t)) + \frac{1}{2} p(t) h(t) \right] \\ &\quad - s \left(\frac{1-F(t)}{1-H(t)} \right)^2 \left(3p'(t)^2 h(t) + 2 \frac{p(t) p'(t)}{1-H(t)} h^2(t) \right. \\ &\quad \left. + p(t) p'(t) h'(t) - p(t) p''(t) h(t) \right) e_K + O(s^2). \end{aligned}$$

Demostración.

La expresión de las funciones $\rho_1(\cdot, Z)$ y $\rho_2(\cdot, Z)$ viene dada en (3.19) y (3.21) respectivamente. Como la esperanza de ambas funciones es cero, la covarianza se reduce a la esperanza del producto de ambas funciones. Si definimos previamente las funciones

$$m(t) = \frac{1-F(t)}{1-H(t)} \left(p'(t) + \frac{p(t)h(t)}{1-H(t)} \right) \text{ y } n(t) = \frac{1-F(t)}{1-H(t)} \left(p'(t)H(t) + \frac{p(t)h(t)}{1-H(t)} \right),$$

este producto es igual a

$$\begin{aligned} \rho_1(t, Z), \rho_2(t, Z) &= \\ & \int K_s(t-v_1) m(v_1) \mathbf{1}_{\{Z \leq v_1\}} dv_1 \int K_s(t-v_2) \frac{1-F(v_2)}{1-H(v_2)} p'(v_2) \mathbf{1}_{\{Z \leq v_2\}} dv_2 \\ & - \int K_s(t-v_1) m(v_1) \mathbf{1}_{\{Z \leq v_1\}} dv_1 \int K_s(t-v_2) \frac{1-F(v_2)}{1-H(v_2)} H(v_2) p'(v_2) dv_2 \\ & + K_s(t-Z) \frac{1-F(Z)}{1-H(Z)} p(Z) \int K_s(t-v_2) \frac{1-F(v_2)}{1-H(v_2)} p'(v_2) \mathbf{1}_{\{Z \leq v_2\}} dv_2 \\ & - K_s(t-Z) \frac{1-F(Z)}{1-H(Z)} p(Z) \int K_s(t-v_2) \frac{1-F(v_2)}{1-H(v_2)} H(v_2) p'(v_2) dv_2 \\ & - \int K_s(t-v_2) \frac{1-F(v_2)}{1-H(v_2)} p'(v_2) \mathbf{1}_{\{Z \leq v_2\}} dv_2 \int K_s(t-v_1) n(v_1) dv_1 \\ & + \int K_s(t-v_2) \frac{1-F(v_2)}{1-H(v_2)} H(v_2) p'(v_2) dv_2 \int K_s(t-v_1) n(v_1) dv_1 \\ & = I_1 - I_2 + I_3 - I_4 - I_5 + I_6. \end{aligned}$$

La esperanza de la resta de las dos últimas integrales es cero, es decir $\mathbb{E}[I_5] = \mathbb{E}[I_6]$, mientras que la primera de las cuatro integrales restantes es, por la propiedad 3.2.10:

$$\begin{aligned} \mathbb{E}[I_1] &= \iint K_s(t-v_1) K_s(t-v_2) m(v_1) \frac{1-F(v_2)}{1-H(v_2)} p'(v_2) H(v_1 \wedge v_2) dv_1 dv_2 \\ &= m(t) \frac{1-F(t)}{1-H(t)} p'(t) H(t) - 2sm(t) \frac{1-F(t)}{1-H(t)} p'(t) h(t) e_K + O(s^2). \end{aligned}$$

Aplicando la propiedad 3.2.7, la esperanza de la segunda integral es igual a

$$\begin{aligned} \mathbb{E}[I_2] &= \int K_s(t-v_1) m(v_1) H(v_1) dv_1 \int K_s(t-v_2) \frac{1-F(v_2)}{1-H(v_2)} H(v_2) p'(v_2) dv_2 \\ &= m(t) H(t) \frac{1-F(t)}{1-H(t)} H(t) p'(t) + O(s^2). \end{aligned}$$

La esperanza del tercer sumando es, por la propiedad 3.2.9,

$$\mathbb{E}[I_3] = \iint K_s(t-v_2) K_s(t-v_1) \frac{1-F(v_2)}{1-H(v_2)} p'(v_2) \frac{1-F(v_1)}{1-H(v_1)} p(v_1) h(v_1) \mathbf{1}_{\{v_1 \leq v_2\}} dv_2$$

$$\begin{aligned}
&= \frac{1}{2} \left(\frac{1-F(t)}{1-H(t)} \right)^2 p(t) p'(t) h(t) - s \left[\left(\frac{1-F(t)}{1-H(t)} p(t) h(t) \right)' \frac{1-F(t)}{1-H(t)} p'(t) \right. \\
&\quad \left. - \frac{1-F(t)}{1-H(t)} p(t) h(t) \left(\frac{1-F(t)}{1-H(t)} p'(t) \right)' \right] e_K + O(s^2).
\end{aligned}$$

Por último, aplicando de nuevo la propiedad 3.2.10, la esperanza de I_4 es igual a

$$\begin{aligned}
&\mathbb{E}[I_4] \\
&= \int K_s(t-v_1) \frac{1-F(v_1)}{1-H(v_1)} p(v_1) h(v_1) dv_1 \int K_s(t-v_2) \frac{1-F(v_2)}{1-H(v_2)} H(v_2) p'(v_2) dv_2 \\
&= \left(\frac{1-F(t)}{1-H(t)} \right)^2 p(t) p'(t) H(t) h(t) + O(s^2).
\end{aligned}$$

Con la suma de estas cuatro esperanzas se concluye lo establecido en este lema. ■

Lema 3.2.10 *Bajo las condiciones (K1), (H2), (p1) y (f1) se tiene*

$$\text{Cov}(\omega_2(t, Z), \rho_2(t, Z)) = (1-F(t)) f(t) p'(t) \left(\frac{H(t)p(t)}{1-H(t)} - \gamma(t) \right) + O(s^2).$$

Demostración.

Como ya se vio en las demostraciones de lemas anteriores, las esperanzas de $\omega_2(\cdot, Z)$ y $\rho_2(\cdot, Z)$ son cero, de modo que la covarianza entre $\omega_2(\cdot, Z)$ y $\rho_2(\cdot, Z)$ es la esperanza del producto entre ambas funciones:

$$\begin{aligned}
\omega_2(t, Z) \rho_2(t, Z) &= \int K_s(t-v_1) f(v_1) \int_0^{v_1} \frac{p'(v_2)}{1-H(v_2)} \mathbf{1}_{\{Z \leq v_2\}} dv_1 dv_2 \\
&\quad \times \int K_s(t-v_3) \frac{1-F(v_3)}{1-H(v_3)} p'(v_3) \mathbf{1}_{\{Z \leq v_3\}} dv_3 \\
&\quad - \int K_s(t-v_1) f(v_1) \int_0^{v_1} \frac{p'(v_2)}{1-H(v_2)} \mathbf{1}_{\{Z \leq v_2\}} dv_1 dv_2 \\
&\quad \times \int K_s(t-v_3) \frac{1-F(v_3)}{1-H(v_3)} H(v_3) p'(v_3) dv_3 \\
&\quad - \int K_s(t-v) f(v) \left(\frac{p(v)H(v)}{1-H(v)} - \gamma(v) \right) dv \times \rho_2(t, Z) \\
&= I_1 - I_2 - I_3.
\end{aligned}$$

La esperanza del tercer término es cero, $\mathbb{E}[I_3] = 0$, por ser nula la esperanza de $\rho_2(\cdot, Z)$, de modo que bastará calcular la de los dos primeros sumandos. La esperanza del primer sumando es consecuencia de la propiedad 3.2.1 (paso (a)), de tomar esperanzas (paso (b)) y de las propiedades 3.2.3 y 3.2.10 (paso (c)):

$$\begin{aligned}
\mathbb{E}[I_1] &\stackrel{(a)}{=} \mathbb{E} \left[\iint K_s(t-v_1) K_s(t-v_3) f(v_1) \frac{1-F(v_3)}{1-H(v_3)} p'(v_3) \right. \\
&\quad \times \left. \left(\frac{p(v_1)}{1-H(v_1)} - \gamma(v_1) - \frac{p(Z)}{1-H(Z)} + \gamma(Z) \right) \mathbf{1}_{\{Z \leq v_1 \wedge v_3\}} dv_3 dv_1 \right] \\
&\stackrel{(b)}{=} \iint K_s(t-v_1) K_s(t-v_3) f(v_1) \frac{1-F(v_3)}{1-H(v_3)} p'(v_3) \\
&\quad \times \left(\frac{p(v_1)}{1-H(v_1)} - \gamma(v_1) \right) H(v_1 \wedge v_3) dv_3 dv_1 \\
&\quad - \iint K_s(t-v_1) K_s(t-v_3) f(v_1) \frac{1-F(v_3)}{1-H(v_3)} p'(v_3) \\
&\quad \times \int \left(\frac{p(u)}{1-H(u)} - \gamma(u) \right) h(u) \mathbf{1}_{\{u \leq v_1 \wedge v_3\}} dudv_3 dv_1 \\
&\stackrel{(c)}{=} f(t) \frac{1-F(t)}{1-H(t)} p'(t) \left(\frac{p(t)}{1-H(t)} - \gamma(t) \right) H(t) \\
&\quad - 2sf(t) \frac{1-F(t)}{1-H(t)} p'(t) \left(\frac{p(t)}{1-H(t)} - \gamma(t) \right) h(t) e_K \\
&\quad - \frac{1-F(t)}{1-H(t)} f(t) p'(t) \gamma(t) (1-H(t)) \\
&\quad + 2s \frac{1-F(t)}{1-H(t)} f(t) p'(t) [\gamma(t) (1-H(t))]' e_K + O(s^2) \\
&= f(t) \frac{1-F(t)}{1-H(t)} p'(t) \left(\frac{p(t)}{1-H(t)} H(t) - \gamma(t) \right) + O(s^2).
\end{aligned}$$

Por otra parte, la esperanza de la segunda integral es, por la aplicación de la propiedad 3.2.1 (paso (a)) y de las propiedades 3.2.3 y 3.2.7 (paso (b)), igual a

$$\begin{aligned}
\mathbb{E}[I_2] &\stackrel{(a)}{=} \mathbb{E} \left[\int K_s(t-v_1) f(v_1) \left(\frac{p(v_1)}{1-H(v_1)} - \gamma(v_1) - \frac{p(Z)}{1-H(Z)} + \gamma(Z) \right) \mathbf{1}_{\{Z \leq v_1\}} dv_1 \right] \\
&\quad \times \int K_s(t-v_3) \frac{1-F(v_3)}{1-H(v_3)} H(v_3) p'(v_3) dv_3 \\
&\stackrel{(b)}{=} \frac{1-F(t)}{1-H(t)} H(t) p'(t) \\
&\quad \times \left[f(t) \left(\frac{p(t)}{1-H(t)} - \gamma(t) \right) H(t) - f(t) \gamma(t) (1-H(t)) \right] + O(s^2) \\
&= \frac{1-F(t)}{1-H(t)} H(t) p'(t) f(t) \left(\frac{p(t)}{1-H(t)} - \gamma(t) \right) + O(s^2).
\end{aligned}$$

De la diferencia entre $\mathbb{E}[I_1]$ y $\mathbb{E}[I_2]$ se obtiene la expresión de la covarianza. ■

A continuación obtenemos las expresiones de las varianzas y covarianzas que involucran ambas ventanas, la de suavizado s y la de presuavizado b , es decir, las que implican la función $\omega_3(\cdot, Z, \delta)$ o $\rho_3(\cdot, Z, \delta)$. Empezamos calculando ahora la varianza del término $\omega_3(\cdot, Z, \delta)$.

Lema 3.2.11 *Bajo las condiciones (K1), (H2), (p1) y (f1) se tiene*

$$\begin{aligned} \text{Var}(\omega_3(t, Z, \delta)) &= f^2(t) \int_0^t q_1(u) du \\ &\quad - 2(s+b) f^2(t) (q_1(t) + q_1(0)) e_K + O(s^2) + O(b^2), \end{aligned}$$

donde

$$q_1(t) = \frac{p(t)(1-p(t))h(t)}{(1-H(t))^2}. \quad (3.23)$$

Demostración.

De las expresiones (3.6) y (2.14) se deriva que la función $\omega_3(\cdot, Z, \delta)$ es igual a

$$\omega_3(t, Z, \delta) = \int K_s(t-v_1) f(v_1) \int_0^{v_1} K_b(v_2-Z) \frac{\delta-p(v_2)}{1-H(v_2)} dv_2 dv_1.$$

Dado que $\mathbb{E}[g_3(t, Z, \delta)] = O(b^2)$, el término dominante de $\text{Var}(\omega_3(t, Z, \delta))$ es

$$\begin{aligned} \mathbb{E}[\omega_3^2(t, Z, \delta)] &= \mathbb{E} \left[\iint K_s(t-v_1) K_s(t-w_1) f(v_1) f(w_1) \right. \\ &\quad \left. \times \int_0^{v_1} K_b(v_2-Z) \frac{\delta-p(v_2)}{1-H(v_2)} \int_0^{w_1} K_b(w_2-Z) \frac{\delta-p(w_2)}{1-H(w_2)} dw_2 dv_2 dw_1 dv_1 \right] \\ &= \mathbb{E} \left[\iiint K_s(t-v_1) K_s(t-w_1) K_b(v_2-Z) K_b(w_2-Z) f(v_1) f(w_1) \right. \\ &\quad \left. \times \frac{\delta - \delta(p(v_2) + p(w_2)) + p(v_2)p(w_2)}{(1-H(v_2))(1-H(w_2))} \mathbf{1}_{\{0 \leq v_2 \leq v_1\}} \mathbf{1}_{\{0 \leq w_2 \leq w_1\}} dw_2 dv_2 dw_1 dv_1 \right] \\ &= \int \int \int \int \int K_s(t-v_1) K_s(t-w_1) K_b(v_2-u) K_b(w_2-u) f(v_1) f(w_1) h(u) \\ &\quad \times \frac{p(u) - p(u)(p(v_2) + p(w_2)) + p(v_2)p(w_2)}{(1-H(v_2))(1-H(w_2))} \\ &\quad \times \mathbf{1}_{\{0 \leq v_2 \leq v_1\}} \mathbf{1}_{\{0 \leq w_2 \leq w_1\}} dw_2 dv_2 dw_1 dv_1 du. \end{aligned}$$

Realizamos ahora los cambios de variable $v_1 = t - sx_1$, $w_1 = t - sz_1$, $v_2 = u + bx_2$ y $w_2 = u + bz_2$:

$$\begin{aligned}
& \mathbb{E} [\omega_3^2(t, Z, \delta)] \\
&= \int \int \int \int \int K(x_1) K(z_1) K(x_2) K(z_2) f(t - sx_1) f(t - sz_1) h(u) \\
&\quad \times \frac{p(u) - p(u)(p(u + bx_2) + p(u + bz_2)) + p(u + bx_2)p(u + bz_2)}{(1 - H(u + bx_2))(1 - H(u + bz_2))} \\
&\quad \times \mathbf{1}_{\{0 \leq u + bx_2 \leq t - sx_1\}} \mathbf{1}_{\{0 \leq u + bz_2 \leq t - sz_1\}} dx_1 dz_1 dx_2 dz_2 du
\end{aligned}$$

Si definimos la función $m(x, y, z)$ como

$$m(x, y, z) = \frac{p(x) - p(x)(p(y) + p(z)) + p(y)p(z)}{(1 - H(y))(1 - H(z))}$$

podemos escribir el momento de orden 2 de $\omega_3(t, Z, \delta)$ como:

$$\begin{aligned}
& \mathbb{E} [\omega_3^2(t, Z, \delta)] \\
&= \int \int \int \int \int K(x_1) K(z_1) K(x_2) K(z_2) f(t - sx_1) f(t - sz_1) h(u) \\
&\quad \times \zeta(u, u + bx_2, u + bz_2) \mathbf{1}_{\{-bx_2 \leq u \leq t - sx_1 - bx_2\}} \mathbf{1}_{\{-bz_2 \leq u \leq t - sz_1 - bz_2\}} dx_1 dz_1 dx_2 dz_2 du \\
&= f^2(t) \int \int \int \int \int K(x_1) K(z_1) K(x_2) K(z_2) m(u, u + bx_2, u + bz_2) h(u) \\
&\quad \times \mathbf{1}_{\{-b(x_2 \wedge z_2) \leq u \leq t - s(x_1 \vee z_1) - b(x_2 \vee z_2)\}} dx_1 dz_1 dx_2 dz_2 du \\
&\quad - sf(t) f'(t) \int \int \int \int \int (x_1 + z_1) K(x_1) K(z_1) K(x_2) K(z_2) m(u, u + bx_2, u + bz_2) \\
&\quad \times h(u) \mathbf{1}_{\{-b(x_2 \wedge z_2) \leq u \leq t - s(x_1 \vee z_1) - b(x_2 \vee z_2)\}} dx_1 dz_1 dx_2 dz_2 du + O(s^2) \\
&= f^2(t) \int \int \int \int K(x_1) K(z_1) K(x_2) K(z_2) \int m(u, u, u) h(u) \\
&\quad \times \mathbf{1}_{\{-b(x_2 \wedge z_2) \leq u \leq t - s(x_1 \vee z_1) - b(x_2 \vee z_2)\}} du dx_1 dz_1 dx_2 dz_2 + O(s^2) + o(b).
\end{aligned}$$

Definimos ahora la siguiente función:

$$n(t) = \int h(u) m(u, u, u) \mathbf{1}_{\{0 \leq u \leq t\}} du = \int_0^t \frac{p(u)(1 - p(u))h(u)}{(1 - H(u))^2} du = \int_0^t q_1(u) du,$$

de modo que el término dominante en la expresión de $\mathbb{E} [\omega_3^2(t, Z, \delta)]$ se puede escribir también como

$$\begin{aligned}
& f^2(t) \iiint\!\!\!\int K(x_1) K(z_1) K(x_2) K(z_2) \\
& \quad \times n(t - s(x_1 \vee z_1) - b(x_2 \vee z_2)) \, dx_1 dz_1 dx_2 dz_2 \\
& \quad - f^2(t) \iint K(x_2) K(z_2) n(-b(x_2 \wedge z_2)) \, dx_2 dz_2.
\end{aligned}$$

Mediante un nuevo desarrollo de Taylor de grado dos en la función $n(t - s(x_1 \vee z_1) - b(x_2 \vee z_2))$ en torno a $t > 0$, y de grado uno en la función $n(-b(x_2 \wedge z_2))$ en torno a cero, la expresión anterior es igual a

$$\begin{aligned}
& f^2(t) n(t) - f^2(t) n'(t) (s + b) \iint (x_1 \vee z_1) K(x_1) K(z_1) \, dx_1 dz_1 \\
& + \frac{1}{2} f^2(t) \iiint\!\!\!\int [s(x_1 \vee z_1) - b(x_2 \vee z_1)]^2 K(x_1) K(z_1) K(x_2) K(z_2) \\
& \quad \times n''(\tau_{t,s,x_1,z_1,b,x_2,z_1}) \, dx_1 dz_1 dx_2 dz_2 \\
& - f^2(t) n(0) + b f^2(t) n'(0) \iint (x_2 \wedge z_2) K(x_2) K(z_2) \, dx_2 dz_2 \\
& - \frac{1}{2} b^2 \iint (x_2 \wedge z_2)^2 K(x_2) K(z_2) n''(\tau_{b,x_2,z_2}) \, dx_2 dz_2.
\end{aligned}$$

donde $\tau_{t,s,x_1,z_1,b,x_2,z_1}$ y τ_{b,x_2,z_2} son dos puntos intermedios entre $t - s(x_1 \vee z_1) - b(x_2 \vee z_2)$ y t , y entre $-b(x_2 \wedge z_2)$ y 0 respectivamente. Puesto que, por la propiedad 3.2.6, se tiene que

$$\begin{aligned}
\iint (x_1 \vee z_1) K(x_1) K(z_1) \, dx_1 dz_1 &= 2 \iint x_1 K(x_1) K(z_1) \mathbf{1}_{\{x_1 \geq z_1\}} \, dx_1 dz_1 \\
&= 2 \int x_1 K(x_1) \mathbb{K}(x_1) \, dx_1 = 2e_K,
\end{aligned}$$

y

$$\begin{aligned}
\iint (x_2 \wedge z_2) K(x_2) K(z_2) \, dx_2 dz_2 &= 2 \iint z_2 K(z_2) K(x_2) \mathbf{1}_{\{x_2 \geq z_2\}} \, dx_2 dz_2 \\
&= 2 \iint z_2 K(z_2) (1 - \mathbb{K}(z_2)) \, dz_2 = -2e_K,
\end{aligned}$$

donde e_K viene dado por (2.37), entonces el resultado del lema queda demostrado. ■

La siguiente varianza a calcular es $Var(\rho_3(\cdot, Z, \delta))$. Su expresión depende ahora de la relación entre las ventanas de suavizado s y la de presuavizado b .

Lema 3.2.12 *Bajo las condiciones (K1), (H2), (p1) y (f1) se tiene lo siguiente:*

a) Si $\frac{b}{s} = L_n \rightarrow 0$, entonces

$$\text{Var}(\rho_3(t, Z, \delta)) = \frac{1}{s} (1 - F(t))^2 q_1(t) A_K(L_n) + \frac{1}{2} s \left[(1 - F(t))^2 q_1(t) \right]'' m_K + o(s) + o(b).$$

b) Si $\frac{b}{s} = L_n \rightarrow L > 0$, entonces

$$\text{Var}(\rho_3(t, Z, \delta)) = \frac{1}{s} (1 - F(t))^2 q_1(t) A_K(L) + O(s).$$

c) Si $\frac{b}{s} \rightarrow \infty$, entonces

$$\text{Var}(\rho_3(t, Z, \delta)) = \frac{1}{b} (1 - F(t))^2 q_1(t) c_K + O(b),$$

donde

$$A_K(L) = \iiint K(u) K(v) K(w) K(u + L(v - w)) dudvdw, \quad (3.24)$$

la función $q_1(\cdot)$ está dada en (3.23), y las constantes m_K y c_K en (3.15) y (2.7) respectivamente.

Demostración.

Teniendo en cuenta (3.8) y (2.14), la función $\rho_3(\cdot, Z, \delta)$ es igual a

$$\rho_3(t, Z, \delta) = \int K_s(t - v) K_b(v - Z) \frac{1 - F(v)}{1 - H(v)} (\delta - p(v)) dv. \quad (3.25)$$

Como $\mathbb{E}[\rho_3(t, Z, \delta)] = O(b^2)$, la parte dominante de la expresión de la varianza de $\rho_3(t, Z, \delta)$ es

$$\begin{aligned} \mathbb{E}[\rho_3^2(t, Z, \delta)] &= \mathbb{E} \left[\iint K_s(t - v_1) K_s(t - v_2) K_b(v_1 - Z) K_b(v_2 - Z) \right. \\ &\quad \left. \times \frac{1 - F(v_1)}{1 - H(v_1)} \frac{1 - F(v_2)}{1 - H(v_2)} (\delta - p(v_1)) (\delta - p(v_2)) dv_1 dv_2 \right] \\ &= \iiint K_s(t - v_1) K_s(t - v_2) K_b(v_1 - u) K_b(v_2 - u) h(u) \frac{1 - F(v_1)}{1 - H(v_1)} \\ &\quad \times \frac{1 - F(v_2)}{1 - H(v_2)} [p(u) - p(u)(p(v_1) + p(v_2)) + p(v_1)p(v_2)] dv_1 dv_2 du \end{aligned}$$

$$\begin{aligned}
& \stackrel{(a)}{=} \iiint K_s(t-u-bw_1) K_s(t-u-bw_2) K(w_1) K(w_2) h(u) \\
& \quad \times \frac{1-F(u+bw_1)}{1-H(u+bw_1)} \frac{1-F(u+bw_2)}{1-H(u+bw_2)} \\
& \quad \times [p(u) - p(u)(p(u+bw_1) + p(u+bw_2)) + p(u+bw_1)p(u+bw_2)] \\
& \quad \times \mathbf{1}_{\{u+bw_1 \geq 0\}} \mathbf{1}_{\{u+bw_2 \geq 0\}} \mathbf{1}_{\{u \geq 0\}} dw_1 dw_2 du
\end{aligned}$$

donde en el paso (a) realizamos los cambios de variable $v_1 = u + bw_1$ y $v_2 = u + bw_2$.

Debemos ahora distinguir dos posibles situaciones, según el cociente entre las ventanas b/s tienda a un número no negativo, o bien a infinito. En el primer caso, en el que $b/s \rightarrow L \geq 0$, realizamos el cambio de variable $u = t - sz - bw_1$:

$$\begin{aligned}
& \mathbb{E} [\rho_3^2(t, Z, \delta)] \\
& = \frac{1}{s} \iiint K(z) K\left(z + \frac{b}{s}(w_1 - w_2)\right) K(w_1) K(w_2) h(t - sz - bw_1) \\
& \quad \times \frac{1-F(t-sz)}{1-H(t-sz)} \frac{1-F(t-sz-b(w_1-w_2))}{1-H(t-sz-b(w_1-w_2))} \\
& \quad \times [p(t-sz-bw_1) - p(t-sz-bw_1)(p(t-sz) + p(t-sz-b(w_1-w_2))) \\
& \quad + p(t-sz)p(t-sz-b(w_1-w_2))] \\
& \quad \times \mathbf{1}_{\{t-sz \geq 0\}} \mathbf{1}_{\{t-sz-b(w_1-w_2) \geq 0\}} \mathbf{1}_{\{t-sz-bw_1 \geq 0\}} dw_1 dw_2 dz \\
& = \frac{1}{s} \iiint K(z) K\left(z + \frac{b}{s}(w_1 - w_2)\right) K(w_1) K(w_2) \\
& \quad \times m(t-sz, t-sz-bw_1, t-sz-b(w_1-w_2)) \\
& \quad \times \mathbf{1}_{\{t-sz \geq 0\}} \mathbf{1}_{\{t-sz-b(w_1-w_2) \geq 0\}} \mathbf{1}_{\{t-sz-bw_1 \geq 0\}} dw_1 dw_2 dz,
\end{aligned}$$

donde se ha definido la función $m(\cdot, \cdot, \cdot)$ como

$$m(x, y, z) = \frac{1-F(x)}{1-H(x)} \frac{1-F(z)}{1-H(z)} [p(y) - p(y)(p(x) + p(z)) + p(x)p(z)] h(y).$$

Usando la siguiente notación:

$$m_{ijk}(x_1, x_2, x_3) = \frac{\partial^3}{\partial x_i \partial x_j \partial x_k} m(x_1, x_2, x_3)$$

para las correspondientes derivadas parciales, realizamos un desarrollo de Taylor de orden dos sobre la función $m(\cdot, \cdot, \cdot)$ en las tres variables en torno al punto (t, t, t) :

$$\begin{aligned} m(t - sz, t - sz - bw_1, t - sz - b(w_1 - w_2)) &= m(t, t, t) - szm_1(t, t, t) \\ &\quad - (sz + bw_1)m_2(t, t, t) - [sz + b(w_1 - w_2)]m_3(t, t, t) \\ &\quad + \left\{ \frac{1}{2}s^2z^2m_{11} + \frac{1}{2}(sz + bw_1)^2m_{22} \right. \\ &\quad + \frac{1}{2}[sz + b(w_1 - w_2)]^2m_{33} + sz(sz + bw_1)m_{12} \\ &\quad + sz[sz + b(w_1 - w_2)]m_{13} + (sz + bw_1)[sz + b(w_1 - w_2)]m_{23} \left. \right\} \\ &\quad (t - sz\theta, t - (sz + bw_1)\theta, t - [sz + b(w_1 - w_2)]\theta) \end{aligned}$$

donde $0 < \theta < 1$. Si sustituimos este desarrollo en la expresión de $\mathbb{E}[\rho_3^2(t, Z, \delta)]$, para un $t > 0$ fijo y para un n suficientemente grande para que las funciones indicadoras sean uno, ésta se reduce a

$$\begin{aligned} \mathbb{E}[\rho_3^2(t, Z, \delta)] &= \frac{1}{s}m(t, t, t)A_K\left(\frac{b}{s}\right) - [m_1 + m_2 + m_3](t, t, t)A_{K,1}\left(\frac{b}{s}\right) \\ &\quad - \frac{b}{s}m_2(t, t, t)A_{K,2}\left(\frac{b}{s}\right) + b[m_{21} + m_{22} + 3m_{23} + 2m_{33} + 2m_{13}](t, t, t)A_{K,3}\left(\frac{b}{s}\right) \\ &\quad + \frac{b^2}{s}\left[\frac{1}{2}m_{22} + m_{33} + m_{23}\right](t, t, t)A_{K,4}\left(\frac{b}{s}\right) - \frac{b^2}{s}[m_{33} + m_{23}](t, t, t)A_{K,5}\left(\frac{b}{s}\right) \\ &\quad + \frac{1}{2}s[m_{11} + m_{22} + m_{33} + 2m_{12} + 2m_{13} + 2m_{23}](t, t, t)A_{K,6}\left(\frac{b}{s}\right) + o(s) + O\left(\frac{b^3}{s}\right), \end{aligned}$$

donde

$$\begin{aligned} A_K(L) &= \iiint K(u)K(v)K(w)K(u + L(v - w))dudvdw \\ A_{K,1}(L) &= \iiint uK(u)K(v)K(w)K(u + L(v - w))dudvdw \\ A_{K,2}(L) &= \iiint vK(u)K(v)K(w)K(u + L(v - w))dudvdw \\ A_{K,3}(L) &= \iiint uvK(u)K(v)K(w)K(u + L(v - w))dudvdw \\ A_{K,4}(L) &= \iiint v^2K(u)K(v)K(w)K(u + L(v - w))dudvdw \\ A_{K,5}(L) &= \iiint vwK(u)K(v)K(w)K(u + L(v - w))dudvdw \\ A_{K,6}(L) &= \iiint u^2K(u)K(v)K(w)K(u + L(v - w))dudvdw, \end{aligned}$$

y siendo

$$\iiint u(v - w)K(u)K(v)K(w)K(u + L(v - w))dudvdw = 2A_{K,3}(L).$$

En las observaciones que siguen a esta demostración se probará que $A_{K,1}(L) = A_{K,2}(L) = 0$ para todo $L \geq 0$. En el caso en el que $b/s = L_n \rightarrow 0$, realizamos un desarrollo de Taylor en el integrando de las anteriores funciones dependientes de $K(\cdot)$,

$$\begin{aligned} A_{K,3}(L) &= L_n \iint uK(u) K'(u) v^2 K(v) dudv + O(L_n^2) = o(1), \\ A_{K,4}(L) &= c_K d_K + O(L_n) = O(1), \\ A_{K,5}(L) &= O(L_n) = o(1), \\ A_{K,6}(L) &= m_K + O(L_n^2), \end{aligned}$$

ya que

$$\int uK(u) K'(u) du = -\frac{1}{2} \int K^2(u) du = -\frac{1}{2} c_K.$$

Por otra parte,

$$m(t, t, t) = \left(\frac{1 - F(t)}{1 - H(t)} \right)^2 p(t) (1 - p(t)) h(t) = (1 - F(t))^2 q_1(t),$$

donde $q_1(\cdot)$ es la función dada en (3.23). Además,

$$[m_{21} + m_{22} + 3m_{23} + 2m_{33} + 2m_{13}](t, t, t) = \left[(1 - F(t))^2 q_1(t) \right]''$$

y

$$[m_{11} + m_{22} + m_{33} + 2m_{12} + 2m_{13} + 2m_{23}](t, t, t) = \left[(1 - F(t))^2 q_1(t) \right]''.$$

Con esto se demuestra que, si $b/s \rightarrow L > 0$, entonces

$$\text{Var}(\rho_3(t, Z, \delta)) = \frac{1}{s} (1 - F(t))^2 q_1(t) A_K(L) + O(s),$$

mientras que si $b/s = L_n \rightarrow 0$, entonces

$$\begin{aligned} \text{Var}(\rho_3(t, Z, \delta)) &= \frac{1}{s} (1 - F(t))^2 q_1(t) A_K(L_n) \\ &\quad + \frac{1}{2} s \left[(1 - F(t))^2 q_1(t) \right]'' m_K + o(s) + o(b). \end{aligned}$$

Estudiamos ahora el otro caso, en el que $b/s \rightarrow \infty$, es decir, $s/b \rightarrow 0$. Para ello, definimos previamente la siguiente función:

$$n(x, y, z) = \frac{1 - F(x)}{1 - H(x)} \frac{1 - F(y)}{1 - H(y)} [p(z) - p(z)(p(x) + p(y)) + p(x)p(y)] h(z).$$

Los cambios de variable a realizar en la expresión inicial de $\mathbb{E} [\rho_3^2(t, Z, \delta)]$ son $v_1 = t - sw_1$ y $v_2 = t - sw_2$ en el paso (a), y $u = t - sw_1 - bz$ en el paso (b). Finalmente, en el paso (c) sustituimos parte del integrando por la función $n(\cdot, \cdot, \cdot)$ definida previamente:

$$\begin{aligned}
\mathbb{E} [\rho_3^2(t, Z, \delta)] &\stackrel{(a)}{=} \iiint K(w_1) K(w_2) K_b(t - sw_1 - u) K_b(t - sw_2 - u) \\
&\quad \times h(u) \frac{1 - F(t - sw_1)}{1 - H(t - sw_1)} \frac{1 - F(t - sw_2)}{1 - H(t - sw_2)} \\
&\quad \times [p(u) - p(u)(p(t - sw_1) + p(t - sw_2)) + p(t - sw_1)p(t - sw_2)] \\
&\quad \times \mathbf{1}_{\{t - sw_1 \geq 0\}} \mathbf{1}_{\{t - sw_2 \geq 0\}} \mathbf{1}_{\{u \geq 0\}} dw_1 dw_2 du \\
&\stackrel{(b)}{=} \frac{1}{b} \iiint K(w_1) K(w_2) K(z) K\left(z + \frac{s}{b}(w_1 - w_2)\right) \\
&\quad \times h(t - sw_1 - bz) \frac{1 - F(t - sw_1)}{1 - H(t - sw_1)} \frac{1 - F(t - sw_2)}{1 - H(t - sw_2)} \\
&\quad \times [p(t - sw_1 - bz) - p(t - sw_1 - bz)(p(t - sw_1) + p(t - sw_2)) \\
&\quad + p(t - sw_1)p(t - sw_2)] \mathbf{1}_{\{t - sw_1 \geq 0\}} \mathbf{1}_{\{t - sw_2 \geq 0\}} \mathbf{1}_{\{t - sw_1 - bz \geq 0\}} dw_1 dw_2 dz \\
&\stackrel{(c)}{=} \frac{1}{b} \iiint K(w_1) K(w_2) K(z) K\left(z + \frac{s}{b}(w_1 - w_2)\right) \\
&\quad \times n(t - sw_1, t - sw_2, t - sw_1 - bz) \\
&\quad \times \mathbf{1}_{\{t - sw_1 \geq 0\}} \mathbf{1}_{\{t - sw_2 \geq 0\}} \mathbf{1}_{\{t - sw_1 - bz \geq 0\}} dw_1 dw_2 dz.
\end{aligned}$$

Realizamos nuevamente un desarrollo de Taylor de grado dos en la función $n(t - sw_1, t - sw_2, t - sw_1 - bz)$ en torno al punto (t, t, t) , lo insertamos en la expresión de $\mathbb{E} [\rho_3^2(t, Z, \delta)]$, simplificamos términos y obtenemos lo siguiente:

$$\begin{aligned}
\mathbb{E} [\rho_3^2(t, Z, \delta)] &= \frac{1}{b} n(t, t, t) A_K\left(\frac{s}{b}\right) - \frac{s}{b} [n_1 + n_2 + n_3](t, t, t) A_{K,2}\left(\frac{s}{b}\right) \\
&\quad - n_3(t, t, t) A_{K,1}\left(\frac{s}{b}\right) + O(s) + O(b) + O\left(\frac{s^2}{b}\right).
\end{aligned}$$

Como la función $n(\cdot, \cdot, \cdot)$ verifica que $n(t, t, t) = (1 - F(t))^2 q_1(t)$, entonces

$$\text{Var}(\rho_3(t, Z, \delta)) = \frac{1}{b} (1 - F(t))^2 q_1(t) c_K + O(s) + O(b),$$

dando por concluida la demostración. ■

Observación 3.2.1

Las funciones $A_{K,1}(L)$ y $A_{K,2}(L)$ son iguales a cero para cualquier valor de $L \geq 0$, ya que ambas son integrales de la forma

$$\iiint u\varphi(u, v, w) \, dudvdw$$

con $\varphi(-u, v, w) = \varphi(u, -v, -w)$, y por tanto

$$\begin{aligned} & \iiint_{\mathbb{R} \times \mathbb{R} \times \mathbb{R}} u\varphi(u, v, w) \, dudvdw \\ &= \iiint_{\mathbb{R}^- \times \mathbb{R} \times \mathbb{R}} u\varphi(u, v, w) \, dudvdw + \iiint_{\mathbb{R}^+ \times \mathbb{R} \times \mathbb{R}} u\varphi(u, v, w) \, dudvdw \\ &= \iiint_{\mathbb{R}^+ \times \mathbb{R} \times \mathbb{R}} -u\varphi(-u, v, w) \, dudvdw + \iiint_{\mathbb{R}^+ \times \mathbb{R} \times \mathbb{R}} u\varphi(u, v, w) \, dudvdw \\ &= -\iiint_{\mathbb{R}^+ \times \mathbb{R} \times \mathbb{R}} u\varphi(u, -v, -w) \, dudvdw + \iiint_{\mathbb{R}^+ \times \mathbb{R} \times \mathbb{R}} u\varphi(u, v, w) \, dudvdw \\ &= -\iiint_{\mathbb{R}^+ \times \mathbb{R} \times \mathbb{R}} u\varphi(u, v, w) \, dudvdw + \iiint_{\mathbb{R}^+ \times \mathbb{R} \times \mathbb{R}} u\varphi(u, v, w) \, dudvdw = 0. \end{aligned}$$

Observación 3.2.2

La observación anterior se puede demostrar también usando propiedades de simetría de la función núcleo $K(\cdot)$ y de la convolución de funciones. Si el parámetro L de las funciones $A_{K,1}(L)$ y $A_{K,2}(L)$ es cero, entonces

$$A_{K,1}(0) = \int uK^2(u) \, du = 0 \text{ y } A_{K,2}(0) = \int vK(v) \, dv \int K^2(u) \, du = 0.$$

En caso contrario, ambas funciones son igualmente cero para cualquier valor $L > 0$, ya que

$$\begin{aligned} A_{K,1}(L) &= \iiint uK(u)K(v)K(w)K(u+L(v-w)) \, dudvdw \\ &= \iiint uK(u)K_L(v)K_L(w)K(u+v-w) \, dudvdw \\ &= \iint uK(u)K_L(w)(K_L * K)(w-u) \, dudw \\ &= \int uK(u)(K_L * K_L * K)(u) \, du = 0, \end{aligned}$$

donde $K_L(\cdot) = \frac{1}{L}K(\frac{\cdot}{L})$ y $*$ denota el operador convolución. La demostración es análoga para la función $A_{K,2}(\cdot)$.

Lema 3.2.13 *Bajo las condiciones (K1), (H2), (p1) y (f1) se tiene lo siguiente:*

a) Si $\frac{b}{s} = L_n \rightarrow 0$, entonces

$$\text{Cov}(\rho_1(t, Z), \rho_3(t, Z, \delta)) = o(s) + o(b).$$

b) Si $\frac{b}{s} \rightarrow L > 0$ o bien $\frac{b}{s} \rightarrow \infty$, entonces

$$\text{Cov}(\rho_1(t, Z), \rho_3(t, Z, \delta)) = O(b).$$

Demostración.

La esperanza de $\rho_1(\cdot, Z)$ es cero, así que la covarianza entre esta función y $\rho_3(\cdot, Z, \delta)$ es igual a

$$\begin{aligned} & \text{Cov}(\rho_1(t, Z), \rho_3(t, Z, \delta)) \\ &= \mathbb{E}[\rho_1(t, Z) \rho_3(t, Z, \delta)] \\ &= \mathbb{E} \left[\left(K_s(t - Z) \frac{1 - F(Z)}{1 - H(Z)} p(Z) \right. \right. \\ & \quad + \int K_s(t - v_1) (1 - F(v_1)) \left(\frac{p(v_1)}{1 - H(v_1)} \right)' \mathbf{1}_{\{Z \leq v_1\}} dv_1 \\ & \quad - \int K_s(t - v_2) (1 - F(v_2)) \left(\frac{p(v_2) H(v_2)}{1 - H(v_2)} \right)' dv_2 \\ & \quad \left. \times \int K_s(t - v_3) K_b(v_3 - Z) \frac{1 - F(v_3)}{1 - H(v_3)} (\delta - p(v_3)) dv_3 \right] \\ &= \mathbb{E}[I_1 + I_2 - I_3]. \end{aligned}$$

La tercera esperanza es igual a

$$\mathbb{E}[I_3] = \int K_s(t - v_2) (1 - F(v_2)) \left(\frac{p(v_2) H(v_2)}{1 - H(v_2)} \right)' dv_2 \times \mathbb{E}[\rho_3(t, Z, \delta)].$$

Dado que la primera integral está acotada, y que $\mathbb{E}[\rho_3(t, Z, \delta)] = O(b^2)$, entonces $\mathbb{E}[I_3]$ es despreciable. Basta entonces con calcular la esperanza de I_1 y I_2 . La esperanza I_1 es

$$\begin{aligned} \mathbb{E}[I_1] &= \iint K_s(t - u) K_s(t - v_3) K_b(v_3 - u) \frac{1 - F(u)}{1 - H(u)} p(u) \\ & \quad \times \frac{1 - F(v_3)}{1 - H(v_3)} (p(u) - p(v_3)) h(u) dv_3 du. \end{aligned}$$

Ahora debemos distinguir dos posibles situaciones, cuando el cociente entre las ventanas b/s tiende a un valor no negativo, y cuando tiende a infinito. En el primero de los casos, cuando $b/s \rightarrow L \geq 0$, definimos previamente la función $m(\cdot, \cdot)$ de la siguiente manera:

$$m(x, y) = \frac{1 - F(x)}{1 - H(x)} p(x) h(x) \frac{1 - F(y)}{1 - H(y)} (p(x) - p(y)),$$

y realizamos los cambios de variable $v_3 = u + bw_3$ y $u = t - su_1$:

$$\begin{aligned} \mathbb{E}[I_1] &= \frac{1}{s} \iint K(u_1) K(w_3) K\left(u_1 - \frac{b}{s}w_3\right) \frac{1 - F(t - su_1)}{1 - H(t - su_1)} p(t - su_1) \\ &\quad \times \frac{1 - F(t - su_1 + bw_3)}{1 - H(t - su_1 + bw_3)} [p(t - su_1) - p(t - su_1 + bw_3)] h(t - su_1) \\ &\quad \times \mathbf{1}_{\{t - su_1 \geq 0\}} \mathbf{1}_{\{t - su_1 + bw_3 \geq 0\}} dw_3 du_1 \\ &= \frac{1}{s} \iint K(u_1) K(w_3) K\left(u_1 - \frac{b}{s}w_3\right) m(t - su_1, t - su_1 + bw_3) \\ &\quad \times \mathbf{1}_{\{t - su_1 \geq 0\}} \mathbf{1}_{\{t - su_1 + bw_3 \geq 0\}} dw_3 du_1. \end{aligned}$$

Si realizamos un desarrollo de Taylor de grado dos en la función $m(t - su_1, t - su_1 + bw_3)$ en torno al punto (t, t) , para un $t > 0$ fijo y para un tamaño muestral suficientemente grande para que las anteriores funciones indicadoras valgan uno, entonces la esperanza de I_1 es igual a

$$\begin{aligned} \mathbb{E}[I_1] &= \frac{1}{s} m(t, t) \iint K(u_1) K(w_3) K\left(u_1 - \frac{b}{s}w_3\right) dw_3 du_1 \\ &\quad - [m_1 + m_2](t, t) \iint u_1 K(u_1) K(w_3) K\left(u_1 - \frac{b}{s}w_3\right) dw_3 du_1 \\ &\quad - \frac{1}{2} s [m_{11} + m_{22} + 2m_{12}](t, t) \iint u_1^2 K(u_1) K(w_3) K\left(u_1 - \frac{b}{s}w_3\right) dw_3 du_1 \\ &\quad + \frac{b}{s} m_2(t, t) \iint w_3 K(u_1) K(w_3) K\left(u_1 - \frac{b}{s}w_3\right) dw_3 du_1 \\ &\quad - b [m_{21} + m_{22}](t, t) \iint u_1 w_3 K(u_1) K(w_3) K\left(u_1 - \frac{b}{s}w_3\right) dw_3 du_1 \\ &\quad + \frac{1}{2} \frac{b^2}{s} m_{22}(t, t) \iint w_3^2 K(u_1) K(w_3) K\left(u_1 - \frac{b}{s}w_3\right) dw_3 du_1 \\ &\quad + O(s^2) + O\left(\frac{b^3}{s}\right). \end{aligned}$$

Pero, puesto que

$$\begin{aligned} m(t, t) &= 0, \quad m_1(t, t) = -m_2(t, t), \\ m_{11}(t, t) + m_{22}(t, t) + 2m_{12}(t, t) &= 0, \\ [m_{21} + m_{22}](t, t) &= - \left[\left(\frac{1 - F(t)}{1 - H(t)} \right)^2 h(t) p(t) p'(t) \right]', \end{aligned} \quad (3.26)$$

y que

$$\iint wK(u)K(u-Lw)K(w)dwdu = 0 \text{ para todo } L \geq 0,$$

entonces

$$\mathbb{E}[I_1] = b \left[\left(\frac{1 - F(t)}{1 - H(t)} \right)^2 p(t) p'(t) h(t) \right]' B_K \left(\frac{b}{s} \right) + o(s) + o(b),$$

donde

$$B_K(x) = \iint uvK(u)K(v)K(u-xv)dvdu.$$

En el caso particular en el que $b/s \rightarrow L > 0$, entonces $\mathbb{E}[I_1] = O(b)$, mientras que si el cociente entre las ventanas es $b/s = L_n \rightarrow 0$ entonces

$$B_K(L_n) = \frac{1}{2}L_n d_K c_K + O(L_n^2) = o(1)$$

y por tanto $\mathbb{E}[I_1] = o(s) + o(b)$.

En cambio, si el cociente entre las ventanas verifica $b/s \rightarrow \infty$, es decir, $s/b \rightarrow 0$, los cambios de variable que realizaremos en la esperanza de I_1 son $u = t - su_1$ y $v_3 = t - sw_3$:

$$\begin{aligned} \mathbb{E}[I_1] &= \frac{1}{b} \iint K(u_1)K(w_3)K\left(\frac{s}{b}(u_1 - w_3)\right) \frac{1 - F(t - su_1)}{1 - H(t - su_1)} p(t - su_1) \\ &\quad \times \frac{1 - F(t - sw_3)}{1 - H(t - sw_3)} (p(t - su_1) - p(t - sw_3)) h(t - su_1) dw_3 du_1 \\ &= \frac{1}{b} \iint K(u_1)K(w_3)K\left(\frac{s}{b}(u_1 - w_3)\right) m(t - su_1, t - sw_3) dw_3 du_1 \end{aligned}$$

Si realizamos un desarrollo de Taylor de grado dos en la función $m(t - su_1, t - sw_3)$ en torno al punto (t, t) , y teniendo en cuenta (3.26), entonces

$$\begin{aligned} \mathbb{E}[I_1] &= \frac{1}{b} m(t, t) \iint K(u_1)K(w_3)K\left(\frac{s}{b}(u_1 - w_3)\right) dw_3 du_1 \\ &\quad - \frac{s}{b} [m_1 + m_2](t, t) \iint u_1 K(u_1)K(w_3)K\left(\frac{s}{b}(u_1 - w_3)\right) dw_3 du_1 + O\left(\frac{s^2}{b}\right) = o(s). \end{aligned}$$

Con respecto a la esperanza del segundo término I_2 , definimos ahora la función $n(\cdot, \cdot, \cdot)$ de la siguiente manera:

$$n(x, y, z) = (1 - F(x)) \left(\frac{p(x)}{1 - H(x)} \right)' \frac{1 - F(y)}{1 - H(y)} (p(z) - p(y)) h(z).$$

Realizamos en primer lugar (paso (a)) los cambios de variable $v_1 = t - sw_1$, $v_3 = t - sw_3$ y $u = v_3 - bu_1$, y a continuación sustituimos parte del integrando por la función $n(\cdot, \cdot, \cdot)$ evaluada en el punto $(t - sw_1, t - sw_3, t - sw_3 - bu_1)$ (paso (b)):

$$\begin{aligned} \mathbb{E}[I_2] &= \iiint K_s(t - v_1) K_s(t - v_3) K_b(v_3 - u) (1 - F(v_1)) \left(\frac{p(v_1)}{1 - H(v_1)} \right)' \\ &\quad \times \frac{1 - F(v_3)}{1 - H(v_3)} (p(u) - p(v_3)) h(u) \mathbf{1}_{\{u \leq v_1\}} dv_1 dv_3 du \\ &\stackrel{(a)}{=} \iiint K(w_1) K(w_3) K(u_1) (1 - F(t - sw_1)) \left(\frac{p(t - sw_1)}{1 - H(t - sw_1)} \right)' \\ &\quad \times \frac{1 - F(t - sw_3)}{1 - H(t - sw_3)} (p(t - sw_3 - bu_1) - p(t - sw_3)) h(t - sw_3 - bu_1) \\ &\quad \times \mathbf{1}_{\{t - sw_1 \geq 0\}} \mathbf{1}_{\{t - sw_3 \geq 0\}} \mathbf{1}_{\{t - sw_3 - bu_1 \geq 0\}} \mathbf{1}_{\{t - sw_3 - bu_1 \leq t - sw_1\}} dw_1 dw_3 du_1 \\ &\stackrel{(b)}{=} \iiint K(w_1) K(w_3) K(u_1) n(t - sw_1, t - sw_3, t - sw_3 - bu_1) \\ &\quad \times \mathbf{1}_{\{t - sw_1 \geq 0\}} \mathbf{1}_{\{t - sw_3 \geq 0\}} \mathbf{1}_{\{t - sw_3 - bu_1 \geq 0\}} \mathbf{1}_{\{s(w_1 - w_3) \leq bu_1\}} dw_1 dw_3 du_1. \end{aligned}$$

La función $n(\cdot, \cdot, \cdot)$ verifica lo siguiente:

$$\begin{aligned} n(t, t, t) &= n_1(t, t, t) = 0 \\ n_2(t, t, t) + n_3(t, t, t) &= 0. \end{aligned} \tag{3.27}$$

Si el cociente entre las ventanas es $b/s = L_n \rightarrow L \geq 0$, entonces para un $t > 0$ fijo y para un tamaño muestral suficientemente grande para que las tres primeras funciones indicadoras sean idénticamente uno, y teniendo en cuenta (3.27), mediante un desarrollo de Taylor podemos concluir que

$$\mathbb{E}[I_2] = -bn_3(t, t, t) C_K \left(\frac{b}{s} \right) + O(s^2) + o(b),$$

donde

$$C_K(x) = \iint uK(u) K(v) \mathbb{K}(v + xu) d\omega du. \tag{3.28}$$

Así, si $b/s \rightarrow L > 0$, entonces $\mathbb{E}[I_2] = O(b)$ mientras que si $b/s = L_n \rightarrow 0$ entonces $C_K(L_n) = L_n c_K d_K + O(L_n^2) = o(1)$ de modo que

$$\mathbb{E}[I_2] = o(s) + o(b).$$

Si el cociente entre las ventanas es $b/s \rightarrow \infty$, entonces para un $t > 0$ fijo y para un tamaño muestral suficientemente grande para que las tres primeras funciones indicadoras sean idénticamente uno, y teniendo en cuenta (3.27), tenemos que

$$\begin{aligned} \mathbb{E}[I_2] &= bn_3(t, t, t) \iiint u_1 K(w_1) K(w_3) K(u_1) \mathbf{1}_{\{u_1 \leq \frac{s}{b}(w_1 - w_3)\}} dw_1 dw_3 du_1 \\ &\quad + O(s^2) + o(b) = O(b) \end{aligned}$$

Se concluye así la demostración del lema. ■

Lema 3.2.14 *Bajo las condiciones (K1), (H2), (p1) y (f1) se tiene lo siguiente:*

a) Si $\frac{b}{s} = L_n \rightarrow 0$, entonces

$$\text{Cov}(\rho_1(t, Z), \omega_3(t, Z, \delta)) = o(s) + o(b).$$

b) Si $\frac{b}{s} \rightarrow L > 0$ o bien $\frac{b}{s} \rightarrow \infty$, entonces

$$\text{Cov}(\rho_1(t, Z), \omega_3(t, Z, \delta)) = O(b).$$

Demostración.

La expresión de esta covarianza, $\text{Cov}(\rho_1(t, Z), \omega_3(t, Z, \delta))$, se obtiene fácilmente, pues la esperanza de la función $\rho_1(\cdot, Z)$ es cero, de modo que

$$\begin{aligned} \text{Cov}(\rho_1(t, Z), \omega_3(t, Z, \delta)) &= \mathbb{E}[\rho_1(t, Z) \omega_3(t, Z, \delta)] \\ &= \mathbb{E} \left[\left(K_s(t - Z) \frac{1 - F(Z)}{1 - H(Z)} p(Z) \right. \right. \\ &\quad \left. \left. + \int K_s(t - v_1) (1 - F(v_1)) \left(\frac{p(v_1)}{1 - H(v_1)} \right)' \mathbf{1}_{\{Z \leq v_1\}} dv_1 \right. \right. \\ &\quad \left. \left. - \int K_s(t - v_2) (1 - F(v_2)) \left(\frac{p(v_2) H(v_2)}{1 - H(v_2)} \right)' dv_2 \right) \right. \\ &\quad \left. \times \int K_s(t - v_3) f(v_3) \int_0^{v_3} K_b(v_4 - Z) \frac{\delta - p(v_4)}{1 - H(v_4)} dv_4 dv_3 \right] \\ &= \mathbb{E}[I_1] + \mathbb{E}[I_2] - \mathbb{E}[I_3]. \end{aligned}$$

La covarianza entre $\rho_1(\cdot, Z)$ y $\omega_3(\cdot, Z, \delta)$ es, por tanto, la suma de tres términos, el tercero de los cuales es

$$\mathbb{E}[I_3] = \int K_s(t - v_2) (1 - F(v_2)) \left(\frac{p(v_2) H(v_2)}{1 - H(v_2)} \right)' dv_2 \times \mathbb{E}[\omega_3(t, Z, \delta)].$$

Dado que la primera integral es de orden $O(1)$, y que $\mathbb{E}[\omega_3(t, Z, \delta)] = O(b^2)$, entonces este tercer sumando es despreciable. Basta entonces con calcular la esperanza de los dos términos restantes.

La esperanza de I_1 es

$$\begin{aligned} \mathbb{E}[I_1] &= \iiint K_s(t - u) K_s(t - v_3) K_b(v_4 - u) \frac{1 - F(u)}{1 - H(u)} p(u) \\ &\quad \times f(v_3) \frac{p(u) - p(v_4)}{1 - H(v_4)} h(u) \mathbf{1}_{\{0 \leq v_4 \leq v_3\}} dv_4 dv_3 du. \end{aligned}$$

Definimos ahora la función $m(\cdot, \cdot, \cdot)$ de la siguiente manera:

$$m(x, y, z) = \frac{1 - F(x)}{1 - H(x)} p(x) h(x) f(y) \frac{p(x) - p(z)}{1 - H(z)}.$$

Entonces, el primer paso en el cálculo de $\mathbb{E}[I_1]$ consiste en realizar los siguientes cambios de variable: $u = t - su_1$, $v_3 = t - sw_3$ y $v_4 = u + bw_4$ (paso (a)) y en la sustitución de parte del integrando por la función $m(\cdot, \cdot, \cdot)$ (paso (b)):

$$\begin{aligned} \mathbb{E}[I_1] &\stackrel{(a)}{=} \iiint K(u_1) K(w_3) K(w_4) \frac{1 - F(t - su_1)}{1 - H(t - su_1)} p(t - su_1) \\ &\quad \times f(t - sw_3) \frac{p(t - su_1) - p(t - su_1 + bw_4)}{1 - H(t - su_1 + bw_4)} h(t - su_1) \\ &\quad \times \mathbf{1}_{\{t - sw_3 \geq 0\}} \mathbf{1}_{\{t - su_1 \geq 0\}} \mathbf{1}_{\{0 \leq t - su_1 + bw_4 \leq t - sw_3\}} dw_4 dw_3 du_1 \\ &\stackrel{(b)}{=} \iiint K(u_1) K(w_3) K(w_4) m(t - su_1, t - sw_3, t - su_1 + bw_4) \\ &\quad \times \mathbf{1}_{\{t - sw_3 \geq 0\}} \mathbf{1}_{\{t - su_1 \geq 0\}} \mathbf{1}_{\{t - su_1 + bw_4 \geq 0\}} \mathbf{1}_{\{s(u_1 - w_3) \geq bw_4\}} dw_4 dw_3 du_1. \end{aligned}$$

Teniendo en cuenta que

$$\begin{aligned} m(t, t, t) &= m_2(t, t, t) = m_1(t, t, t) + m_3(t, t, t) = 0, \\ m_3(t, t, t) &= \frac{1 - F(t)}{(1 - H(t))^2} p(t) p'(t) h(t) f(t), \end{aligned}$$

entonces, si el cociente entre las ventanas verifica que $b/s = L_n \rightarrow L \geq 0$, la esperanza de I_1 es igual a

$$\mathbb{E}[I_1] = bm_3(t, t, t) C_K(L_n) + o(s) + o(b),$$

donde $C_K(\cdot)$ es la función dada en (3.28). Así, si $b/s \rightarrow 0$ entonces $\mathbb{E}[I_1] = o(s) + o(b)$, mientras que si $b/s \rightarrow L > 0$ entonces $\mathbb{E}[I_1] = O(b)$.

En caso de que las ventanas verifiquen $b/s \rightarrow \infty$, la esperanza de I_1 es igual a

$$\mathbb{E}[I_1] = bm_3(t, t, t) \iiint w_4 K(u_1) K(w_3) K(w_4) \mathbf{1}_{\{w_4 \leq \frac{s}{b}(u_1 - w_3)\}} dw_4 dw_3 du_1 + o(s) + o(b),$$

y por tanto $\mathbb{E}[I_1] = O(b)$.

La misma idea se aplicará a la segunda esperanza $\mathbb{E}[I_2]$:

$$\begin{aligned} \mathbb{E}[I_2] &= \iiint K_s(t - v_1) K_s(t - v_3) K_b(v_4 - u) (1 - F(v_1)) \left(\frac{p(v_1)}{1 - H(v_1)} \right)' \\ &\quad \times f(v_3) \frac{p(u) - p(v_4)}{1 - H(v_4)} h(u) \mathbf{1}_{\{0 \leq u \leq v_1\}} \mathbf{1}_{\{0 \leq v_4 \leq v_3\}} dv_4 dv_3 dv_1 du. \end{aligned}$$

Definimos ahora la función $n(\cdot, \cdot, \cdot, \cdot)$:

$$n(x, y, z, t) = (1 - F(x)) \left(\frac{p(x)}{1 - H(x)} \right)' f(y) \frac{p(t) - p(z)}{1 - H(z)} h(t).$$

Los pasos en el cálculo de $\mathbb{E}[I_2]$ son los cambios de variable $v_1 = t - sw_1$, $v_3 = t - sw_3$ y $v_4 = u + bw_4$ (paso (a)), el reemplazo de parte del integrando por la función $n(\cdot, \cdot, \cdot, \cdot)$ y un desarrollo de Taylor de orden 2 en dicha función (paso (c)):

$$\begin{aligned} \mathbb{E}[I_2] &\stackrel{(a)}{=} \iiint K(w_1) K(w_3) K(w_4) (1 - F(t - sw_1)) \left(\frac{p(t - sw_1)}{1 - H(t - sw_1)} \right)' h(u) \\ &\quad \times f(t - sw_3) \frac{p(u) - p(u + bw_4)}{1 - H(u + bw_4)} \mathbf{1}_{\{0 \leq u \leq t - sw_1\}} \mathbf{1}_{\{0 \leq u + bw_4 \leq t - sw_3\}} dw_4 dw_3 dw_1 du \\ &\stackrel{(b)}{=} \iiint K(w_1) K(w_3) K(w_4) n(t - sw_1, t - sw_3, u + bw_4, u) \\ &\quad \times \mathbf{1}_{\{-bw_4 \leq u \leq t - sw_1\}} \mathbf{1}_{\{u \leq t - sw_3 - bw_4\}} dw_4 dw_3 dw_1 du \\ &\stackrel{(c)}{=} b \int_0^t n_3(t, t, u, u) du \iiint w_4 K(w_1) K(w_3) K(w_4) dw_4 dw_3 dw_1 + O(s^2) + o(b) \\ &= O(s^2) + o(b), \end{aligned}$$

puesto que $n(t, t, u, u) = n_1(t, t, u, u) = n_2(t, t, u, u) = 0$. Con esto se concluye la demostración de este lema. ■

Lema 3.2.15 *Bajo las condiciones (K1), (H2), (p1) y (f1) se tiene lo siguiente:*

a) Si $\frac{b}{s} = L_n \rightarrow 0$, entonces

$$\text{Cov}(\rho_2(t, Z), \rho_3(t, Z, \delta)) = o(s) + o(b).$$

b) Si $\frac{b}{s} \rightarrow L > 0$ o bien $\frac{b}{s} \rightarrow \infty$, entonces

$$\text{Cov}(\rho_2(t, Z), \rho_3(t, Z, \delta)) = O(b).$$

Demostración.

Al ser la esperanza de $\rho_2(\cdot, Z)$ cero, la covarianza entre $\rho_2(\cdot, Z)$ y $\rho_3(\cdot, Z, \delta)$ es igual a la esperanza del producto de ambas funciones. Como además el último sumando en la expresión de $\rho_2(\cdot, Z)$ es de orden $O(1)$, y la esperanza de $\rho_3(\cdot, Z, \delta)$ es de orden $O(b^2)$, el término dominante de la covarianza entre $\rho_2(\cdot, Z)$ y $\rho_3(\cdot, Z, \delta)$ es el siguiente:

$$\begin{aligned} \text{Cov}(\rho_2(t, Z), \rho_3(t, Z, \delta)) &= \mathbb{E}[\rho_2(t, Z) \rho_3(t, Z, \delta)] \\ &= \mathbb{E} \left[\int K_s(t - v_1) \frac{1 - F(v_1)}{1 - H(v_1)} p'(v_1) \mathbf{1}_{\{Z \leq v_1\}} dv_1 \right. \\ &\quad \left. \times \int K_s(t - v_2) K_b(v_2 - Z) \frac{1 - F(v_2)}{1 - H(v_2)} (\delta - p(v_2)) dv_2 \right] + O(b^2) = \mathbb{E}[I_1] + O(b^2). \end{aligned}$$

Definimos ahora la función $m(\cdot, \cdot, \cdot)$ como

$$m(x, y, z) = \frac{1 - F(x)}{1 - H(x)} p'(x) \frac{1 - F(y)}{1 - H(y)} (p(z) - p(y)) h(z).$$

Para calcular la esperanza que representa el término dominante $\text{Cov}(\rho_2(t, Z), \rho_3(t, Z, \delta))$, tomamos la esperanza de I_1 (paso (a)), realizamos los cambios de variable $v_1 = t - sw_1$, $v_2 = t - sw_2$ y $u = v_2 - bu_1$ y sustituimos el integrando por la función $m(\cdot, \cdot, \cdot)$ (paso (b)):

$$\begin{aligned} \mathbb{E}[I_1] &\stackrel{(a)}{=} \iiint K_s(t - v_1) K_s(t - v_2) K_b(v_2 - u) \frac{1 - F(v_1)}{1 - H(v_1)} p'(v_1) \\ &\quad \times \frac{1 - F(v_2)}{1 - H(v_2)} (p(u) - p(v_2)) h(u) \mathbf{1}_{\{u \leq v_1\}} dv_1 dv_2 du \\ &\stackrel{(b)}{=} \iiint K(w_1) K(w_2) K(u_1) m(t - sw_1, t - sw_2, t - sw_2 - bu_1) \\ &\quad \times \mathbf{1}_{\{0 \leq t - sw_1\}} \mathbf{1}_{\{0 \leq t - sw_2\}} \mathbf{1}_{\{0 \leq t - sw_2 - bu_1\}} \mathbf{1}_{\{s(w_1 - w_2) \leq bu_1\}} dw_2 dw_1 du_1, \end{aligned}$$

Un desarrollo de Taylor de $m(t - sw_1, t - sw_2, t - sw_2 - bu_1)$ en torno al punto (t, t, t) bajo las hipótesis (H2), (p1) y (f1), y dado que

$$m(t, t, t) = m_1(t, t, t) = m_2(t, t, t) + m_3(t, t, t) = 0,$$

podemos concluir que si el cociente entre las ventanas es $b/s = L_n \rightarrow L \geq 0$ entonces

$$\mathbb{E}[I_1] = -bm_3(t, t, t)C_K(L_n) + o(s) + o(b).$$

En el caso particular en que $b/s \rightarrow 0$ entonces $\mathbb{E}[I_1] = o(s) + o(b)$, mientras que si $b/s \rightarrow L > 0$ entonces $\mathbb{E}[I_1] = O(b)$.

Cuando $b/s \rightarrow \infty$ entonces la esperanza de I_1 es

$$\mathbb{E}[I_1] = bm_3(t, t, t) \iiint u_1 K(w_1) K(w_2) K(u_1) \mathbf{1}_{\{u_1 \leq \frac{s}{b}(w_1 - w_2)\}} dw_2 dw_1 du_1 + o(s) + o(b)$$

y, por tanto, $\mathbb{E}[I_1] = O(b)$. De este modo queda demostrado el lema. ■

Lema 3.2.16 *Bajo las condiciones (K1), (H2), (p1) y (f1) se tiene*

$$\text{Cov}(\omega_3(t, Z, \delta), \rho_2(t, Z)) = o(s) + o(b).$$

Demostración.

Como para la covarianza anterior, al ser la esperanza de $\rho_2(\cdot, Z)$ cero, la covarianza entre $\rho_2(\cdot, Z)$ y $\omega_3(\cdot, Z, \delta)$ es igual a la esperanza del producto de ambas funciones. Como además el último sumando en la expresión de $\rho_2(\cdot, Z)$ de orden $O(1)$, y la esperanza de $\omega_3(\cdot, Z, \delta)$ es de orden $O(b^2)$, los términos dominantes de la covarianza entre $\rho_2(\cdot, Z)$ y $\omega_3(\cdot, Z, \delta)$ son los siguientes:

$$\begin{aligned} \text{Cov}(\omega_3(t, Z, \delta), \rho_2(t, Z)) &= \mathbb{E}[\omega_3(t, Z, \delta) \rho_2(t, Z)] \\ &= \mathbb{E} \left[\int K_s(t - v_1) f(v_1) \int_0^{v_1} K_b(v_2 - Z) \frac{\delta - p(v_2)}{1 - H(v_2)} dv_2 dv_1 \right. \\ &\quad \left. \times \int K_s(t - v_3) \frac{1 - F(v_3)}{1 - H(v_3)} p'(v_3) \mathbf{1}_{\{Z \leq v_3\}} dv_3 \right] + O(b^2) = \mathbb{E}[I_1] + O(b^2). \end{aligned}$$

Definamos la función $m(\cdot, \cdot, \cdot, \cdot)$ como

$$m(x, y, z, t) = f(x) \frac{1 - F(y)}{1 - H(y)} p'(y) \frac{p(t) - p(z)}{1 - H(z)} h(t).$$

En el primer paso del cálculo de $\mathbb{E}[I_1]$ llevamos a cabo los cambios de variable $v_1 = t - sw_1$, $v_2 = u + bw_2$ y $v_3 = t - sw_3$ y sustituimos el integrando resultante por la función $m(\cdot, \cdot, \cdot, \cdot)$:

$$\begin{aligned}
\mathbb{E}[I_1] &= \iiint\iiint K_s(t-v_1) K_s(t-v_3) K_b(v_2-u) f(v_1) \frac{p(u)-p(v_2)}{1-H(v_2)} \\
&\quad \times \frac{1-F(v_3)}{1-H(v_3)} p'(v_3) h(u) \mathbf{1}_{\{0 \leq v_2 \leq v_1\}} \mathbf{1}_{\{0 \leq u \leq v_3\}} dudv_3 dv_2 dv_1 \\
&= \iiint\iiint K(w_1) K(\omega_3) K(w_2) m(t-sw_1, t-sw_3, u+bw_2, u) \\
&\quad \times \mathbf{1}_{\{0 \leq u+bw_2 \leq t-sw_1\}} \mathbf{1}_{\{0 \leq u \leq t-s\omega_3\}} dw_1 dw_2 dw_3 du.
\end{aligned}$$

Mediante un desarrollo de Taylor de $m(t-sw_1, t-sw_3, u+bw_2, u)$ en torno al punto (t, t, u, u) , y teniendo en cuenta que

$$m(t, t, u, u) = m_1(t, t, u, u) = m_2(t, t, u, u) = 0,$$

la esperanza anterior se puede escribir como

$$\begin{aligned}
\mathbb{E}[I_1] &= b \iiint\iiint w_2 K(w_1) K(w_2) K(w_3) \\
&\quad \times \int m_3(t, t, u, u) \mathbf{1}_{\{-bw_2 \leq u \leq t-sw_1-bw_2\}} \mathbf{1}_{\{0 \leq u \leq t-s\omega_3\}} dudw_3 dw_2 dw_1 = O(s^2) + o(b).
\end{aligned}$$

Se ha demostrado por tanto lo establecido en el lema. ■

Lema 3.2.17 *Bajo las condiciones (K1), (H2), (p1) y (f1) se tiene lo siguiente:*

a) Si $\frac{b}{s} = L_n \rightarrow 0$, entonces

$$\text{Cov}(\omega_1(t, Z), \rho_3(t, Z, \delta)) = o(s) + o(b).$$

b) Si $\frac{b}{s} \rightarrow L > 0$ o bien $\frac{b}{s} \rightarrow \infty$, entonces

$$\text{Cov}(\omega_1(t, Z), \rho_3(t, Z, \delta)) = O(b).$$

Demostración.

De nuevo, calculamos la covarianza entre $\omega_1(\cdot, Z)$ y $\rho_3(\cdot, Z, \delta)$ mediante la esperanza del producto de ambas funciones, puesto que la esperanza de la función $\omega_1(\cdot, Z)$ es cero. Dado que además $\mathbb{E}[\rho_3(t, Z, \delta)] = O(b^2)$ y que $\omega_1(\cdot, Z)$ se puede descomponer en dos sumandos, el segundo de los cuales es de orden $O(1)$, entonces:

$$\begin{aligned}
\text{Cov}(\rho_3(t, Z, \delta), \omega_1(t, Z)) &= \mathbb{E}[\rho_3(t, Z, \delta) \omega_1(t, Z)] \\
&= \mathbb{E} \left[\int K_s(t - v_1) K_b(v_1 - Z) \frac{1 - F(v_1)}{1 - H(v_1)} (\delta - p(v_1)) dv_1 \right. \\
&\quad \left. \times \int K_s(t - v_2) f(v_2) \frac{p(v_2)}{1 - H(v_2)} (\mathbf{1}_{\{Z \leq v_2\}} - H(v_2)) dv_2 \right] \\
&= \iiint K_s(t - v_1) K_s(t - v_2) K_b(v_1 - u) \frac{1 - F(v_1)}{1 - H(v_1)} \\
&\quad \times (p(u) - p(v_1)) f(v_2) \frac{p(v_2)}{1 - H(v_2)} h(u) \mathbf{1}_{\{u \leq v_2\}} dv_2 dv_1 du + O(b^2) \\
&= \mathbb{E}[I_1] + O(b^2).
\end{aligned}$$

Definamos la función $m(\cdot, \cdot, \cdot)$ como

$$m(x, y, z) = \frac{1 - F(x)}{1 - H(x)} f(y) \frac{p(y)}{1 - H(y)} (p(z) - p(x)) h(z).$$

En el cálculo de $\mathbb{E}[I_1]$ realizamos ahora los cambios de variable $v_1 = t - sw_1$, $v_2 = t - sw_2$ y $u = v_1 - bu_1 = t - sw_1 - bu_1$ (paso (a)) y sustituimos el integrando por la función $m(\cdot, \cdot, \cdot)$ definida antes (paso (b)):

$$\begin{aligned}
\mathbb{E}[I_1] &\stackrel{(a)}{=} \iiint K(w_1) K(w_2) K(u_1) \frac{1 - F(t - sw_1)}{1 - H(t - sw_1)} \\
&\quad \times [p(t - sw_1 - bu_1) - p(t - sw_1)] f(t - sw_2) \frac{p(t - sw_2)}{1 - H(t - sw_2)} \\
&\quad \times h(t - sw_1 - bu_1) \mathbf{1}_{\{s(w_2 - w_1) \leq bu_1\}} dw_2 dw_1 du_1 \\
&\stackrel{(b)}{=} \iiint K(w_1) K(w_2) K(u_1) m(t - sw_1, t - sw_2, t - sw_1 - bu_1) \\
&\quad \times \mathbf{1}_{\{s(w_2 - w_1) \leq bu_1\}} dw_2 dw_1 du_1.
\end{aligned}$$

Mediante un desarrollo de Taylor de la función $m(\cdot, \cdot, \cdot)$ en torno al punto (t, t, t) , dado que

$$m(t, t, t) = m_2(t, t, t) = m_1(t, t, t) + m_3(t, t, t) = 0,$$

y si el cociente entre las ventanas es $b/s = L_n \rightarrow L \geq 0$, entonces podemos escribir la esperanza anterior como

$$\mathbb{E}[I_1] = -bm_3(t, t, t) C_K(L_n) + o(s) + o(b),$$

siendo $C_K(\cdot)$ la función dada en (3.28). De esta forma, cuando $b/s \rightarrow 0$ se tiene que $\mathbb{E}[I_1] = o(s) + o(b)$, mientras que si $b/s \rightarrow L > 0$ resulta $\mathbb{E}[I_1] = O(b)$.

En caso contrario, cuando $b/s \rightarrow \infty$, entonces

$$\begin{aligned} \mathbb{E}[I_1] &= bm_3(t, t, t) \iiint u_1 K(w_1) K(w_2) K(u_1) \left(\mathbf{1}_{\{u_1 \leq \frac{s}{b}(w_2 - w_1)\}} \right) dw_2 dw_1 du_1 \\ &\quad + o(s) + o(b) = O(b). \end{aligned}$$

Con esto queda demostrado el lema. ■

Lema 3.2.18 *Bajo las condiciones (K1), (H2), (p1) y (f1) se tiene lo siguiente:*

$$\text{Cov}(\omega_2(t, Z), \rho_3(t, Z, \delta)) = o(s) + o(b).$$

Demostración.

Para obtener esta covarianza basta calcular la esperanza del producto de $\omega_2(\cdot, Z)$ y $\rho_3(\cdot, Z, \delta)$, puesto que $\mathbb{E}[\omega_2(\cdot, Z)] = 0$, con $\omega_2(\cdot, Z)$ y $\rho_3(\cdot, Z, \delta)$ dados en (3.20) y (3.25) respectivamente. El segundo término en la expresión de $\omega_2(\cdot, Z)$ es de orden $O(1)$, mientras que $\mathbb{E}[\rho_3(t, Z, \delta)] = O(b^2)$, de manera que los términos dominantes de esta covarianza son los que se obtienen de calcular la esperanza del producto entre la primera de las integrales de $\omega_2(\cdot, Z)$ y el término $\rho_3(\cdot, Z, \delta)$:

$$\begin{aligned} &\text{Cov}(\omega_2(t, Z), \rho_3(t, Z, \delta)) \\ &= \mathbb{E} \left[\int K_s(t-v) f(v) \left(\int_Z^v \frac{p'(w)}{1-H(w)} dw \right) \mathbf{1}_{\{Z \leq v\}} dv \times \rho_3(t, Z, \delta) \right] + O(b^2) \\ &\stackrel{(a)}{=} \mathbb{E} \left[\int K_s(t-v_1) f(v_1) \left(\frac{p(v_1)}{1-H(v_1)} - \gamma(v_1) - \frac{p(Z)}{1-H(Z)} + \gamma(Z) \right) \mathbf{1}_{\{Z \leq v_1\}} dv_1 \right. \\ &\quad \left. \times \int K_s(t-v_2) K_b(v_2-Z) (1-F(v_2)) \frac{\delta - p(v_2)}{1-H(v_2)} dv_2 \right] + O(b^2) \\ &\stackrel{(b)}{=} \iiint K_s(t-v_1) K_s(t-v_2) K_b(v_2-u) f(v_1) (1-F(v_2)) \frac{p(u) - p(v_2)}{1-H(v_2)} \\ &\quad \times \left(\frac{p(v_1)}{1-H(v_1)} - \gamma(v_1) - \frac{p(u)}{1-H(u)} + \gamma(u) \right) h(u) \mathbf{1}_{\{u \leq v_1\}} dv_2 dv_1. + O(b^2) \end{aligned}$$

sin más que aplicando la propiedad 3.2.1 (paso (a)) y el operador esperanza (paso (b)).

Definimos ahora la función $m(\cdot, \cdot, \cdot)$ como

$$m(x, y, z) = f(x) \frac{1 - F(y)}{1 - H(y)} h(z) (p(z) - p(y)) \\ \times \left(\frac{p(x)}{1 - H(x)} - \gamma(x) - \frac{p(z)}{1 - H(z)} + \gamma(z) \right)$$

y realizamos los cambios de variable $v_1 = t - sw_1, v_2 = t - sw_2$ y $u = t - sw_2 - bu_1$. Resulta así que el término dominante de $Cov(\omega_2(t, Z), \rho_3(t, Z, \delta))$ es igual a:

$$\iiint K(w_1) K(w_2) K(u_1) m(t - sw_1, t - sw_2, t - sw_2 - bu_1) \\ \times \mathbf{1}_{\{t - sw_1 \geq 0\}} \mathbf{1}_{\{t - sw_2 \geq 0\}} \mathbf{1}_{\{t - sw_2 - b_1 \geq 0\}} \mathbf{1}_{\{s(w_1 - w_2) \leq bu_1\}} dw_3 dw_1 du_1.$$

Haciendo desarrollos de Taylor en la función $m(\cdot, \cdot, \cdot)$ en torno al punto (t, t, t) , bajo las condiciones (H2), (p1) y (f1), y teniendo en cuenta que

$$m(t, t, t) = m_1(t, t, t) = m_2(t, t, t) = m_3(t, t, t) = 0,$$

se concluye la demostración. ■

Lema 3.2.19 *Bajo las condiciones (K1), (H2), (p1) y (f1) se tiene lo siguiente:*

$$Cov(\omega_1(t, Z) \omega_3(t, Z, \delta)) = o(s) + o(b).$$

Demostración.

La esperanza de la función $\omega_1(\cdot, Z)$, dada en (3.18), es cero de modo que la covarianza entre $\omega_1(\cdot, Z)$ y $\omega_3(\cdot, Z, \delta)$ es la esperanza del producto entre ambas. Además, dado que $\mathbb{E}[\omega_3(t, Z, \delta)] = O(b^2)$ y que el segundo sumando en el que se descompone $\omega_1(\cdot, Z)$ es de orden $O(1)$, entonces la covarianza entre $\omega_1(\cdot, Z)$ y $\omega_3(\cdot, Z, \delta)$ es igual a:

$$Cov(\omega_1(t, Z) \omega_3(t, Z, \delta)) = \mathbb{E}[\omega_1(t, Z) \omega_3(t, Z, \delta)] \\ = \iiint K_s(t - v_1) K_s(t - v_2) K_b(v_3 - u) f(v_1) \frac{f(v_2) p(v_2)}{1 - H(v_2)} \\ \times \frac{p(u) - p(v_3)}{1 - H(v_3)} h(u) \mathbf{1}_{\{0 \leq u \leq v_2\}} \mathbf{1}_{\{0 \leq v_3 \leq v_1\}} dv_3 dv_2 dv_1 du + O(b^2).$$

Definimos previamente la función $m(\cdot, \cdot, \cdot, \cdot)$ como

$$m(x, y, z, t) = f(x) \frac{f(y) p(y) p(t) - p(z)}{1 - H(y) (1 - H(z))} h(t).$$

Con los cambios de variable $v_1 = t - sw_1$, $v_2 = t - sw_2$ y $v_3 = u + bw_3$, la covarianza anterior es igual a

$$\begin{aligned} & Cov(\omega_1(t, Z)\omega_3(t, Z, \delta)) \\ &= \iiint\iiint K(w_1)K(w_2)K(w_3)m(t-sw_1, t-sw_2, u+bw_3, u) \\ & \quad \times \mathbf{1}_{\{0 \leq u \leq t-sw_2\}} \mathbf{1}_{\{0 \leq u+bw_3 \leq t-sw_1\}} dw_3 dw_2 dw_1 du + O(b^2). \end{aligned}$$

De nuevo aplicamos a la función $m(\cdot, \cdot, \cdot, \cdot)$ un desarrollo de Taylor en torno al punto (t, t, u, u) . Dado que

$$m(t, t, u, u) = m_1(t, t, u, u) = m_2(t, t, u, u) = 0$$

y por las hipótesis (H2), (p1) y (f1), podemos concluir que

$$\begin{aligned} Cov(\omega_1(t, Z)\omega_3(t, Z, \delta)) &= b \iiint\iiint w_3 K(w_1)K(w_2)K(w_3) \\ & \quad \times \int m_3(t, t, u, u) \mathbf{1}_{\{0 \leq u \leq t-sw_2\}} \mathbf{1}_{\{-bw_3 \leq u \leq t-sw_1+bw_3\}} dudw_3 dw_2 dw_1 + O(s^2) + o(b) \\ &= b \iiint\iiint w_3 K(w_1)K(w_2)K(w_3) \\ & \quad \times (n(t, t-sw_2, t-sw_1+bw_3) + n(t, t-sw_2, -bw_3)) dudw_3 dw_2 dw_1 + O(s^2) + o(b), \end{aligned}$$

si se define la función $n(\cdot, \cdot, \cdot)$ como

$$n(t, x, y) = \int m_3(t, t, u, u) \mathbf{1}_{\{0 \leq u \leq x\}} \mathbf{1}_{\{0 \leq u \leq y\}} du.$$

Por desarrollos de Taylor de esta función en torno al punto (t, t, t) , entonces la covarianza anterior es igual a

$$Cov(\omega_1(t, Z)\omega_3(t, Z, \delta)) = o(s) + o(b).$$

■

Lema 3.2.20 *Bajo las condiciones (K1), (H2), (p1) y (f1) se tiene lo siguiente:*

$$Cov(\omega_2(t, Z)\omega_3(t, Z, \delta)) = o(s) + o(b).$$

Demostración.

Dado que la esperanza de $\omega_2(\cdot, Z)$ es cero, basta con calcular la esperanza del producto entre $\omega_2(\cdot, Z)$ y $\omega_3(\cdot, Z, \delta)$. Pero al ser $\mathbb{E}[\omega_3(t, Z, \delta)] = O(b^2)$ y el segundo sumando de la expresión de $\omega_2(\cdot, Z)$, dada en (3.20), de orden $O(1)$ entonces la covarianza entre $\omega_2(\cdot, Z)$ y $\omega_3(\cdot, Z, \delta)$ es igual a

$$\begin{aligned} \text{Cov}(\omega_2(t, Z) \omega_3(t, Z, \delta)) &= \mathbb{E}[\omega_2(t, Z) \omega_3(t, Z, \delta)] \\ &= \mathbb{E} \left[\int K_s(t - v_3) f(v_3) \left(\frac{p(v_3)}{1 - H(v_3)} - \gamma(v_3) - \frac{p(Z)}{1 - H(Z)} + \gamma(Z) \right) \mathbf{1}_{\{Z \leq v_3\}} dv_3 \right. \\ &\quad \times \left. \int K_s(t - v_1) f(v_1) \int_0^{v_1} K_b(v_2 - Z) \frac{\delta - p(v_2)}{1 - H(v_2)} dv_2 dv_1 \right] + O(b^2) \\ &= \iiint \int K_s(t - v_1) K_s(t - v_3) K_b(v_2 - u) f(v_1) f(v_3) \frac{p(u) - p(v_2)}{1 - H(v_2)} \\ &\quad \times \left(\frac{p(v_3)}{1 - H(v_3)} - \gamma(v_3) - \frac{p(u)}{1 - H(u)} + \gamma(u) \right) h(u) \mathbf{1}_{\{0 \leq v_2 \leq v_1\}} \mathbf{1}_{\{u \leq v_3\}} dudv_3 dv_2 dv_1 \Big] \\ &\quad + O(b^2). \end{aligned}$$

Definimos ahora la función $m(\cdot, \cdot, \cdot, \cdot)$ de la siguiente manera:

$$m(x, y, z, t) = f(x) f(y) \frac{p(z) - p(t)}{1 - H(t)} \left(\frac{p(y)}{1 - H(y)} - \gamma(y) - \frac{p(z)}{1 - H(z)} + \gamma(z) \right) h(z),$$

y realizamos los cambios de variable $v_1 = t - sw_1$, $v_3 = t - sw_3$ y $v_2 = u + bw_2$, de modo que la covarianza anterior es igual a

$$\begin{aligned} &\iiint \int K(w_1) K(w_3) K(u_1) m(t - sw_1, t - sw_3, u, u + bw_2) \\ &\quad \times \mathbf{1}_{\{0 \leq u + bw_2 \leq t - sw_1\}} \mathbf{1}_{\{0 \leq u \leq t - sw_3\}} dudw_3 dw_2 dw_1 + O(b^2). \end{aligned}$$

Teniendo ahora en cuenta que la función $m(\cdot, \cdot, \cdot, \cdot)$ verifica que

$$m(t, t, u, u) = m_1(t, t, u, u) = m_2(t, t, u, u) = m_4(t, t, u, u) = 0,$$

se demuestra lo establecido en el lema. ■

La última covarianza que se necesita calcular es la covarianza entre $\omega_3(\cdot, Z, \delta)$ y $\rho_3(\cdot, Z, \delta)$. Su expresión asintótica depende de nuevo de la relación entre las ventanas s y b .

Lema 3.2.21 *Bajo las condiciones (K1), (H2), (p1) y (f1) se tiene lo siguiente:*

a) Si $\frac{b}{s} \rightarrow 0$, entonces

$$\begin{aligned} \text{Cov}(\omega_3(t, Z, \delta), \rho_3(t, Z, \delta)) &= \frac{1}{2} f(t) (1 - F(t)) q_1(t) \\ &\quad - s \{ f(t) [(1 - F(t)) q_1(t)]' - f'(t) (1 - F(t)) q_1(t) \} e_K + o(s) + o(b). \end{aligned}$$

b) Si $\frac{b}{s} = L_n \rightarrow L > 0$, entonces

$$\text{Cov}(\omega_3(t, Z, \delta), \rho_3(t, Z, \delta)) = \frac{1}{2} f(t) (1 - F(t)) q_1(t) + O(s).$$

c) Si $\frac{b}{s} \rightarrow \infty$, entonces

$$\text{Cov}(\omega_3(t, Z, \delta), \rho_3(t, Z, \delta)) = \frac{1}{2} f(t) (1 - F(t)) q_1(t) + O(b).$$

Demostración.

Las esperanzas $\mathbb{E}[\omega_3(\cdot, Z, \delta)]$ y $\mathbb{E}[\rho_3(\cdot, Z, \delta)]$ son ambas de orden $O(b^2)$, de modo que la parte dominante de la covarianza entre $\omega_3(\cdot, Z, \delta)$ y $\rho_3(\cdot, Z, \delta)$ es $\mathbb{E}[\omega_3(\cdot, Z, \delta) \rho_3(\cdot, Z, \delta)]$. El producto entre ambas funciones es

$$\begin{aligned} &\int K_s(t - v_1) f(v_1) \int_0^{v_1} K_b(v_2 - Z) \frac{\delta - p(v_2)}{1 - H(v_2)} dv_2 dv_1 \\ &\quad \times \int K_s(t - v_3) K_b(v_3 - Z) \frac{1 - F(v_3)}{1 - H(v_3)} (\delta - p(v_3)) dv_3 \\ &= \iiint K_s(t - v_1) K_s(t - v_3) K_b(v_2 - Z) K_b(v_3 - Z) f(v_1) \\ &\quad \times \frac{1 - F(v_3)}{1 - H(v_3)} \frac{(\delta - p(v_3)) (\delta - p(v_2))}{1 - H(v_2)} \mathbf{1}_{\{0 \leq v_2 \leq v_1\}} dv_3 dv_2 dv_1 \end{aligned}$$

y, por tanto, su esperanza es

$$\begin{aligned} &\mathbb{E}[\omega_3(t, Z, \delta) \rho_3(t, Z, \delta)] \\ &= \iiint K_s(t - v_1) K_s(t - v_3) K_b(v_2 - u) K_b(v_3 - u) f(v_1) \frac{1 - F(v_3)}{1 - H(v_3)} \\ &\quad \times \frac{p(u) - p(u)(p(v_2) + p(v_3)) + p(v_2)p(v_3)}{1 - H(v_2)} h(u) \mathbf{1}_{\{0 \leq v_2 \leq v_1\}} dv_3 dv_2 dv_1 du. \end{aligned}$$

Mediante los cambios de variable $v_1 = t - sw_1$, $v_3 = t - sw_3$, $v_2 = u + bw_2$ y $u = v_3 - bu_1$ obtenemos lo siguiente:

$$\begin{aligned}
\mathbb{E}[\omega_3(t, Z, \delta) \rho_3(t, Z, \delta)] &= \iiint\!\!\!\int K(w_1) K(w_2) K(w_3) K(u_1) f(t - sw_1) \\
&\times \frac{1 - F(t - sw_3)}{1 - H(t - sw_3)} \frac{h(t - sw_3 - bu_1)}{1 - H(t - sw_3 - b(u_1 - w_2))} \\
&\times \{p(t - sw_3 - bu_1) - p(t - sw_3 - bu_1) [p(t - sw_3 - b(u_1 - w_2)) + p(t - sw_3)] \\
&+ p(t - sw_3 - b(u_1 - w_2)) p(t - sw_3)\} \mathbf{1}_{\{0 \leq t - sw_3 - b(u_1 - w_2) \leq t - sw_1\}} dw_3 dw_2 dw_1 du_1. \\
&= \iiint\!\!\!\int K(w_1) K(w_2) K(w_3) K(u_1) \\
&\times m(t - sw_1, t - sw_3, t - sw_3 - bu_1, t - sw_3 - b(u_1 - w_2)) \\
&\times \mathbf{1}_{\{0 \leq t - sw_3 - b(u_1 - w_2)\}} \mathbf{1}_{\{s(w_1 - w_3) \leq b(u_1 - w_2)\}} dw_3 dw_2 dw_1 du_1.
\end{aligned}$$

donde

$$m(x, y, z, t) = f(x) \frac{1 - F(y)}{1 - H(y)} \frac{h(z)}{1 - H(t)} [p(z) - p(z)(p(t) + p(y)) + p(t)p(y)].$$

Realizamos un desarrollo de Taylor de grado dos de la función $m(t - sw_1, t - sw_3, t - sw_3 - bu_1, t - sw_3 - b(u_1 - w_2))$ en torno al punto (t, t, t, t) , suponiendo que $t > 0$ es fijo y que el tamaño muestral es suficientemente grande para que la primera función indicadora sea igual a uno, de modo que

$$\begin{aligned}
\mathbb{E}[\omega_3(t, Z, \delta) \rho_3(t, Z, \delta)] &= m(t, t, t, t) \iiint\!\!\!\int K(w_1) K(w_2) K(w_3) K(u_1) \\
&\times \mathbf{1}_{\{s(w_1 - w_3) \leq b(u_1 - w_2)\}} dw_3 dw_2 dw_1 du_1 \\
&- sm_1(t, t, t, t) \iiint\!\!\!\int w_1 K(w_1) K(w_2) K(w_3) K(u_1) \\
&\times \mathbf{1}_{\{s(w_1 - w_3) \leq b(u_1 - w_2)\}} dw_3 dw_2 dw_1 du_1 \\
&- s[m_2 + m_3 + m_4](t, t, t, t) \iiint\!\!\!\int w_3 K(w_1) K(w_2) K(w_3) K(u_1) \quad (3.29) \\
&\times \mathbf{1}_{\{s(w_1 - w_3) \leq b(u_1 - w_2)\}} dw_3 dw_2 dw_1 du_1 \\
&- bm_3(t, t, t, t) \iiint\!\!\!\int u_1 K(w_1) K(w_2) K(w_3) K(u_1) \\
&\times \mathbf{1}_{\{s(w_1 - w_3) \leq b(u_1 - w_2)\}} dw_3 dw_2 dw_1 du_1 \\
&- bm_4(t, t, t, t) \iiint\!\!\!\int (u_1 - w_2) K(w_1) K(w_2) K(w_3) K(u_1) \\
&\times \mathbf{1}_{\{s(w_1 - w_3) \leq b(u_1 - w_2)\}} dw_3 dw_2 dw_1 du_1 + O(s^2) + O(b^2),
\end{aligned}$$

donde

$$\begin{aligned} m(t, t, t, t) &= f(t) (1 - F(t)) q_1(t) \\ m_1(t, t, t, t) &= f'(t) (1 - F(t)) q_1(t) \\ [m_2 + m_3 + m_4](t, t, t, t) &= f(t) [(1 - F(t)) q_1(t)]'. \end{aligned}$$

y $q_1(\cdot)$ dada en (3.23). De nuevo, debemos distinguir dos posibles casos, aquel en el que el cociente b/s tiende a un número L no negativo, y cuando b/s tiende a infinito.

En el primero de los casos, cuando $b/s \rightarrow L > 0$, entonces la primera integral en (3.29) es igual a

$$\iiint K(w_1) K(w_2) K(w_3) K(u_1) \mathbf{1}_{\{w_1 - w_3 \leq L(u_1 - w_2)\}} dw_1 dw_2 dw_3 du_1 = \frac{1}{2}, \quad (3.30)$$

puesto que se puede escribir, como función de L , de la siguiente manera:

$$\eta(L) = \iiint K(w_2) K(w_3) K(u_1) \mathbb{K}(w_3 + L(u_1 - w_2)) dw_2 dw_3 du_1,$$

siendo, por la propiedad 3.2.5, $\eta(0) = \frac{1}{2}$. Para demostrar (3.30) veamos que la derivada primera de $\eta(\cdot)$ es idénticamente nula:

$$\begin{aligned} \frac{\partial}{\partial L} \eta(L) &= u_1 \iiint K(w_2) K(w_3) K(u_1) K(w_3 + L(u_1 - w_2)) dw_2 dw_3 du_1 \\ &\quad - w_2 \iiint K(w_2) K(w_3) K(u_1) K(w_3 + L(u_1 - w_2)) dw_2 dw_3 du_1. \end{aligned}$$

El primer sumando de $\frac{\partial}{\partial L} \eta(L)$ es igual al segundo, por la simetría del núcleo evaluado en u_1 y w_2 (paso (a)) y en w_3 y $w_3 - L(w_2 - u_1)$ (paso (b)):

$$\begin{aligned} &u_1 \iiint K(w_2) K(w_3) K(u_1) K(w_3 + L(u_1 - w_2)) dw_2 dw_3 du_1 \\ &\stackrel{(a)}{=} u_1 \iiint K(w_2) K(w_3) K(u_1) K(w_3 - L(w_2 - u_1)) dw_2 dw_3 du_1 \\ &\stackrel{(b)}{=} u_1 \iiint K(w_2) K(w_3) K(u_1) K(w_3 + L(w_2 - u_1)) dw_2 dw_3 du_1 \end{aligned}$$

Como consecuencia, la función $\eta(\cdot)$ es constantemente igual a $\frac{1}{2}$ y queda probado (3.30).

Entonces, cuando $b/s \rightarrow L > 0$, se tiene que

$$\text{Cov}(\omega_3(t, Z, \delta), \rho_3(t, Z, \delta)) = \frac{1}{2}f(t)(1 - F(t))q_1(t) + O(s).$$

En el caso en el que $b/s = L_n \rightarrow 0$, la primera de las integrales en (3.29) es también igual a

$$\iiint K(u)K(v)K(\omega)\mathbb{K}(u + L_n(v - \omega))dudvd\omega = \frac{1}{2},$$

mientras que las otras cuatro son iguales a

$$\begin{aligned} & \iiint w_1 K(w_1)K(w_2)K(w_3)K(u_1)\mathbf{1}_{\{w_1 \leq w_3 + L_n(u_1 - w_2)\}}dw_3dw_2dw_1du_1 \\ &= - \iiint w_1 K(w_1)K(w_2)K(w_3)K(u_1)\mathbf{1}_{\{w_3 \leq w_1 + L_n(u_1 - w_2)\}}dw_3dw_2dw_1du_1 \\ &= - \iiint w_1 K(w_1)K(w_2)K(u_1)\mathbb{K}(w_1 + L_n(u_1 - w_2))dw_2dw_1du_1 \\ &= -e_K + O(L_n^2) \end{aligned}$$

para la primera de ellas,

$$\begin{aligned} & \iiint w_3 K(w_1)K(w_2)K(w_3)K(u_1)\mathbf{1}_{\{w_1 \leq w_3 + L_n(u_1 - w_2)\}}dw_3dw_2dw_1du_1 \\ &= \iiint w_3 K(w_2)K(w_3)K(u_1)\mathbb{K}(w_3 + L_n(u_1 - w_2))dw_3dw_2dw_1du_1 \\ &= e_K + O(L_n^2) \end{aligned}$$

en el caso de la segunda, y

$$\iiint vK(u)K(v)K(\omega)\mathbb{K}(u + L_n(v - \omega))dudvd\omega = L_n c_K d_K + O(L_n^2) = o(1)$$

y

$$\begin{aligned} & \iiint (v - \omega)K(u)K(v)K(\omega)\mathbb{K}(u + L_n(v - \omega))dudvd\omega \\ &= 2L_n c_K d_K + O(L_n^2) = o(1) \end{aligned}$$

respectivamente para las dos últimas, de manera que

$$\begin{aligned} \text{Cov}(\omega_3(t, Z, \delta), \rho_3(t, Z, \delta)) &= \frac{1}{2} f(t) (1 - F(t)) q_1(t) \\ &\quad - s \{ f(t) [(1 - F(t)) q_1(t)]' - f'(t) (1 - F(t)) q_1(t) \} e_K + o(s) + o(b). \end{aligned}$$

Por último, cuando $b/s \rightarrow \infty$, es decir, cuando $s/b \rightarrow 0$ se tiene que la primera integral en (3.29) es igual a

$$\begin{aligned} &\int \int \int \int K(u_1) K(u_2) K(w_1) K(w_2) \mathbf{1}_{\{\frac{s}{b}(u_1 - u_2) \leq w_1 - w_2\}} du_1 du_2 dw_1 dw_2 \\ &= \int \int \int K(u_1) K(u_2) K(w_1) \mathbb{K}\left(w_1 - \frac{s}{b}(u_1 - u_2)\right) dw_1 du_1 du_2 = \frac{1}{2}. \end{aligned}$$

Entonces

$$\text{Cov}(\omega_3(t, Z, \delta), \rho_3(t, Z, \delta)) = \frac{1}{2} f(t) (1 - F(t)) q_1(t) + O(b).$$

Recopilamos todos estos resultados y obtenemos la expresión final de la covarianza entre $\omega_3(\cdot, Z, \delta)$ y $\rho_3(\cdot, Z, \delta)$. ■

Estamos ahora en condiciones de calcular las expresiones asintóticas del sesgo y varianza de $\overline{f_n^P}(\cdot)$. Para ello necesitaremos definir previamente las siguientes funciones:

$$A(t) = f(t) \frac{1 - F(t)}{1 - H(t)}, \quad (3.31)$$

$$B(t) = f^2(t) \int_0^t \frac{f(v)}{(1 - F(v))(1 - H(v))} dv - \frac{f^2(t)}{1 - H(t)}, \quad (3.32)$$

$$C(t) = f^2(t) q_1(t) e_K, \quad (3.33)$$

$$\begin{aligned} D(t) &= \frac{1}{2} \left[(1 - F(t))^2 q_2(t) \right]'' m_K \\ &\quad + 2 \left[f(t) (1 - F(t)) q_2'(t) - f'(t) (1 - F(t)) q_2(t) - 2f^2(t) q_2(t) \right] e_K, \end{aligned}$$

donde m_K se definió en (3.15), y $q_1(\cdot)$, $q_2(\cdot)$, c_K , d_K y e_K fueron dados en (2.40), (2.7), (2.36) y (2.37), en el capítulo anterior.

Teorema 3.2.2 (Sesgo y varianza asintóticos de $\overline{f_n^P}(\cdot)$) Bajo las condiciones (K1), (H3), (f2) y (p2) el sesgo de $\overline{f_n^P}(\cdot)$ es

$$\begin{aligned} \mathbb{E} \left[\overline{f_n^P}(t) - f(t) \right] &= \frac{1}{2} s^2 f''(t) d_K + b^2 d_K [(1 - F(t)) \alpha(t)]' \\ &\quad + O(s^4) + O(b^4) + O(s^2 b^2), \end{aligned} \quad (3.34)$$

y la varianza es, según el cociente entre las ventanas,

(a) Si $b/s \rightarrow 0$

$$\text{Var} \left(\overline{f_n^P}(t) \right) = \frac{1}{ns} A(t) c_K + \frac{1}{n} B(t) - 2 \frac{b}{n} C(t) + \frac{s}{n} D(t) + o \left(\frac{b}{n} \right) + o \left(\frac{s}{n} \right). \quad (3.35)$$

(b) Si $b/s \rightarrow L \in (0, \infty)$

$$\text{Var} \left(\overline{f_n^P}(t) \right) = \frac{1}{ns} [A(t) p(t) c_K + A(t) (1 - p(t)) A_K(L)] + \frac{1}{n} B(t) + O \left(\frac{s}{n} \right). \quad (3.36)$$

(c) Si $b/s \rightarrow \infty$

$$\text{Var} \left(\overline{f_n^P}(t) \right) = \frac{1}{ns} A(t) p(t) c_K + \frac{1}{nb} A(t) (1 - p(t)) c_K + \frac{1}{n} B(t) + o \left(\frac{1}{n} \right). \quad (3.37)$$

Demostración.

El sesgo del estimador $\overline{f_n^P}(t) = f(t) + \beta_n(t) + \sigma_n(t)$ se calcula fácilmente, ya que es igual a

$$\mathbb{E} \left[\overline{f_n^P}(t) - f(t) \right] = \beta_n(t) + \mathbb{E} [\sigma_n(t)],$$

con $\beta_n(\cdot)$ y $\sigma_n(\cdot)$ dados en (3.2) y (3.3). Usando (f2), podemos efectuar un desarrollo de Taylor con resto en forma integral, y expresar $f(t - sv)$ de la siguiente forma:

$$\begin{aligned} f(t - sv) &= f(t) - sv f'(t) + \frac{1}{2} s^2 v^2 f''(t) - \frac{1}{6} s^3 v^3 f'''(t) \\ &\quad + \frac{1}{6} s^4 v^4 \int_0^1 (1-x)^3 f^{(4)}(t - vsx) dx. \end{aligned}$$

Bajo la hipótesis (K1), entonces

$$\beta_n(t) = \frac{1}{2} s^2 f''(t) d_K + O(s^4).$$

Con respecto a $\mathbb{E}[\sigma_n(\cdot)]$, empezamos con una integración por partes (paso (a)) y con un desarrollo de Taylor de la función $\mathbb{E}[\varepsilon_1(t - vs)]'$ (paso (b)):

$$\begin{aligned} \mathbb{E}[\sigma_n(t)] &= \frac{1}{s} \int \mathbb{E}[\varepsilon_1(t - sv)] K'(v) dv \stackrel{(a)}{=} \int \mathbb{E}[\varepsilon_1(t - sv)]' K(v) dv \\ &\stackrel{(b)}{=} \mathbb{E}[\varepsilon_1(t)]' + s^2 \int \int_0^1 \mathbb{E}[\varepsilon_1(t - svx)]''' v^2 K(v) dx dv \end{aligned} \quad (3.38)$$

donde la derivada de la función $\mathbb{E}[\varepsilon_1(t)]$ es igual a

$$\begin{aligned} \mathbb{E}[\varepsilon_1(t)]' &= \frac{\partial}{\partial t} [(1 - F(t)) \mathbb{E}[g_3(t, Z_1, \delta_1)]] \\ &= -f(t) \mathbb{E}[g_3(t, Z_1, \delta_1)] + (1 - F(t)) \frac{\partial}{\partial t} \mathbb{E}[g_3(t, Z_1, \delta_1)] \end{aligned}$$

Bajo las hipótesis (H3) y (p2), la esperanza de $g_3(t, Z, \delta)$ es

$$\begin{aligned}\mathbb{E}[g_3(t, Z, \delta)] &= \mathbb{E}\left[\int_0^t K_b(v-Z) \frac{\delta - p(v)}{1-H(v)} dv\right] \\ &= \iint_0^t K_b(v-u) \frac{p(u)-p(v)}{1-H(v)} h(u) dvdu = d_K \alpha(t) b^2 + O(b^4),\end{aligned}$$

y su derivada

$$\begin{aligned}\frac{\partial}{\partial t} \mathbb{E}[g_3(t, Z, \delta)] &= \frac{\partial}{\partial t} \iint_0^t K_b(v-Z) \frac{p(u)-p(v)}{1-H(v)} h(u) dvdu \\ &= \int K(v) \frac{p(t-vb) - p(t)}{1-H(t)} h(t-vb) dv \\ &= b^2 \frac{\frac{1}{2} p''(t) h(t) + p'(t) h'(t)}{1-H(t)} d_K + O(b^4) = b^2 \alpha'(t) d_K + O(b^4).\end{aligned}$$

Entonces,

$$\mathbb{E}[\varepsilon_1(t)]' = b^2 d_K [(1-F(t)) \alpha(t)]' + O(b^4).$$

El término de orden $O(b^4)$ es función de términos residuales provenientes de desarrollos de Taylor de las funciones $h(\cdot)$, $p(\cdot)$ y $f(\cdot)$ de modo que, bajo las hipótesis (H3), (p2) y (f2) y dado que el intervalo $[0, t_H]$ es compacto, podemos concluir que

$$\sup_{t \in [0, t_H]} |\mathbb{E}[\varepsilon_1(t)]'| = O(b^2). \quad (3.39)$$

Mediante cálculos y argumentos análogos, obtenemos

$$\sup_{t \in [0, t_H]} \mathbb{E}[\varepsilon_1(t)]''' = O(b^2) \quad (3.40)$$

y, por tanto,

$$\mathbb{E}[\sigma_n(t)] = b^2 d_K [(1-F(t)) \alpha(t)]' + O(b^4) + O(s^2 b^2).$$

En resumen,

$$\mathbb{E}\left[\overline{f_n^P}(t) - f(t)\right] = \frac{1}{2} s^2 f''(t) d_K + b^2 d_K [(1-F(t)) \alpha(t)]' + O(s^4) + O(b^4) + O(s^2 b^2).$$

La expresión de la varianza de $\overline{f_n^P}(\cdot)$ se obtiene reuniendo los resultados dados en los lemas 3.2.1-3.2.21 y la expresión (3.9). ■

Propiedad 3.2.11

Cuando la ventana de presuavizado es muy pequeña ($b/s \rightarrow 0$), las expresiones anteriores se reducen a las correspondientes del estimador de la densidad con pesos Kaplan-Meier:

$$\begin{aligned}\mathbb{E} [f_n^{KM}(t) - f(t)] &= \frac{1}{2} s^2 f''(t) d_K + o(s^2), \\ \text{Var} (f_n^{KM}(t)) &= \frac{1}{ns} f(t) \frac{1-F(t)}{1-H(t)} c_K + o((ns)^{-1}).\end{aligned}$$

Propiedad 3.2.12

En el caso de no censura, este resultado es una extensión del teorema 1 de Cao (1993) con datos completos, dado que $A(t) = f(t)$, $p(t) = 1$ y $B(t) = -f^2(t)$:

$$\begin{aligned}\mathbb{E} [\hat{f}_n(t) - f(t)] &= \frac{1}{2} s^2 f''(t) d_K + o(s^2), \\ \text{Var} (\hat{f}_n(t)) &= \frac{1}{ns} f(t) c_K + o((ns)^{-1}).\end{aligned}$$

3.2.2. Normalidad asintótica

La normalidad asintótica puntual del estimador presuavizado $f_n^P(\cdot)$ se probará a partir de los teorema 3.2.1 y 3.2.2.

Teorema 3.2.3 (Normalidad asintótica de $f_n^P(\cdot)$) Bajo las condiciones (K1), (H3), (H4), (f2), (p2), (v1) y (v3), entonces

$$\sqrt{ns} (f_n^P(t) - f(t)) \xrightarrow{d} N(B(t), \sigma^2(t)),$$

donde

$$B(t) = \begin{cases} 0 & \text{si } ns^5 \rightarrow 0 \\ \frac{1}{2} C^{5/2} f''(t) d_K & \text{si } s = Cn^{-1/5}, nb^5 \rightarrow 0 \\ \frac{1}{2} C^{5/2} f''(t) d_K + C^{1/2} B^2 [(1-F(t)) \alpha(t)]' d_K & \text{si } s = Cn^{-1/5}, b = Bn^{-1/5} \end{cases}$$

$$\sigma^2(t) = \begin{cases} f(t) \frac{1-F(t)}{1-H(t)} c_K & \text{si } \frac{b}{s} \rightarrow 0 \\ \frac{f^2(t)}{h(t)} c_K + f(t) \left(\frac{1-F(t)}{1-H(t)} - \frac{f(t)}{h(t)} \right) A_K(L) & \text{si } \frac{b}{s} \rightarrow L \in (0, \infty) \\ \frac{f^2(t)}{h(t)} c_K & \text{si } \frac{b}{s} \rightarrow \infty \end{cases}$$

y los términos $\alpha(\cdot)$, $A_K(\cdot)$, c_K y d_K vienen dados en (2.38), (3.24) y (2.7).

Demostración.

En primer lugar, aplicamos el teorema 3.2.1 para obtener una descomposición de $f_n^P(\cdot) - f(\cdot)$:

$$f_n^P(t) - f(t) = \beta_n(t) + \sigma_n(t) + e_n(t),$$

donde el término despreciable en la representación verifica que

$$\sup_{t \in [0, t_H]} |e_n(t)| = O\left(s^{-1} \left(b^2 + (nb)^{-1/2} (\log n)^{1/2}\right)^2 (\log n)^2\right) \quad c.s.$$

Puesto que, bajo la hipótesis (v3),

$$\sqrt{ns} s^{-1} \left(b^2 + (nb)^{-1/2} (\log n)^{1/2}\right)^2 (\log n)^2 \rightarrow 0$$

entonces la distribución asintótica del proceso $\sqrt{ns}(f_n^P(\cdot) - f(\cdot))$ coincide con la de

$$\begin{aligned} \sqrt{ns}(\beta_n(t) + \sigma_n(t)) &= \sqrt{ns}\beta_n(t) + \sqrt{ns}\frac{1}{ns}\sum_{i=1}^n \int \varepsilon_i(t - vs) K'(v) dv \\ &= \sqrt{ns}\frac{1}{ns}\sum_{i=1}^n \int (\varepsilon_i(t - vs) - \mathbb{E}[\varepsilon_i(t - vs)]) K'(v) dv \\ &\quad + \sqrt{ns}\frac{1}{ns}\sum_{i=1}^n \int \mathbb{E}[\varepsilon_i(t - vs)] K'(v) dv + \sqrt{ns}\beta_n(t) \\ &= (I) + (II) + (III) \end{aligned}$$

El sesgo de $\sqrt{ns}(f_n^P(\cdot) - f(\cdot))$ viene de los términos (II) y (III). Estudiamos el último de ellos:

$$\begin{aligned} (III) &= \sqrt{ns}\beta_n(t) = \sqrt{ns}\left(\int K_s(t - v) dF(v) - f(t)\right) \\ &= \sqrt{ns}\left(\frac{1}{2}s^2 f''(t) d_K + o(s^2)\right) = \frac{1}{2}\sqrt{ns^5} f''(t) d_K + o(\sqrt{ns^5}). \end{aligned}$$

A su vez, el término (II) es igual a:

$$(II) = \sqrt{ns}\frac{1}{ns}\sum_{i=1}^n \int \mathbb{E}[\varepsilon_i(t - vs)] K'(v) dv = \sqrt{ns}\frac{1}{s} \int \mathbb{E}[\varepsilon_1(t - vs)] K'(v) dv.$$

La esperanza de la función $\varepsilon_1(\cdot)$, dada en (3.4), es igual a

$$\begin{aligned}\mathbb{E}[\varepsilon_1(t)] &= (1 - F(t)) \mathbb{E}[g_3(t, Z_1, \delta_1)] = \iint_0^t K_b(v - u) \frac{p(u) - p(v)}{1 - H(v)} h(u) dv du \\ &= d_K (1 - F(t)) \alpha(t) b^2 + o(b^2).\end{aligned}$$

Bajo las condiciones de regularidad sobre las funciones $h(\cdot)$, $p(\cdot)$ y $f(\cdot)$ dadas en las hipótesis (H3), (p2) y (f2) y aplicando desarrollos de Taylor, podemos permutar el término $o(b^2)$ con la integral y escribir el término (II) de la siguiente manera:

$$\begin{aligned}(II) &= \sqrt{ns} \left(\frac{1}{s} \int d_K (1 - F(t - vs)) \alpha(t - vs) b^2 K'(v) dv + o(b^2) \right) \\ &= \sqrt{ns} \left(\frac{1}{s} d_K b^2 \int (1 - F(t)) \alpha(t) - vs [(1 - F(t)) \alpha(t)]' \right. \\ &\quad \left. + \frac{1}{2} v^2 s^2 [(1 - F(t)) \alpha(t)]'' + o(s^2) \right) K'(v) dv + o(b^2) \\ &= \sqrt{ns} [b^2 d_K [(1 - F(t)) \alpha(t)]' + o(sb^2) + o(b^2)] \\ &= \sqrt{nsb^4} [(1 - F(t)) \alpha(t)]' d_K + o(\sqrt{nsb^4}),\end{aligned}$$

puesto que

$$\begin{aligned}\mathbb{E}[\varepsilon_1(t)] &= \mathbb{E}[(1 - F(t)) (g_1(t, Z_1) - g_2(t, Z_1) + g_3(t, Z_1, \delta_1))] \\ &= d_K (1 - F(t)) \alpha(t) b^2 + o(b^2).\end{aligned}$$

En resumen,

$$\begin{aligned}\mathbb{E}[\sqrt{ns} (f_n^P(t) - f(t))] &= \frac{1}{2} \sqrt{ns^5} f''(t) d_K + \sqrt{nsb^4} d_K [(1 - F(t)) \alpha(t)]' \\ &\quad + o(\sqrt{ns^5}) + o(\sqrt{nsb^4}).\end{aligned}$$

a) Supongamos que $s = Cn^{-1/5} + o(n^{-1/5})$ y $b = Bn^{-1/5} + o(n^{-1/5})$. Entonces,

$$\begin{aligned}\frac{1}{2} \sqrt{ns^5} f''(t) d_K &= \frac{1}{2} C^{5/2} f''(t) d_K + o(1), \\ \sqrt{nsb^4} d_K [(1 - F(t)) \alpha(t)]' &= C^{1/2} B^2 d_K [(1 - F(t)) \alpha(t)]' + o(1),\end{aligned}$$

y por tanto,

$$\mathbb{E}[\sqrt{ns} (f_n^P(t) - f(t))] = \frac{1}{2} C^{5/2} f''(t) d_K + C^{1/2} B^2 d_K [(1 - F(t)) \alpha(t)]' + o(1).$$

b) Si $s = Cn^{-1/5} + o(n^{-1/5})$ y $nb^5 \rightarrow 0$, tenemos lo siguiente:

$$\frac{1}{2}\sqrt{ns^5}f''(t)d_K = \frac{1}{2}C^{5/2}f''(t)d_K + o(1).$$

Con respecto al término con $\sqrt{nsb^4}$, usamos que $b/s \rightarrow 0$ y que $ns^5 \rightarrow C$, y por tanto

$$nsb^4 = ns^5 \left(\frac{b}{s}\right)^4 \rightarrow 0,$$

es decir,

$$\mathbb{E}[\sqrt{ns}(f_n^P(t) - f(t))] = \frac{1}{2}C^{5/2}f''(t)d_K + o(1).$$

c) En el caso en que $ns^5 \rightarrow 0$ y $b = Bn^{-1/5} + o(n^{-1/5})$, tenemos que $\frac{1}{2}\sqrt{ns^5}f''(t)d_K = o(1)$. Por otra parte, usando que $s/b \rightarrow 0$ y que $nb^5 \rightarrow B$, tenemos también que

$$nsb^4 = nb^5 \frac{s}{b} \rightarrow 0,$$

lo que implica que

$$\mathbb{E}[\sqrt{ns}(f_n^P(t) - f(t))] = o(1).$$

d) Sea $ns^5 \rightarrow 0$ y $nb^5 \rightarrow 0$, entonces $\frac{1}{2}\sqrt{ns^5}f''(t)d_K = o(1)$. Distinguiamos ahora tres posibles situaciones, según sea el cociente entre las ventanas:

d.1) Si $\frac{b}{s} \rightarrow 0$ entonces

$$\sqrt{nsb^4}d_K [(1 - F(t))\alpha(t)]' = o(\sqrt{ns^5}) = o(1).$$

d.2) Si $\frac{b}{s} \rightarrow L \in (0, \infty)$ entonces

$$\sqrt{nsb^4}d_K [(1 - F(t))\alpha(t)]' = O(\sqrt{ns^5}) = o(1).$$

d.3) Si $\frac{b}{s} \rightarrow \infty$ entonces

$$\sqrt{nsb^4}d_K [(1 - F(t))\alpha(t)]' = O(\sqrt{nb^5}) = o(1).$$

En resumen:

a) Si $ns^5 \rightarrow 0$ entonces

$$\mathbb{E}[\sqrt{ns}(f_n^P(t) - f(t))] = o(1).$$

b) Si $s = Cn^{-1/5}$ y $nb^5 \rightarrow 0$ entonces

$$\mathbb{E} [\sqrt{ns} (f_n^P(t) - f(t))] = \frac{1}{2} C^{5/2} f''(t) d_K + o(1).$$

c) Si $s = Cn^{-1/5}$ y $b = Bn^{-1/5}$ entonces

$$\mathbb{E} [\sqrt{ns} (f_n^P(t) - f(t))] = \frac{1}{2} C^{5/2} f''(t) d_K + C^{1/2} B^2 d_K [(1 - F(t)) \alpha(t)]' + o(1).$$

Por último, estudiamos el término (I) . Éste puede escribirse como una suma de variables aleatorias i.i.d. de media cero

$$(I) = \sum_{i=1}^n \eta_{i,n}(t),$$

donde las variables $\eta_{i,n}(t)$ están definidas por

$$\eta_{i,n}(t) = (ns)^{-1/2} \int (\varepsilon_i(t - vs) - \mathbb{E}[\varepsilon_i(t - vs)]) K'(v) dv.$$

Probando que $\text{Var} [\eta_{i,n}(t)] = \sigma_{i,n}^2(t) < +\infty$ para cada $i = 1, \dots, n$ y $\sigma_n^2(t) = \sum_{i=1}^n \text{Var} [\eta_{i,n}(t)]$ es un valor acotado, podemos aplicar el Teorema Central del Límite para disposiciones triangulares (ver teorema 7.2 de Billingsley (1968), pag. 42) y obtener

$$\frac{\sum_{i=1}^n \eta_{i,n}(t)}{\sum_{i=1}^n \text{Var} [\eta_{i,n}(t)]} \xrightarrow{d} N(0, 1)$$

siempre que la condición de Lindeberg se verifique.

En primer lugar, probaremos que la varianza $\text{Var} [\eta_{i,n}(\cdot)]$ y $\sigma_n^2(\cdot)$ son finitas:

a) Si $\frac{b}{s} \rightarrow 0$ entonces

$$\begin{aligned} \text{Var} [\eta_{1,n}(t)] &= \text{Var} \left((ns)^{-1/2} \int \varepsilon_1(t - vs) K'(v) dv \right) \\ &= (ns)^{-1} \text{Var} \left(s \int K_s(t - v) d\varepsilon_1(v) \right) \\ &= n^{-1} f(t) \frac{1 - F(t)}{1 - H(t)} c_K + O\left(\frac{s}{n}\right), \end{aligned}$$

y, en consecuencia,

$$\sigma_n^2(t) = \sum_{i=1}^n \text{Var} [\eta_{i,n}(t)] = f(t) \frac{1 - F(t)}{1 - H(t)} c_K + O(s).$$

b) Supongamos $\frac{b}{s} \rightarrow L \in (0, \infty)$. Entonces

$$\text{Var} [\eta_{1,n}(t)] = n^{-1} [A(t)p(t)c_K + A(t)(1-p(t))A_K(L)] + O\left(\frac{s}{n}\right).$$

Así,

$$\sigma_n^2(t) = \sum_{i=1}^n \text{Var} [\eta_{i,n}(t)] = A(t)p(t)c_K + A(t)(1-p(t))A_K(L) + O(s).$$

c) Cuando $\frac{b}{s} \rightarrow \infty$ tenemos que

$$\text{Var} [\eta_{1,n}(t)] = n^{-1} \frac{f^2(t)}{h(t)} c_K + O\left(\frac{s}{n}\right),$$

y por tanto

$$\sigma_n^2(t) = \sum_{i=1}^n \text{Var} [\eta_{i,n}(t)] = \frac{f^2(t)}{h(t)} c_K + o(1).$$

En cualquiera de los tres casos, las condiciones $\sigma_{1,n}^2(t) = \text{Var} [\eta_{1,n}(t)] < \infty$ y $\sigma_n^2(t)$ acotado (como sucesión de n para un $t > 0$ fijo) se cumplen. Ahora, probaremos que también se verifica la condición de Lindeberg:

$$\frac{1}{\sigma_n^2(t)} \sum_{i=1}^n \int_{\{|\eta_{i,n}(t)| > \varepsilon \sigma_n(t)\}} \eta_{i,n}^2(t) dP \rightarrow 0 \quad \forall \varepsilon > 0. \quad (3.41)$$

Si fijamos $\varepsilon > 0$ y definimos las siguientes funciones

$$I_{i,n} = \mathbf{1}_{\{|\eta_{i,n}(t)| > \varepsilon \sigma_n(t)\}} \quad \text{y} \quad \eta_n = \sum_{i=1}^n \eta_{i,n}^2 I_{i,n},$$

entonces la condición de Lindeberg dada en (3.41) puede escribirse como

$$\frac{1}{\sigma_n^2(t)} \mathbb{E}(\eta_n) \rightarrow 0.$$

Puesto que las funciones $g_1(\cdot, Z)$, $g_2(\cdot, Z)$ y $g_3(\cdot, Z, \delta)$ están acotadas para todo $t \in [0, t_H]$, entonces también lo está la función

$$\varepsilon_1(t) = (1 - F(t)) (g_1(t, Z) - g_2(t, Z) + g_3(t, Z, \delta)).$$

Usando además que $n \rightarrow \infty$, entonces:

$$\begin{aligned} & \exists n_0 \in \mathbb{N} : \forall n \geq n_0 \Rightarrow I_{i,n}(\omega) = \mathbf{1}_{\{|\eta_{i,n}(\omega)| > \varepsilon \sigma_n(\omega)\}} = 0 \quad \forall \omega, \forall i = 1, \dots, n \\ & \Rightarrow \exists n_0 \in \mathbb{N} : \forall n \geq n_0 \Rightarrow \eta_n(\omega) = 0 \quad \forall \omega \\ & \Rightarrow \exists n_0 \in \mathbb{N} : \forall n \geq n_0 \Rightarrow \mathbb{E}(\eta_n) = 0, \end{aligned}$$

de modo que

$$\lim_{n \rightarrow \infty} \frac{1}{\sigma_n^2(t)} \mathbb{E}(\eta_n) = 0 \quad \forall \varepsilon > 0.$$

■

Observación 3.2.3

Si $ns^5 \rightarrow \infty$ o bien $nsb^4 \rightarrow \infty$, entonces el sesgo dominaría a la varianza, y la distribución límite sería degenerada.

3.2.3. Representación asintótica del MISE

Todos los métodos de estimación no paramétricos, y en concreto la estimación tipo núcleo de la función de densidad, dependen de forma crucial de la selección del parámetro de suavización o ventana. Su adecuada selección, por tanto, es de suma importancia, y se estudiará en profundidad en el siguiente capítulo de esta memoria.

La idea que subyace en la selección del parámetro de suavización es la minimización de algún tipo de error cometido en el proceso de estimación. En el caso de la estimación presuavizada de la densidad con datos censurados, se puede comprobar por (3.34)-(3.37) que se puede hacer el sesgo de $f_n^P(\cdot)$ tan pequeño como se quiera tomando las ventanas s y b arbitrariamente pequeñas, pero a costa de aumentar la varianza.

Un criterio razonable para elegir el parámetro ventana puede ser el de minimizar algún tipo de distancia entre la función a estimar y su estimación no paramétrica. Uno de los criterios globales más aplicados es la minimización del error cuadrático medio integrado:

$$\begin{aligned} MISE(f_n^P(\cdot)) &= \mathbb{E} \left[\int (f_n^P(v) - f(v))^2 \omega(v) dv \right] \\ &= \int \text{Sesgo}^2(f_n^P(v)) \omega(v) dv + \int \text{Var}(f_n^P(v)) \omega(v) dv. \end{aligned}$$

Se ha introducido en la definición de $MISE(f_n^P(\cdot))$ una función de ponderación, $\omega(\cdot)$, con el fin de poder atribuir una menor importancia a los errores cometidos en ciertas zonas del soporte de la variable que presenten mayores dificultades de estimación.

En el siguiente teorema se da una representación asintótica para el $MISE(f_n^P(\cdot))$. A partir de ella, se darán expresiones explícitas para las ventanas óptimas en el siguiente capítulo. Definamos previamente las siguientes funciones:

$$M(t) = d_K f''(t) \quad \text{y} \quad N(t) = d_K [(1 - F(t)) \alpha(t)]', \quad (3.42)$$

donde d_K y $\alpha(\cdot)$ están definidas en (2.7) y (2.38) respectivamente.

Además de las hipótesis enunciadas al comienzo de este capítulo, se requerirá la siguiente hipótesis sobre la función de peso $\omega(\cdot)$:

(ω_1) La función de peso $\omega(\cdot)$ es una función continua, derivable que verifica

$$\int \omega(v) dv < \infty \quad \text{y} \quad \int |\omega'(v)| dv < \infty.$$

Teorema 3.2.4 (*Representación asintótica de $MISE(f_n^P(\cdot))$*) Bajo las condiciones (K1), (H3), (H4), (f2), (p2), (v1), (v4) y (ω_1), la función $MISE$ del estimador presuavizado $f_n^P(\cdot)$ admite las siguientes representaciones:

a) Si $b/s \rightarrow 0$ entonces

$$\begin{aligned} MISE(f_n^P(\cdot)) &= \int \left(\frac{1}{2} s^2 M(v) + b^2 N(v) \right)^2 \omega(v) dv \\ &\quad + \frac{1}{ns} c_K \int A(v) \omega(v) dv + O(s^6) + o((ns)^{-1}). \end{aligned}$$

b) Si $b/s \rightarrow L > 0$ entonces

$$\begin{aligned} MISE(f_n^P(\cdot)) &= \int \left(\frac{1}{2} s^2 M(v) + b^2 N(v) \right)^2 \omega(v) dv \\ &\quad + \frac{1}{ns} \left[c_K \int A(v) p(v) \omega(v) dv + A_K(L) \int A(v) (1 - p(v)) \omega(v) dv \right] \\ &\quad + O(s^6) + o((ns)^{-1}). \end{aligned}$$

c) Si $b/s \rightarrow \infty$ entonces

$$\begin{aligned} MISE(f_n^P(\cdot)) &= \int \left(\frac{1}{2} s^2 M(v) + b^2 N(v) \right)^2 \omega(v) dv \\ &\quad + s^2 b^2 \int M(v) N(v) \omega(v) dv + \frac{1}{ns} c_K \int A(v) p(v) \omega(v) dv \\ &\quad + \frac{1}{nb} c_K \int A(v) (1 - p(v)) \omega(v) dv + O(b^6) + o((nb)^{-1}), \end{aligned}$$

donde las funciones $M(\cdot)$, $N(\cdot)$, $A(\cdot)$ y $A_K(\cdot)$ están dadas en (3.42), (3.31) y (3.24) respectivamente.

Demostración.

Bajo las condiciones del teorema 3.2.1, se tiene la siguiente descomposición del estimador presuavizado de la función de densidad:

$$f_n^P(t) = f(t) + \beta_n(t) + \sigma_n(t) + e_n(t).$$

La clásica descomposición del *MISE* en el sesgo al cuadrado integrado y varianza integrada da lugar a la siguiente relación:

$$\begin{aligned} MISE(f_n^P(\cdot)) &= \mathbb{E} \left[\int (f_n^P(v) - f(v))^2 \omega(v) dv \right] \\ &= \mathbb{E} \left[\int (\beta_n(v) + \sigma_n(v) + e_n(v))^2 \omega(v) dv \right] \\ &= S_1 + S_2 + S_3 + 2(S_4 + S_5 + S_6) \end{aligned}$$

siendo

$$\begin{aligned} S_1 &= \int \beta_n^2(v) \omega(v) dv \\ S_2 &= \int \mathbb{E}[\sigma_n^2(v)] \omega(v) dv \\ S_3 &= \int \mathbb{E}[e_n^2(v)] \omega(v) dv \\ S_4 &= \int \beta_n(v) \mathbb{E}[e_n(v)] \omega(v) dv \\ S_5 &= \int \mathbb{E}[\sigma_n(v) e_n(v)] \omega(v) dv \\ S_6 &= \int \beta_n(v) \mathbb{E}[\sigma_n(v)] \omega(v) dv. \end{aligned}$$

El término S_1 representa la parte principal del sesgo al cuadrado integrado, y S_2 la parte dominante de la varianza integrada. Se probará que los otros tres términos, S_3 , S_4 y S_5 , son de orden despreciable con respecto a la parte dominante de la representación, proveniente de S_1 , S_2 y S_6 .

Con respecto a S_1 se tiene lo siguiente:

$$S_1 = \int \beta_n^2(v) \omega(v) dv = \int \left[\int (f(v-us) - f(v)) K(u) du \right]^2 \omega(v) dv.$$

Bajo la condición (f2), podemos efectuar un desarrollo de Taylor con resto en forma integral, y expresar $f(v-us)$ de la siguiente manera:

$$f(v - us) = f(v) - usf'(v) + \frac{1}{2}s^2u^2f''(v) - \frac{1}{6}s^3u^3f'''(v) \\ + \frac{1}{6}s^4u^4 \int_0^1 (1-z)^3 f^{(4)}(v - usz) dz.$$

Sustituyendo esta expresión en S_1 y dado que, por (K1), $K(\cdot)$ es una función simétrica,

$$S_1 = \frac{1}{4}s^4 d_K^2 \int f''(v)^2 \omega(v) dv \\ + \frac{1}{6}s^6 d_K \int f''(v) \int u^4 K(u) \int_0^1 (1-z)^3 f^{(4)}(v - suz) dz du \omega(v) dv \\ + \frac{1}{36}s^8 \int \left[\int u^4 K(u) \int_0^1 (1-z)^3 f^{(4)}(v - suz) dz du \right]^2 \omega(v) dv.$$

El primer término es dominante, mientras que el valor absoluto del segundo se puede acotar por

$$\frac{1}{6}s^6 d_K \frac{1}{4} \left(\int u^4 K(u) du \right) \sup_{t \in [0, t_H]} |f^{(4)}(t)| \sup_{t \in [0, t_H]} |f''(t)| \int \omega(v) dv,$$

y el valor absoluto del tercero por

$$\frac{1}{36}s^8 \frac{1}{4^2} \left(\int u^4 K(u) du \right)^2 \sup_{t \in [0, t_H]} |f^{(4)}(t)|^2 \int \omega(v) dv.$$

De este modo se llega a

$$S_1 = \frac{1}{4}s^4 d_K^2 \int f''(v)^2 \omega(v) dv + O(s^6) = \frac{1}{4}s^4 \int M^2(v) \omega(v) dv + O(s^6).$$

A continuación, obtendremos un desarrollo para S_2 . Dado que el momento de orden dos de $\sigma_n(\cdot)$ se puede expresar como

$$\mathbb{E}[\sigma_n^2(t)] = \text{Var}(\sigma_n(t)) + \mathbb{E}[\sigma_n(t)]^2,$$

obtenemos la siguiente descomposición para $S_2 = S_{2,1} + S_{2,2}$, donde

$$S_{2,1} = \int \text{Var}(\sigma_n(t)) \omega(v) dv \quad \text{y} \quad S_{2,2} = \int \mathbb{E}[\sigma_n(v)]^2 \omega(v) dv.$$

Por el teorema 3.2.2, se tiene la siguiente expresión para $\text{Var}(\sigma_n(t))$:

a) Si $b/s \rightarrow 0$

$$\text{Var}[\sigma_n(t)] = \frac{1}{ns} A(t) c_K + \frac{1}{n} B(t) + o(n^{-1}).$$

b) Si $b/s \rightarrow L \in (0, \infty)$

$$\text{Var} [\sigma_n(t)] = \frac{1}{ns} [A(t)p(t)c_K + A(t)(1-p(t))A_K(L)] + \frac{1}{n}B(t) + O\left(\frac{s}{n}\right).$$

c) Si $b/s \rightarrow \infty$

$$\text{Var} [\sigma_n(t)] = \frac{1}{ns}A(t)p(t)c_K + \frac{1}{nb}A(t)(1-p(t))c_K + \frac{1}{n}B(t) + o(n^{-1}),$$

siendo $A(\cdot)$ y $B(\cdot)$ las funciones dadas en (3.31) y (3.32) respectivamente.

Todos estos términos despreciables provienen de desarrollos de Taylor, de modo que bajo las condiciones (H2), (p2) y (f2), la expresión del término $S_{2,1}$ es:

a) Si $b/s \rightarrow 0$

$$S_{2,1} = \frac{1}{ns}c_K \int A(v)\omega(v)dv + \frac{1}{n} \int B(v)\omega(v)dv + o(n^{-1}).$$

b) Si $b/s \rightarrow L \in (0, \infty)$

$$\begin{aligned} S_{2,1} &= \frac{1}{ns} \left[c_K \int A(v)p(v)\omega(v)dv + A_K(L) \int A(v)(1-p(v))\omega(v)dv \right] \\ &\quad + \frac{1}{n} \int B(v)\omega(v)dv + O\left(\frac{s}{n}\right). \end{aligned}$$

c) Si $b/s \rightarrow \infty$

$$\begin{aligned} S_{2,1} &= \frac{1}{ns}c_K \int A(v)p(v)\omega(v)dv + \frac{1}{nb}c_K \int A(v)(1-p(v))\omega(v)dv \\ &\quad + \frac{1}{n} \int B(v)\omega(v)dv + o(n^{-1}). \end{aligned}$$

A continuación, calcularemos $S_{2,2} = \int \mathbb{E}[\sigma_n(v)]^2 \omega(v)dv$. Sustituyendo en $S_{2,2}$ la expresión de $\mathbb{E}[\sigma_n(v)]$ dada en (3.38) (paso (a)), y desarrollando el cuadrado (paso (b)) resulta

$$\begin{aligned} S_{2,2} &= \int \mathbb{E}[\sigma_n(v)]^2 \omega(v)dv \\ &\stackrel{(a)}{=} \int \left(\mathbb{E}[\varepsilon_1(v)]' + s^2 \iint_0^1 \mathbb{E}[\varepsilon_1(v-svz)]''' z^2 K(z) dx dz \right)^2 \omega(v)dv \end{aligned} \quad (3.43)$$

$$\begin{aligned}
& \stackrel{(b)}{=} \int (\mathbb{E} [\varepsilon_1 (v)]')^2 \omega (v) dv \\
& + 2s^2 \int \mathbb{E} [\varepsilon_1 (v)]' \iint_0^1 \mathbb{E} [\varepsilon_1 (v - svz)]''' z^2 K (z) \omega (v) dx dz dv \\
& + s^4 \int \left(\iint_0^1 \mathbb{E} [\varepsilon_1 (v - szx)]''' z^2 K (z) dx dz \right)^2 \omega (v) dv.
\end{aligned}$$

El término dominante de $S_{2,2}$, dado por el primer sumando en (3.43), es igual a

$$\int (\mathbb{E} [\varepsilon_1 (v)]')^2 \omega (v) dv,$$

donde

$$\mathbb{E} [\varepsilon_1 (t)]' = b^2 d_K [(1 - F (t)) \alpha (t)]' + O (b^4) = b^2 N (t) + O (b^4),$$

siendo $N (\cdot)$ la función dada en (3.42). Acotamos el valor absoluto del segundo sumando en (3.43) de la siguiente manera:

$$\begin{aligned}
& 2s^2 \left| \int \mathbb{E} [\varepsilon_1 (v)]' \iint_0^1 \mathbb{E} [\varepsilon_1 (v - svz)]''' z^2 K (z) \omega (v) dx dz dv \right| \\
& \leq 2s^2 \sup_{t \in [0, t_H]} |\mathbb{E} [\varepsilon_1 (t)]'| \sup_{t \in [0, t_H]} |\mathbb{E} [\varepsilon_1 (t)]'''| d_K \int \omega (v) dv.
\end{aligned}$$

Utilizando las cotas (3.39) y (3.40) podemos concluir que el valor absoluto del segundo sumando en (3.43) es de orden $O (s^2 b^4)$.

De igual modo, y repitiendo los mismos argumentos, se consigue con facilidad una cota para el valor absoluto del último sumando en (3.43):

$$\begin{aligned}
& s^4 \left| \int \left(\iint_0^1 \mathbb{E} [\varepsilon_1 (v - szx)]''' z^2 K (z) dx dz \right)^2 \omega (v) dv \right| \\
& \leq s^4 \left(\sup_{t \in [0, t_H]} \mathbb{E} [\varepsilon_1 (t)]''' \right)^2 d_K \int \omega (v) dv.
\end{aligned}$$

Como consecuencia,

$$S_{2,2} = b^4 \int N^2 (v) \omega (v) dv + O (b^6) + O (s^2 b^4).$$

Resumiendo, la expresión final del término S_2 es:

a) Si $b/s \rightarrow 0$

$$\begin{aligned} S_2 &= \frac{1}{ns} c_K \int A(v) \omega(v) dv + \frac{1}{n} \int B(v) \omega(v) dv \\ &\quad + b^4 d_K^2 \int ([(1-F(v))\alpha(v)]')^2 \omega(v) dv + O(b^6) + O(s^2 b^4) + o(n^{-1}). \end{aligned}$$

b) Si $b/s \rightarrow L \in (0, \infty)$

$$\begin{aligned} S_2 &= \frac{1}{ns} \left[c_K \int A(v) p(v) \omega(v) dv + A_K(L) \int A(v) (1-p(v)) \omega(v) dv \right] \\ &\quad + \frac{1}{n} \int B(v) \omega(v) dv + b^4 d_K^2 \int ([(1-F(v))\alpha(v)]')^2 \omega(v) dv \\ &\quad + O(b^6) + O(s^2 b^4) + O\left(\frac{s}{n}\right). \end{aligned}$$

c) Si $b/s \rightarrow \infty$

$$\begin{aligned} S_2 &= \frac{1}{ns} c_K \int A(v) p(v) \omega(v) dv + \frac{1}{nb} c_K \int A(v) (1-p(v)) \omega(v) dv \\ &\quad + \frac{1}{n} \int B(v) \omega(v) dv + b^4 d_K^2 \int ([(1-F(v))\alpha(v)]')^2 \omega(v) dv \\ &\quad + O(b^6) + O(s^2 b^4) + o(n^{-1}). \end{aligned}$$

A continuación, demostraremos que los términos S_3, S_4 y S_5 son despreciables. Para S_3 , se consigue una cota asintótica utilizando la tasa para el error $e_n(\cdot)$ dada en el teorema 3.2.1:

$$|S_3| \leq \sup_{0 \leq t \leq t_H} \mathbb{E} [e_n^2(t)] \int \omega(v) dv = O\left(\left(b^4 s^{-1} + (nbs)^{-1}\right)^2\right).$$

Los órdenes de S_4 y S_5 se extraen por aplicación de la desigualdad de Cauchy-Schwarz:

$$\begin{aligned} |S_4| &\leq S_1^{1/2} S_3^{1/2} = O\left(s^2 \left(b^4 s^{-1} + (nbs)^{-1}\right)\right) = O\left(b^4 s + (nb)^{-1} s\right). \\ |S_5| &\leq S_2^{1/2} S_3^{1/2} = O\left(\left(b^2 + (ns)^{-1/2}\right) \left(b^4 + (nb)^{-1}\right) s^{-1}\right) \\ &= O\left(b^6 s^{-1} + n^{-1} b s^{-1} + n^{-1/2} b^4 s^{-3/2} + n^{-3/2} b^{-1} s^{-3/2}\right) \end{aligned}$$

Bajo la condición (v4) podemos concluir que

$$S_3 = o\left((ns)^{-1/2}\right), \quad S_4 = o\left((ns)^{-1/2}\right) \quad \text{y} \quad S_5 = o\left((ns)^{-1/2}\right).$$

Y por último, obtenemos el término dominante de S_6 . Para ello recordamos que $\beta_n(\cdot)$ admite la siguiente representación

$$\begin{aligned} \beta_n(t) &= \int f(t - vs) K(v) dv - f(t) \\ &= \frac{1}{2}s^2 d_K f''(t) + \frac{1}{6}s^4 \int v^4 K(v) f^{(4)}(t - vs\omega_{t,s,v}) dv \end{aligned}$$

donde $\omega_{s,t,v} \in (0, 1)$ depende de t, s y v .

Por su parte, la esperanza de $\sigma_n(\cdot)$ es igual a

$$\mathbb{E}[\sigma_n(t)] = \mathbb{E}[\varepsilon_1(t)]' + s^2 \int \mathbb{E}[\varepsilon_1(t - sv\theta_{s,t,v})]''' v^2 K(v) dv.$$

donde $\theta_{s,t,v} \in (0, 1)$ depende de t, s y v . Entonces, podemos expresar el término S_6 de la siguiente manera:

$$\begin{aligned} S_6 &= \int \beta_n(v) \mathbb{E}[\sigma_n(v)] \omega(v) dv \\ &= \int \left(\frac{1}{2}s^2 d_K f''(v) + \frac{1}{6}s^4 \int v_1^4 K(v_1) f^{(4)}(v - v_1 s\omega_{t,s,v_1}) dv_1 \right) \\ &\quad \times \left(\mathbb{E}[\varepsilon_1(v)]' + s^2 \int \mathbb{E}[\varepsilon_1(t - sv_2\theta_{s,t,v_2})]''' v_2^2 K(v_2) dv_2 \right) \omega(z) dz \\ &= \frac{1}{2}s^2 d_K \int f''(v) \mathbb{E}[\varepsilon_1(v)]' \omega(v) dv \\ &\quad + \frac{1}{6}s^4 \iint \mathbb{E}[\varepsilon_1(v)]' v_1^4 K(v_1) f^{(4)}(v - v_1 s\omega_{t,s,v_1}) dv_1 \omega(v) dv \\ &\quad + \frac{1}{2}s^4 d_K \iint f''(v) \mathbb{E}[\varepsilon_1(t - sv_2\theta_{s,t,v_2})]''' v_2^2 K(v_2) dv_2 \omega(v) dv \\ &\quad + \frac{1}{6}s^6 \iiint \mathbb{E}[\varepsilon_1(t - sv_2\theta_{s,t,v_2})]''' f^{(4)}(v - v_1 s\omega_{t,s,v_1}) K(v_2) v_1^4 K(v_1) \omega(v) dv_1 dv_2 dv. \end{aligned}$$

Repitiendo los razonamientos seguidos en la acotación del término $S_{2,2}$ y teniendo en cuenta (3.39) y (3.40) podemos concluir que

$$\begin{aligned} S_6 &= \frac{1}{2}d_K^2 s^2 b^2 \int f''(v) [(1 - F(v)) \alpha(v)]' \omega(v) dv + O(s^4 b^2) + O(s^2 b^4) \\ &= \frac{1}{2}s^2 b^2 \int M(v) N(v) \omega(v) dv + O(s^4 b^2) + O(s^2 b^4). \end{aligned}$$

Recopilando las cotas asintóticas para $S_i, i = 1, \dots, 6$ y bajo las hipótesis (v4), se tiene la representación dada en el presente teorema. ■

Observación 3.2.4

Cuando la ventana de presuavizado es muy pequeña ($b/s \rightarrow 0$), esta expresión del MISE de $f_n^P(\cdot)$ se reduce a la correspondiente del estimador tipo núcleo de la densidad con pesos Kaplan-Meier:

$$\begin{aligned} MISE(f_n^{KM}(\cdot)) &= \frac{1}{4}s^4 d_K^2 \int f''(v)^2 \omega(v) dv + \frac{1}{ns} c_K \int f(v) \frac{1-F(v)}{1-H(v)} \omega(v) dv \\ &+ o((ns)^{-1}) + O(s^6). \end{aligned}$$

Observación 3.2.5

Asimismo, esta expresión del MISE de $f_n^P(\cdot)$ se reduce a la correspondiente del estimador de Parzen-Rosenblatt de la densidad cuando no hay censura:

$$MISE(f_n^{PR}(\cdot)) = \frac{1}{4}s^4 d_K^2 \int f''(v)^2 \omega(v) dv + \frac{1}{ns} c_K + O(n^{-1}) + O(s^6).$$

3.3. Estudio de simulación

Hemos realizado un estudio de simulación para comparar el comportamiento en la práctica de los estimadores de la función de densidad con pesos Kaplan-Meier $f_n^{KM}(\cdot)$ y pesos presuavizados $f_n^P(\cdot)$. Los modelos con los que hemos trabajado son los mismos cuatro modelos usados en el estudio de simulación para la función de distribución.

En los modelos 1, 2 y 4 la variable observable sigue la distribución $Z \stackrel{d}{=} U[0, 1]$, mientras que en el modelo 3 se tiene que $Z \stackrel{d}{=} W(1.75, 5)$, donde $W(a, b)$ denota la distribución de Weibull con parámetro de escala a y parámetro de forma b .

Tabla 3.1. Funciones de densidad $f(\cdot)$ y probabilidad condicional de no censura $p(\cdot)$ en los modelos 1-4.

Función de densidad	Función $p(t)$
$f(t) = (1-t)^{3/8} e^{-1/4t^2} \left(\frac{3}{8}(1-t)^{-1} + \frac{1}{2}t \right)$	$p(t) = \frac{1}{2} \left[1 - \left(t - \frac{1}{2} \right)^2 \right]$
$f(t) = 3te^{-3/2t^2}$	$p(t) = \left[\frac{3}{4} - 3 \left(t - \frac{1}{2} \right)^2 \right] \mathbf{1}_{\{0 \leq t \leq 1\}}$
$f(t) = \frac{5}{2}t^4 e^{-t^5} + 80t^4 e^{-32t^5}$	$p(t) = \frac{\frac{5}{2}e^{-t^5} + 80e^{-32t^5}}{41.0327(e^{-t^5} + e^{-32t^5})}$
$f(t) = (20t^3 - 12t^2 + 1) e^{-5t^4 + 4t^3 - t}$	$p(t) = (20t^3 - 12t^2 + 1)(1-t)$

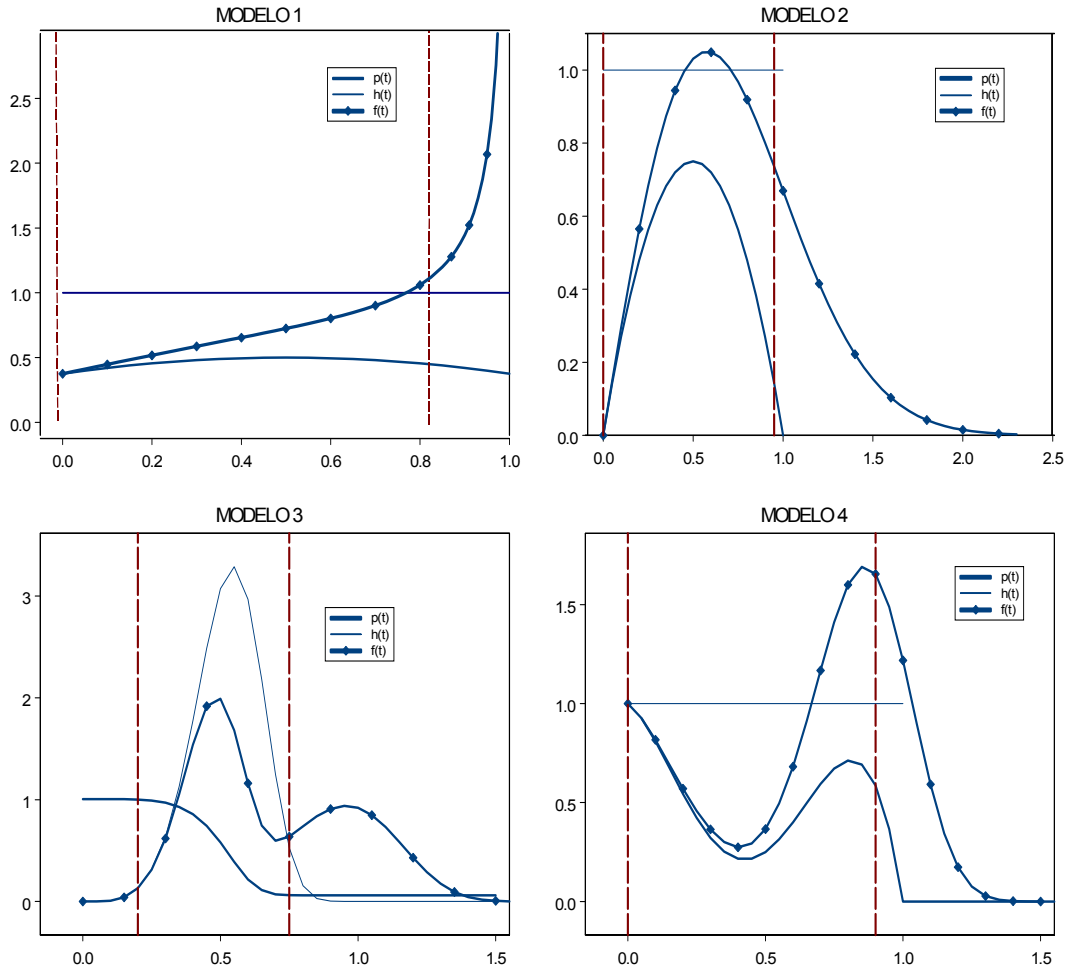


Figura 3.1. Funciones $p(\cdot)$, $h(\cdot)$ y $f(\cdot)$ para los modelos 1-4. La función peso $\omega(\cdot)$ es la función indicadora del intervalo marcado por las líneas verticales.

En la tabla 3.1 se muestran las funciones de densidad teóricas a estimar, $f(\cdot)$, y las probabilidades condicionales de no censura $p(\cdot)$ para cada uno de los modelos. Además, en la figura 3.1 se representan la función de densidad $f(\cdot)$ de la variable de interés (línea con puntos), la función de densidad de la variable observable $h(\cdot)$ (línea continua gruesa) y la función de probabilidad condicional de no censura $p(\cdot)$ (línea continua fina) para cada uno de los modelos. Para todas ellas, la función peso $\omega(\cdot)$ es la función indicadora del intervalo cuyos extremos vienen dados por las líneas verticales. Estos intervalos son los mismos que los tomados en el estudio de simulación del capítulo anterior.

Para cada modelo se ha calculado la eficiencia relativa del estimador de Kaplan-Meier $f_n^{KM}(\cdot)$ con respecto al estimador presuavizado $f_n^P(\cdot)$ como el cociente entre $MISE(f_n^P(\cdot))$ y $MISE(f_n^{KM}(\cdot))$ en función de las ventanas de presuavizado s y b , necesarias para obtener $f_n^P(\cdot)$, y de la ventana s^{KM} con la que se calcula $f_n^{KM}(\cdot)$:

$$RE^{KM,P}(s, b; s^{KM}) = \frac{MISE(f_n^P(\cdot))}{MISE(f_n^{KM}(\cdot))}.$$

Valores de $RE^{KM,P}(s, b; s^{KM})$ menores que 1 indican una mayor eficiencia del estimador presuavizado $f_n^P(\cdot)$ calculado con el par de ventanas (s, b) con respecto al estimador tipo núcleo con pesos Kaplan-Meier $f_n^{KM}(\cdot)$ para la correspondiente ventana s^{KM} . Para ello se ha simulado mediante Montecarlo el $MISE$ de ambos estimadores a partir de $m = 500$ muestras de tamaños $n = 50, 100, 200$ y 1000 .

En el caso del estimador de Kaplan-Meier, se ha obtenido el valor mínimo del $MISE$, $MISE_{\min}(f_n^{KM}(\cdot))$, y la ventana de suavización s_{MISE}^{KM} en la que alcanza ese valor. A continuación, se ha calculado la eficiencia relativa de $f_n^{KM}(\cdot)$ con respecto a $f_n^P(\cdot)$ tomando $s^{KM} = s_{MISE}^{KM}$. Los resultados se presentan en las figuras 3.2-3.5, en las que valores de $RE^{KM,P}(s, b; s_{MISE}^{KM})$ por debajo del valor 1 indican un mejor comportamiento para las correspondientes ventanas s y b del estimador presuavizado $f_n^P(\cdot)$ frente al estimador tipo núcleo con pesos Kaplan-Meier $f_n^{KM}(\cdot)$ en el caso óptimo en el que se conociese el valor exacto de la ventana óptima s_{MISE}^{KM} . Se marca con un aspa (\times) el punto que indica las ventanas $(s, b)_{MISE}$ donde dicho cociente alcanza su mínimo.

Se puede comprobar con facilidad el mejor comportamiento del estimador presuavizado de la densidad para un amplio rango de ventanas. Además, la elección de la ventana de presuavizado b tiene un efecto de segundo orden frente a la de la ventana de suavizado s , sobre todo en el modelo 1. Es decir, una vez elegida la ventana de suavizado s la eficiencia de $f_n^P(\cdot)$ con respecto a $f_n^{KM}(\cdot)$ varía muy poco para ventanas de presuavizado b muy diferentes entre sí. Esto es debido a la forma aproximadamente plana de la función $p(\cdot)$ en dicho modelo, de modo que no varía mucho la estimación de $p(\cdot)$ usando diferentes ventanas b , y por tanto, tampoco la estimación final de $f(\cdot)$.

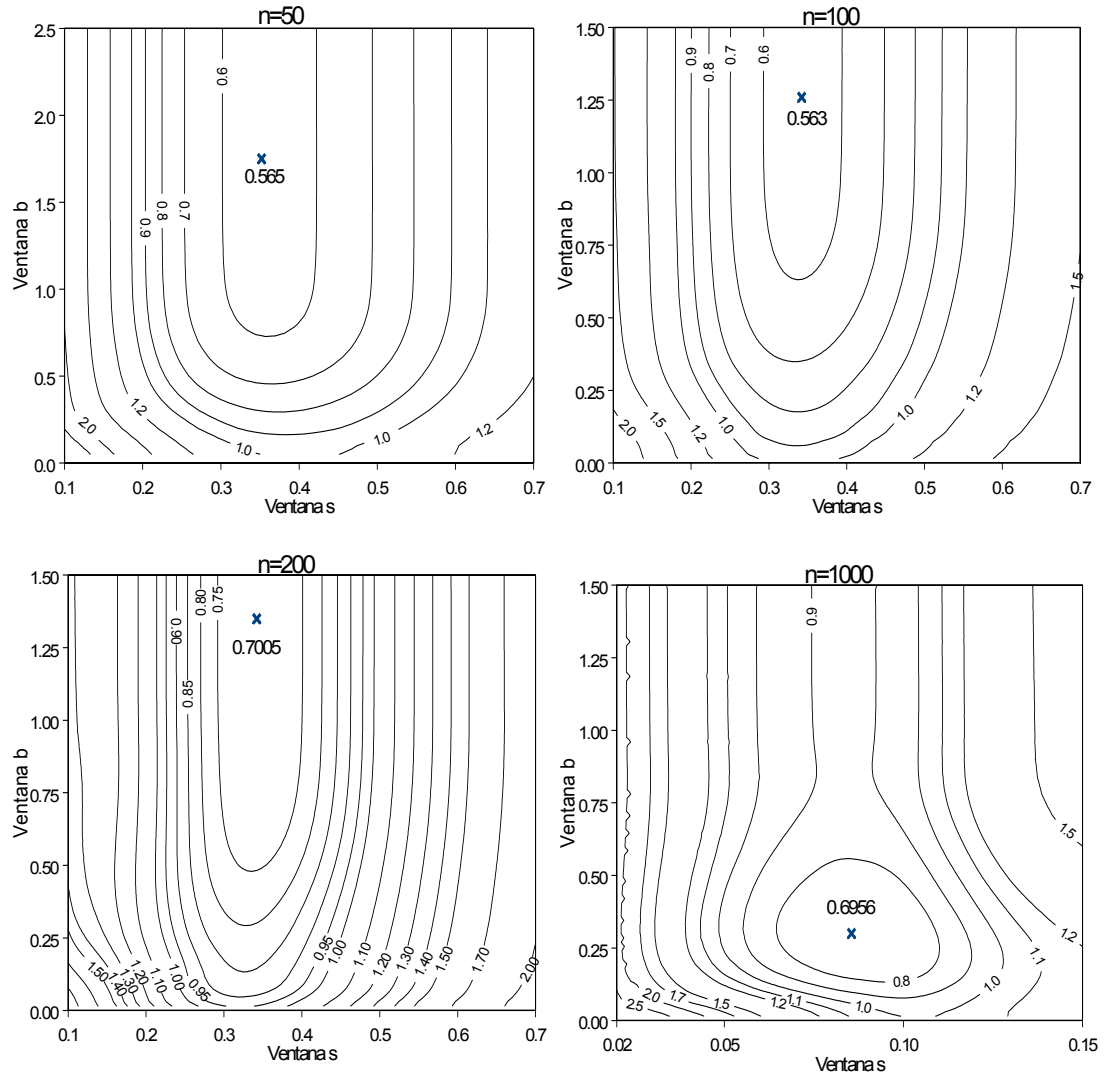


Figura 3.2. Eficiencia relativa, $RE^{KM,P}(s, b; s_{MISE}^{KM})$, para el modelo 1 con tamaños muestrales $n = 50, 100, 200$ y 1000 . El valor mínimo está marcado con un aspa (\times).

Tabla 3.2 Eficiencia relativa de $f_n^{KM}(\cdot)$ con respecto a $f_n^P(\cdot)$ obtenido por simulación para el modelo 1, usando ventanas óptimas.

	$n = 50$	$n = 100$	$n = 200$	$n = 1000$
$RE^{KM,P}(s_{MISE}^P, b_{MISE}^P; s_{MISE}^{KM})$	0.5650	0.5630	0.7005	0.6956

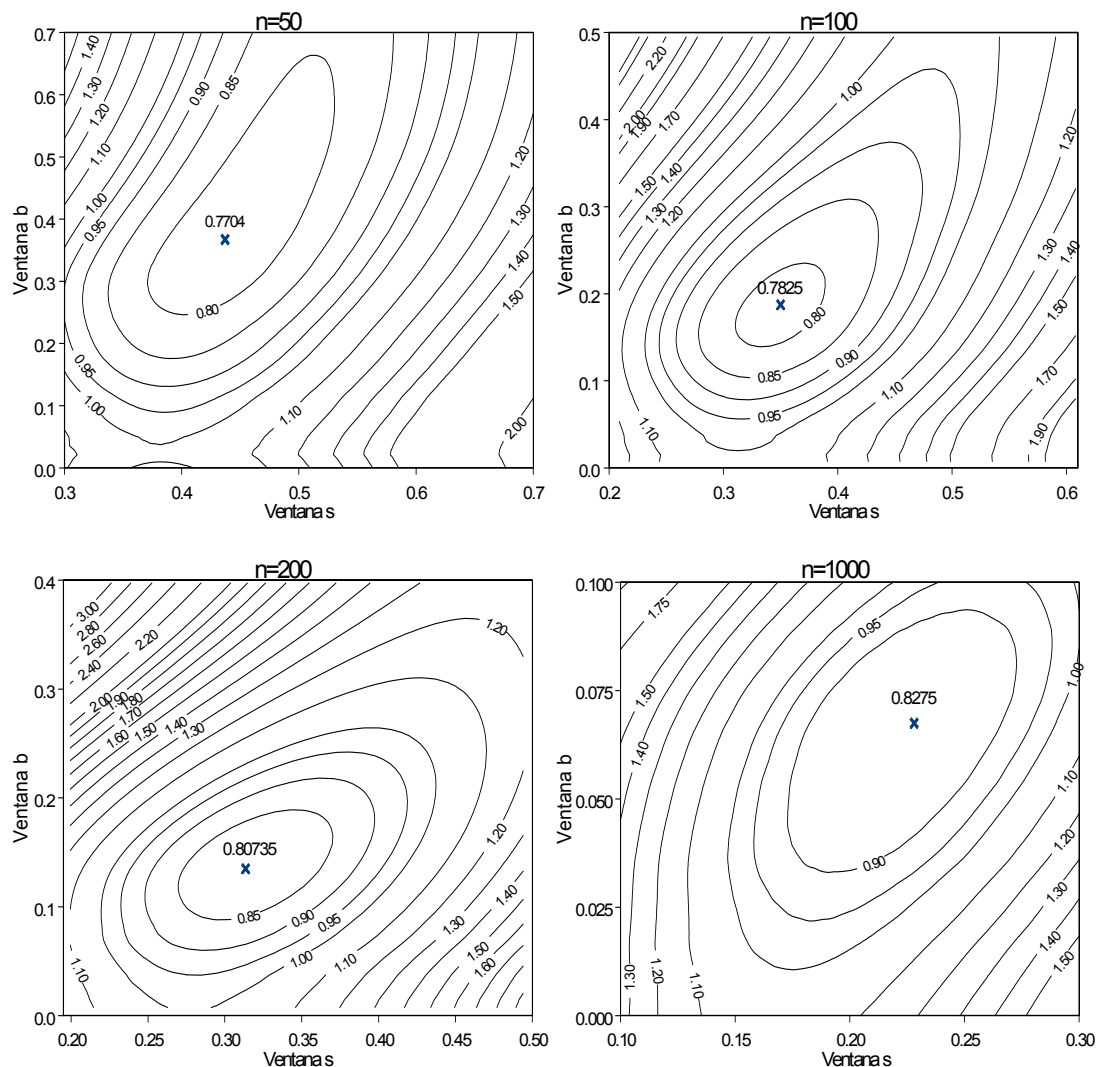


Figura 3.3. Eficiencia relativa, $RE^{KM,P}(s, b; s_{MISE}^{KM})$, para el modelo 2 con tamaños muestrales $n = 50, 100, 200$ y 1000 . El valor mínimo está marcado con un aspa (\times).

Tabla 3.3. Eficiencia relativa de $f_n^{KM}(\cdot)$ con respecto a $f_n^P(\cdot)$ obtenido por simulación para el modelo 2, usando ventanas óptimas.

	$n = 50$	$n = 100$	$n = 200$	$n = 1000$
$RE^{KM,P}(s_{MISE}^P, b_{MISE}^P; s_{MISE}^{KM})$	0.7704	0.7825	0.8073	0.8275

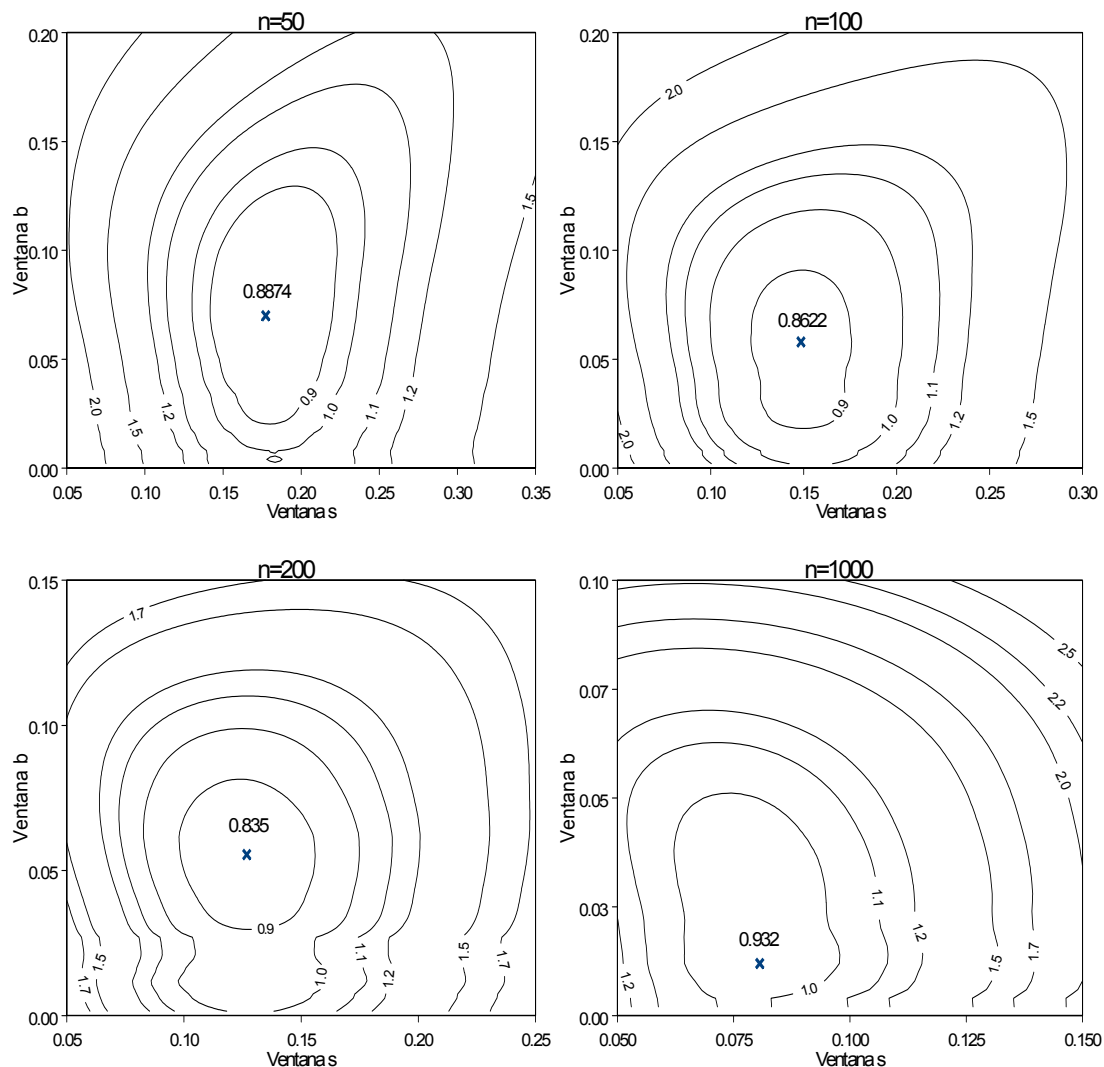


Figura 3.4. Eficiencia relativa, $RE^{KM,P}(s, b; s_{MISE}^{KM})$, para el modelo 3 con tamaños muestrales $n = 50, 100, 200$ y 1000 . El valor mínimo está marcado con un aspa (\times).

Tabla 3.4. Eficiencia relativa de $f_n^{KM}(\cdot)$ con respecto a $f_n^P(\cdot)$ obtenido por simulación para el modelo 3, usando ventanas óptimas.

	$n = 50$	$n = 100$	$n = 200$	$n = 1000$
$RE^{KM,P}(s_{MISE}^P, b_{MISE}^P; s_{MISE}^{KM})$	0.8874	0.8622	0.8350	0.9320

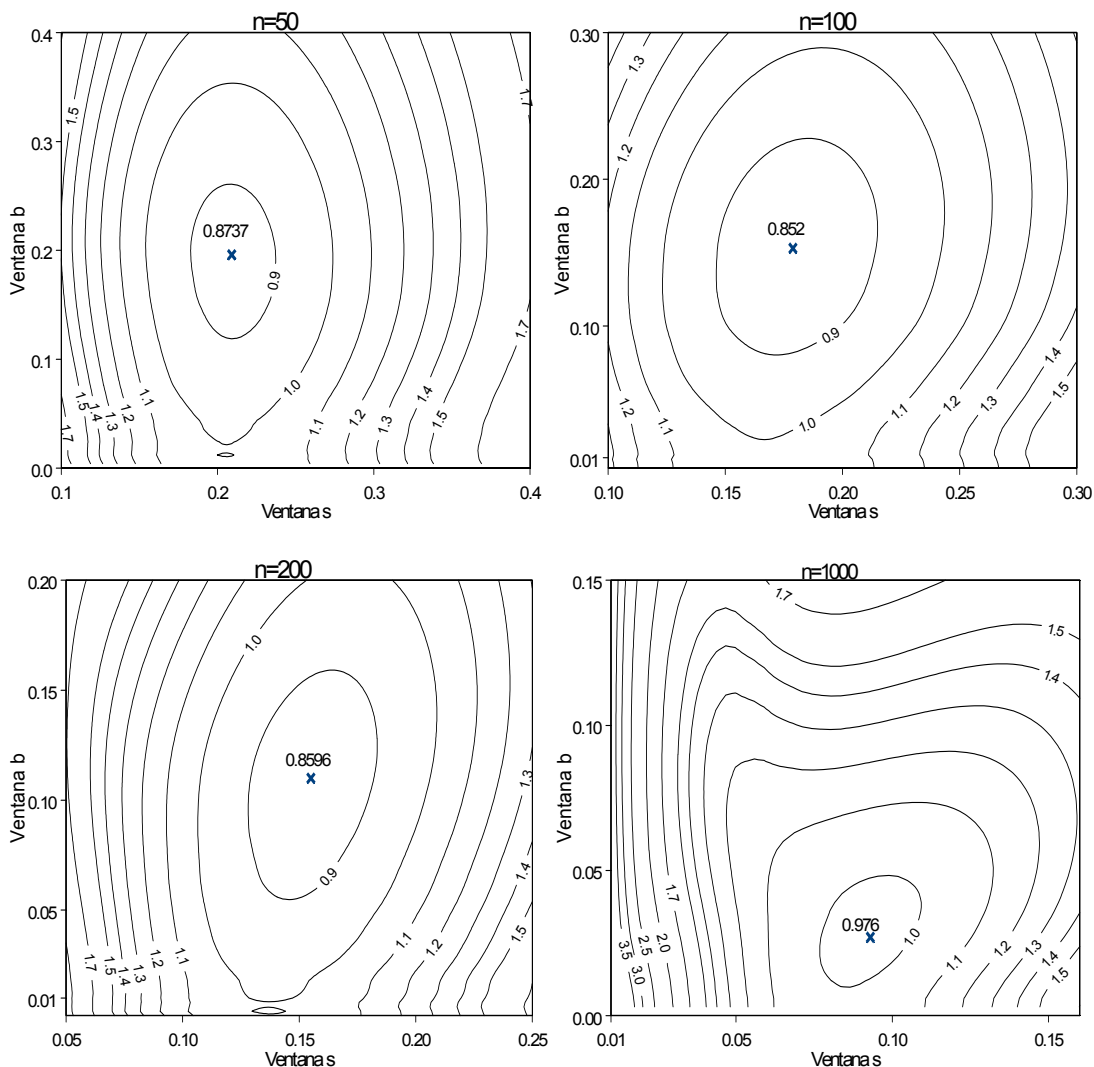


Figura 3.5. Eficiencia relativa, $RE^{KM,P}(s, b; s_{MISE}^{KM})$, para el modelo 4 con tamaños muestrales $n = 50, 100, 200$ y 1000 . El valor mínimo está marcado con un aspa (\times).

Tabla 3.5. Eficiencia relativa de $f_n^{KM}(\cdot)$ con respecto a $f_n^P(\cdot)$ obtenido por simulación para el modelo 4, usando ventanas óptimas.

	$n = 50$	$n = 100$	$n = 200$	$n = 1000$
$RE^{KM,P}(s_{MISE}^P, b_{MISE}^P; s_{MISE}^{KM})$	0.874	0.852	0.859	0.976

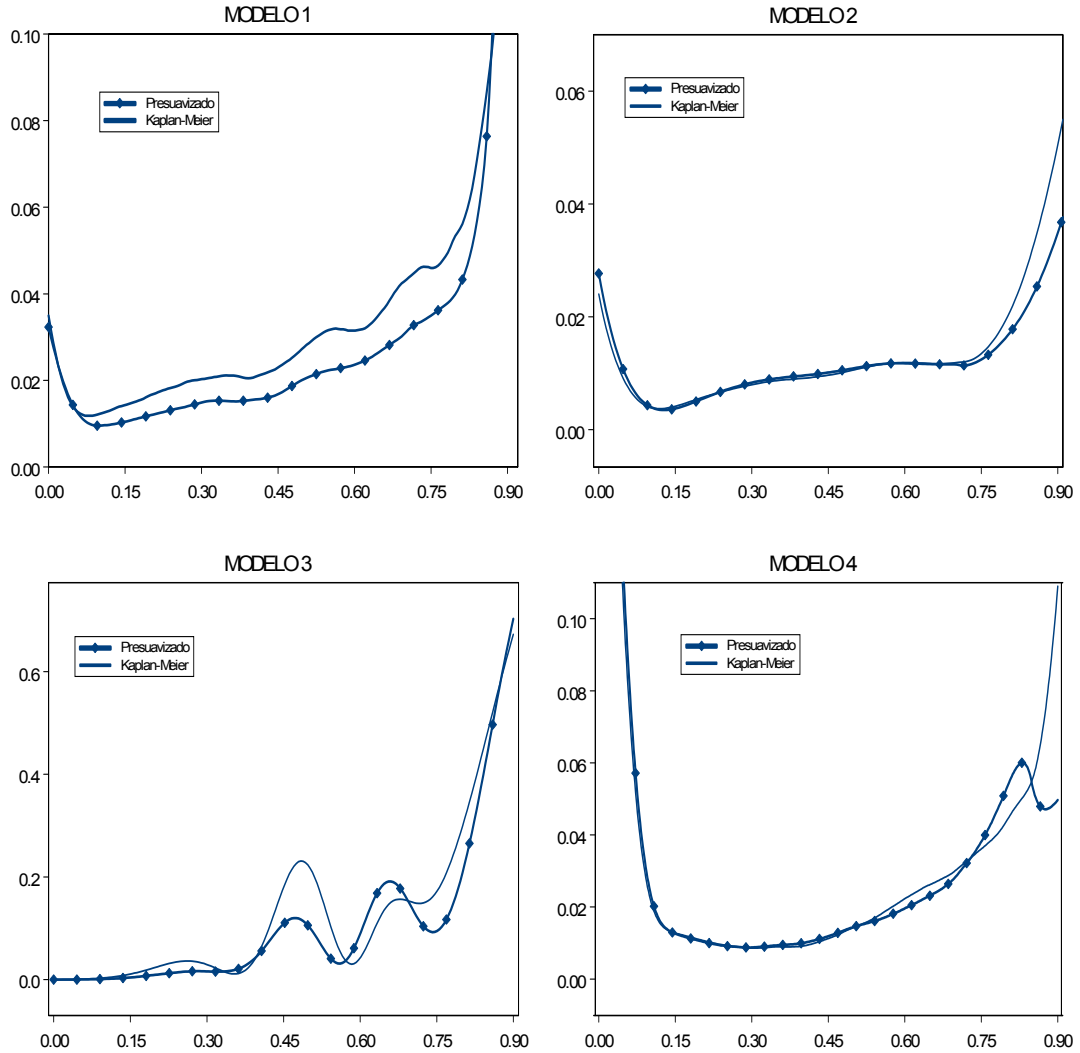


Figura 3.6 Error cuadrático medio del estimador presuavizado $f_n^P(\cdot)$ (línea con puntos) y del estimador con pesos Kaplan-Meier (línea continua) en los modelos 1-4 con $n = 200$.

En las tablas 3.2-3.5 figuran los valores mínimos de $MISE(f_n^P(\cdot))$ y $MISE(f_n^{KM}(\cdot))$ obtenidos por simulación para los cuatro modelos y con tamaños muestrales $n = 50, 100, 200$ y 1000 respectivamente. Estos valores aumentan con el tamaño muestral hasta aproximadamente 0.7 en el modelo 1, y hasta $0.83, 0.93$ y 0.98 en los demás modelos. Como se verá en el siguiente capítulo, esto se debe a la eficiencia de $f_n^P(\cdot)$ con respecto a $f_n^{KM}(\cdot)$ de primer orden para el modelo 1, y de tercer orden para los demás modelos.

Además, para un tamaño muestral igual a $n = 200$, se ha estimado el error cuadrático medio:

$$MSE(\hat{f}_n(t)) = \mathbb{E} \left[\left(\hat{f}_n(t) - f(t) \right)^2 \right]$$

para cada uno de los modelos anteriores, con las ventanas $(s, b)_{MISE}$ y s_{MISE}^{KM} que minimizan $MISE(f_n^P(\cdot))$ y $MISE(f_n^{KM}(\cdot))$ respectivamente, obtenidas anteriormente por simulación. En la figura 3.6 se puede comprobar cómo, eligiendo adecuadamente las ventanas de suavizado y presuavizado (s, b) , el estimador $f_n^P(\cdot)$ comete un error menor que $f_n^{KM}(\cdot)$ en una amplia región del intervalo de observación. Este mejor comportamiento de $f_n^P(\cdot)$ es muy llamativo en el modelo 1, en el que la estimación de $f(\cdot)$ es mejor mediante la presuavización en la práctica totalidad del intervalo de observación.

Capítulo 4

Selección del parámetro de suavización

Mientras la elección en la práctica de la función núcleo no es muy importante, la selección del parámetro de suavización o ventana es un problema crucial, y el estudio de procedimientos que permitan elegir la ventana óptima según algún criterio constituye uno de los problemas más importantes en la estadística no paramétrica. En todos ellos se escoge el parámetro de suavización de manera que minimice alguna medida de distancia entre la función de densidad verdadera y su estimador. Sin embargo, no existe ningún método que proporcione una expresión para la ventana óptima de manera que en su cálculo sólo se utilice la información proporcionada por la muestra, puesto que, generalmente, la función de discrepancia depende de la curva teórica, inobservable, y por tanto también la expresión de la ventana.

Existe una gran variedad de técnicas de selección automática de la ventana: métodos plug-in, de validación cruzada (o cross-validation) suavizada o no, métodos basados en el remuestreo bootstrap, de los puntos de inflexión, de doble núcleo ... Además, en función de la medida de distancia, el selector puede ser de dos tipos: global o puntual, según el interés se centre en la estimación de la curva por entero o en un punto particular.

En este capítulo, se estudiará la expresión de las ventanas que minimizan, como medida global de error, el error cuadrático medio integrado, y puesto que dependerán de la función de densidad teórica que se desea estimar, propondremos dos tipos de selectores de las ventanas, unos de tipo plug-in y otros bootstrap. Demostraremos que el selector plug-in es consistente, y finalmente estudiaremos el comportamiento de todos ellos en un estudio de simulación.

4.1. Representación de las ventanas MISE

La medida de distancia global entre la función de densidad verdadera y su estimador probablemente más estudiada y usada en la práctica es el error cuadrático medio integrado (*MISE*):

$$\begin{aligned} MISE(f_n^P(\cdot)) &= \mathbb{E} \left[\int (f_n^P(v) - f(v))^2 \omega(v) dv \right] \\ &= \int \mathbb{E} (f_n^P(v) - f(v))^2 \omega(v) dv + \int Var [f_n^P(v)] \omega(v) dv, \end{aligned} \quad (4.1)$$

donde $\omega(\cdot)$ es una función de peso no negativa. Esta función es muy útil en el contexto de censura, puesto que es probable que existan zonas del soporte de la variable de interés en las que sea particularmente difícil la estimación.

Los parámetros de suavización s y b que minimizan la expresión (4.1) se denominan ventanas *MISE*, y se denotarán por s_{MISE} y b_{MISE} . Sin embargo, al desconocer la expresión exacta de la función $MISE(f_n^P(\cdot))$ y disponer únicamente de su parte dominante asintóticamente, $AMISE(\cdot, \cdot)$, más términos despreciables, obtendremos la expresión de las ventanas que minimizan el $AMISE(\cdot, \cdot)$, y que se denotarán por s_{AMISE} y b_{AMISE} . No obstante, como se verá más adelante, la expresión de estas ventanas *AMISE* será precisamente el término dominante de las ventanas *MISE*.

En la siguiente propiedad se da, a partir del resultado del teorema 3.2.4 del capítulo anterior, la expresión de la función $AMISE(\cdot, \cdot)$, distinguiendo de nuevo tres casos, según sea la relación entre las ventanas s y b . Dicha expresión depende de las siguientes funciones, algunas de ellas ya definidas previamente en capítulos anteriores:

$$A(t) = f(t) \frac{1 - F(t)}{1 - H(t)},$$

$$M(t) = d_K f''(t), \quad (4.2)$$

$$N(t) = d_K [(1 - F(t)) \alpha(t)]', \quad (4.3)$$

$$\alpha(t) = \int_0^t \frac{\frac{1}{2} p''(v) h(v) + p'(v) h'(v)}{1 - H(v)} dv, \quad (4.4)$$

$$A_K(L) = \iiint K(u) K(v) K(w) K(u + L(v - w)) dudvdw,$$

$$A(t, L) = A(t) p(t) c_K + A(t) (1 - p(t)) A_K(L), \quad (4.5)$$

y de las constantes que dependen de la función núcleo

$$c_K = \int K^2(v) dv, \quad d_K = \int v^2 K(v) dv \quad \text{y} \quad e_K = \int v K(v) \mathbb{K}(v) dv.$$

Propiedad 4.1.1

Bajo las condiciones del teorema 3.2.4, se tiene lo siguiente

(a) Si $b/s \rightarrow 0$, entonces

$$MISE(f_n^P(\cdot)) = AMISE(s, b) + O(s^6) + o((ns)^{-1}),$$

donde

$$\begin{aligned} AMISE(s, b) &= \frac{1}{4}s^4 \int M^2(v) \omega(v) dv + b^4 \int N^2(v) \omega(v) dv \\ &+ s^2 b^2 \int M(v) N(v) \omega(v) dv + \frac{1}{ns} c_K \int A(v) \omega(v) dv. \end{aligned} \quad (4.6)$$

(b) Si $b/s \rightarrow L > 0$, entonces

$$MISE(f_n^P(\cdot)) = AMISE(s, b) + O(s^6) + o((ns)^{-1}),$$

donde

$$\begin{aligned} AMISE(s, b) &= \frac{1}{4}s^4 \int M^2(v) \omega(v) dv + b^4 \int N^2(v) \omega(v) dv \\ &+ s^2 b^2 \int M(v) N(v) \omega(v) dv + \frac{1}{ns} \int A(v, L) \omega(v) dv. \end{aligned} \quad (4.7)$$

(c) Si $b/s \rightarrow \infty$, entonces

$$MISE(f_n^P(\cdot)) = AMISE(s, b) + O(b^6) + o((nb)^{-1}),$$

donde

$$\begin{aligned} AMISE(s, b) &= \frac{1}{4}s^4 \int M^2(v) \omega(v) dv + b^4 \int N^2(v) \omega(v) dv \\ &+ s^2 b^2 \int M(v) N(v) \omega(v) dv \\ &+ \frac{1}{ns} c_K \int A(v) p(v) \omega(v) dv + \frac{1}{nb} c_K \int A(v) (1 - p(v)) \omega(v) dv. \end{aligned}$$

Demostración.

Es inmediato teniendo en cuenta la expresión de $MISE(f_n^P(\cdot))$ dada en el teorema 3.2.4. del capítulo anterior. ■

Las ventanas $AMISE$ son las que minimizan la función $AMISE(\cdot, \cdot)$. Su expresión, que depende de la relación entre las ventanas, viene dada en la propiedad 4.1.2.

Observación 4.1.1

Cuando la relación entre las ventanas es $b/s \rightarrow 0$, el orden del término despreciable en (4.6) no admite un balance entre el sesgo y la varianza que permita calcular la ventana de presuavizado b de manera que minimice $AMISE(\cdot, \cdot)$, así que en este caso obtendremos las ventanas s y b de forma que minimicen la parte asintóticamente dominante de $MISE\left(\overline{f_n^P}(\cdot)\right)$ hasta términos de orden superior, y las denominaremos por \overline{s}_{AMISE} y \overline{b}_{AMISE} respectivamente:

$$MISE\left(\overline{f_n^P}(\cdot)\right) = \overline{MISE}(s, b) = \overline{AMISE}(s, b) + o(s^4) + o(s^2b^2) + o(n^{-1}b),$$

donde

$$\begin{aligned} \overline{AMISE}(s, b) &= \frac{1}{4}s^4 \int M^2(v) \omega(v) dv + b^4 \int N^2(v) \omega(v) dv \\ &+ s^2b^2 \int M(v) N(v) \omega(v) dv \\ &+ \frac{1}{ns} c_K \int A(v) \omega(v) dv + \frac{1}{n} \int B(v) \omega(v) dv - 2\frac{b}{n} \int C(v) \omega(v) dv. \end{aligned} \quad (4.8)$$

con $B(\cdot)$ y $C(\cdot)$ dados en (3.32) y (3.33) en el capítulo anterior.

Observación 4.1.2

Cuando la relación entre las ventanas es $b/s \rightarrow \infty$, mediante la minimización en la expresión de $AMISE(\cdot, \cdot)$ con respecto a s y b se obtiene que ambas ventanas deben ser de orden $n^{-1/5}$, lo que contradice la hipótesis de que el cociente entre las ventanas es $b/s \rightarrow \infty$. Por lo tanto, las ventanas óptimas s_{AMISE} y b_{AMISE} no pueden guardar esta relación asintótica.

Propiedad 4.1.2

Bajo las condiciones del teorema 3.2.4, las ventanas $AMISE$ vienen dadas por las siguientes expresiones

(a) Si $\overline{b}_{AMISE}/\overline{s}_{AMISE} \rightarrow 0$, entonces

$$\overline{s}_{AMISE} = \left(\frac{c_K \int A(v) \omega(v) dv}{\int M^2(v) \omega(v) dv} \right)^{1/5} n^{-1/5}, \quad (4.9)$$

$$\overline{b}_{AMISE} = \frac{\int C(v) \omega(v) dv \times \left(\int M^2(v) \omega(v) dv \right)^{2/5}}{\left(c_K \int A(v) \omega(v) dv \right)^{2/5} \times \int M(v) N(v) \omega(v) dv} n^{-3/5}. \quad (4.10)$$

Además,

$$\begin{aligned} & \overline{MISE}(\bar{s}_{AMISE}, \bar{b}_{AMISE}) \\ &= \frac{5}{4} \left(c_K \int A(v) \omega(v) dv \right)^{4/5} \left(\int M^2(v) \omega(v) dv \right)^{1/5} n^{-4/5} + \frac{1}{n} \int B(v) \omega(v) dv \\ & \quad - \frac{\left(\int C(v) \omega(v) dv \right)^2 \left(\int M^2(v) \omega(v) dv \right)^{2/5}}{\left(c_K \int A(v) \omega(v) dv \right)^{2/5} \int M(v) N(v) \omega(v) dv} n^{-8/5} + O(n^{-12/5}). \end{aligned} \quad (4.11)$$

(b) Si $b_{AMISE}/s_{AMISE} \rightarrow L_0 > 0$, entonces

$$s_{AMISE} = \left(\frac{\int A(v, L_0) \omega(v) dv}{4 \int \left(\frac{1}{2} M(v) + L_0^2 N(v) \right)^2 \omega(v) dv} \right)^{1/5} n^{-1/5}, \quad (4.12)$$

$$b_{AMISE} = L_0 \times s_{AMISE}. \quad (4.13)$$

donde $L_0 = \arg \min_{L>0} \psi(L)$, siendo

$$\psi(L) = \left(\int A(v, L) \omega(v) dv \right)^{4/5} \left[\int \left(\frac{1}{2} M(v) + L^2 N(v) \right)^2 \omega(v) dv \right]^{1/5} \quad (4.14)$$

y

$$MISE(s_{AMISE}, b_{AMISE}) = \frac{5}{4^{4/5}} \psi(L_0) n^{-4/5} + o(n^{-4/5}). \quad (4.15)$$

Demostración.

(a) Si $b/s \rightarrow 0$, entonces derivamos (4.8) respecto a s y a b , resolvemos el sistema y obtenemos las expresiones de las ventanas *AMISE* dadas en (4.9) y (4.10), lo que implica que el error mínimo cometido con estas ventanas es (4.11).

(b) Si las ventanas s y b son del mismo orden, entonces tomando asintóticamente $b = L \times s$, la expresión de *AMISE*(\cdot, \cdot) dada en (4.7) se reduce a la siguiente:

$$\begin{aligned} AMISE(s, Ls) &= s^4 \int \left(\frac{1}{2} M(v) + L^2 N(v) \right)^2 \omega(v) dv \\ & \quad + \frac{1}{ns} \int A(v, L) \omega(v) dv. \end{aligned} \quad (4.16)$$

El balance clásico entre los términos con s^4 y $(ns)^{-1}$ en (4.16) proporciona las ventanas óptimas (4.12) y (4.13), que dependen de $L > 0$. Sustituimos en (4.16) el parámetro s por una expresión análoga a (4.12) pero con L en lugar de L_0 , de forma que el mínimo de *AMISE*(\cdot, \cdot) como función de L , se alcanza en $L_0 = \arg \min_{L>0} \psi(L)$.

■

Observación 4.1.3

Estos resultados extienden a los del teorema 3.2 de Lo, Mack y Wang (1989) sin presuavizar, y al teorema 1 de Cao (1993), aplicable al caso sin censura.

Observación 4.1.4

En el caso particular en el que $L = 0$, es decir, cuando $b = 0$ de modo que no se presuaviza en absoluto, el término dominante en (4.15) es

$$\frac{5}{4} \left(c_K \int A(v) \omega(v) dv \right)^{4/5} \left(\int M^2(v) \omega(v) dv \right)^{1/5},$$

que coincide con el término dominante de la función MISE para el estimador tipo núcleo con pesos de tipo Kaplan-Meier $f_n^{KM}(\cdot)$.

Observación 4.1.5

Si la función núcleo $K(\cdot)$ es log-cóncava, es decir, su logaritmo es una función cóncava, entonces

$$\int K(u) K(u+t) du \stackrel{(a)}{=} \int K\left(u - \frac{t}{2}\right) K\left(u + \frac{t}{2}\right) du \stackrel{(b)}{\leq} \int K^2(u) du = c_K.$$

En la primera igualdad (paso (a)) se realizó un cambio de variable, mientras que en el paso (b) se ha aplicado la log-concavidad de $K(\cdot)$, ya que entonces

$$\begin{aligned} & \frac{1}{2} \log K\left(u - \frac{t}{2}\right) + \frac{1}{2} \log K\left(u + \frac{t}{2}\right) \\ & \leq \log K\left(\frac{1}{2}\left(u - \frac{t}{2}\right) + \frac{1}{2}\left(u + \frac{t}{2}\right)\right) = \log K(u), \end{aligned}$$

con lo cual

$$K\left(u - \frac{t}{2}\right) K\left(u + \frac{t}{2}\right) \leq K^2(u) \text{ para cualquier } t.$$

Podemos concluir así lo siguiente,

$$A_K(L) \leq A_K(0) = c_K, \text{ y por tanto } A(t, L) \leq A(t, 0) = A(t) c_K \quad \forall L > 0,$$

lo que implica

$$\left(\int A(v, L) \omega(v) dv \right)^{4/5} \leq \left(c_K \int A(v) \omega(v) dv \right)^{4/5} \quad \forall L > 0. \quad (4.17)$$

Las principales funciones núcleo usadas en la estimación no paramétrica, entre ellas las dadas en la tabla 1.1 del primer capítulo, son log-cóncavas.

Observación 4.1.6

Cuando las ventanas s y b son del mismo orden, es decir, $b/s \rightarrow L_0 > 0$, la eficiencia del estimador presuavizado, $f_n^P(\cdot)$, es de primer orden con respecto al estimador de Kaplan-Meier $f_n^{KM}(\cdot)$, dado que

$$\begin{aligned} \frac{5}{4^{4/5}} \left(\int A(v, L_0) \omega(v) dv \right)^{4/5} \left(\int \left(\frac{1}{2} M(v) + L_0^2 N(v) \right)^2 \omega(v) dv \right)^{1/5} &= \frac{5}{4^{4/5}} \psi(L_0) \\ &< \frac{5}{4^{4/5}} \psi(0) = \frac{5}{4} \left(c_K \int A(v) \omega(v) dv \right)^{4/5} \left(\frac{1}{4} \int M^2(v) \omega(v) dv \right)^{1/5}, \end{aligned}$$

de modo que el término dominante de $MISE(f_n^P(\cdot))$ es menor que el término dominante de $MISE(f_n^{KM}(\cdot))$ cuando $L_0 \neq 0$. Si $L_0 = 0$ estamos en el caso (a), en el que se puede comprobar fácilmente que la eficiencia del estimador presuavizado $f_n^P(\cdot)$ es de tercer orden.

Observación 4.1.7

Si la integral $\int M(v) N(v) \omega(v) dv$ es negativa, entonces la función $\int \left(\frac{1}{2} M(v) + L^2 N(v) \right)^2 \omega(v) dv$ tiene un mínimo en

$$\tilde{L} = \left(-\frac{1}{2} \frac{\int M(v) N(v) \omega(v) dv}{\int N^2(v) \omega(v) dv} \right)^{1/2} > 0,$$

con lo cual, para un núcleo log-cóncavo, usando (4.17) se tiene $\psi(\tilde{L}) < \psi(0)$.

Por tanto, si $\int M(v) N(v) \omega(v) dv < 0$ entonces el valor óptimo en (4.14) se alcanza en algún $L_0 > 0$ y la eficiencia del estimador presuavizado $f_n^P(\cdot)$ es de primer orden con respecto al estimador de Kaplan-Meier $f_n^{KM}(\cdot)$. Obviamente, también podría haber situaciones en las que $\int M(v) N(v) \omega(v) dv \geq 0$ y $L_0 > 0$.

El siguiente teorema proporciona la parte dominante de las ventanas que minimizan $MISE(f_n^P(\cdot))$, que coincide con la expresión de las ventanas que minimizan $AMISE(\cdot, \cdot)$, en el caso en el que ambas ventanas sean del mismo orden. En el caso en que $b_{MISE}/s_{MISE} \rightarrow 0$, la demostración es más laboriosa, puesto que habría que considerar el $MISE(\overline{f_n^P}(\cdot))$ en lugar de $MISE(f_n^P(\cdot))$.

Teorema 4.1.1 *Bajo las condiciones del teorema 3.2.4, y si las ventanas óptimas son del mismo orden, entonces*

$$\frac{s_{MISE} - s_{AMISE}}{s_{MISE}} \rightarrow 0 \quad \text{y} \quad \frac{b_{MISE} - b_{AMISE}}{b_{MISE}} \rightarrow 0$$

donde las ventanas s_{AMISE} y b_{AMISE} vienen dadas por (4.12) y (4.13) respectivamente, siendo L_0 el único mínimo de la función $\psi(\cdot)$.

Demostración.

En el caso en el que las ventanas óptimas son del mismo orden (supondremos $b = s \times L$), mediante una reparametrización de (4.7) se tiene que

$$MISE_{sL}(s, L) = AMISE_{sL}(s, L) + O(s^6) + o((ns)^{-1}), \quad (4.18)$$

con

$$AMISE_{sL}(s, L) = s^4 c_1(L) + \frac{1}{ns} c_2(L), \quad (4.19)$$

siendo

$$c_1(L) = \int \left(\frac{1}{2} M(v) + L^2 N(v) \right)^2 \omega(v) dv \quad (4.20)$$

$$c_2(L) = \int A(v, L) \omega(v) dv \quad (4.21)$$

y donde $M(\cdot)$, $N(\cdot)$ y $A(\cdot, \cdot)$ han sido definidas en (4.2), (4.3) y (4.5) respectivamente.

El balance clásico entre los términos con s^4 y $n^{-1}s^{-1}$ en (4.19) proporciona las siguientes ventanas $AMISE$ óptimas (dependientes del parámetro L):

$$s_{AMISE}(L) = \left(\frac{c_2(L)}{4c_1(L)} \right)^{1/5} n^{-1/5} = c_0(L) n^{-1/5} \quad \text{y} \quad b_{AMISE}(L) = L \times s_{AMISE}(L)$$

donde

$$c_0(L) = \left(\frac{c_2(L)}{4c_1(L)} \right)^{1/5}. \quad (4.22)$$

Sustituyendo estas ventanas en (4.19) obtenemos

$$AMISE_{sL}(s_{AMISE}(L), L) = \frac{5}{4^{4/5}} \left(c_1(L) c_2(L)^4 \right)^{1/5} n^{-4/5} = \frac{5}{4^{4/5}} \psi(L) n^{-4/5},$$

donde $\psi(\cdot)$ ha sido definido en (4.14). Si tomamos

$$L_0 = \arg \min_{L>0} AMISE_{sL}(s_{AMISE}(L), L) = \arg \min_{L>0} \psi(L), \quad (4.23)$$

tenemos que las ventanas $AMISE$ óptimas son las siguientes:

$$s_{AMISE} = c_0(L_0) n^{-1/5} \quad \text{y} \quad b_{AMISE} = L_0 \times s_{AMISE} = L_0 \times c_0(L_0) n^{-1/5}.$$

Por otro lado, por definición, las ventanas $MISE$ (en función del parámetro L) son

$$s_{MISE}(L) = \arg \min_{s>0} MISE_{sL}(s, L) \quad \text{y} \quad b_{MISE}(L) = L \times s_{MISE}(L).$$

Sustituyendo estos valores en la expresión de $MISE_{sL}(\cdot, \cdot)$ dada en (4.18), tenemos que las ventanas MISE son $s_{MISE} = s_{MISE}(L_{MISE,n})$ y $b_{MISE} = b_{MISE}(L_{MISE,n})$ donde

$$L_{MISE,n} = \arg \min_{L>0} MISE_{sL}(s_{MISE}(L), L)$$

es un valor que varía con n . Vamos a demostrar que las ventanas que minimizan la función $MISE_{sL}(\cdot, \cdot)$ dada en (4.18) (s_{MISE} y b_{MISE}) y las que minimizan su parte dominante $AMISE_{sL}(\cdot, \cdot)$ dada en (4.19) (s_{AMISE} y b_{AMISE}) guardan la siguiente relación:

$$\begin{aligned} s_{MISE} &= c_0(L_0)n^{-1/5} + o(n^{-1/5}) = s_{AMISE} + o(n^{-1/5}), \\ b_{MISE} &= L_0 \times c_0(L_0)n^{-1/5} + o(n^{-1/5}) = b_{AMISE} + o(n^{-1/5}). \end{aligned}$$

En primer lugar demostraremos que la parte dominante de $L_{MISE,n}$ es precisamente el valor L_0 dado en (4.23), es decir, $\lim_{n \rightarrow \infty} L_{MISE,n} = L_0$. Para ello definimos la siguiente función

$$MISE_L(L) = MISE_{sL}(s_{MISE}(L), L) = \frac{5}{4^{4/5}}\psi(L)n^{-4/5} + o(n^{-4/5}), \quad (4.24)$$

que por construcción alcanza su mínimo en $L_{MISE,n}$. Entonces

$$\min_{L>0} MISE_L(L) = MISE_L(L_{MISE,n}).$$

Por otro lado, tenemos que

$$\lim_{n \rightarrow \infty} n^{4/5} MISE_L(L) = \frac{5}{4^{4/5}}\psi(L) \quad (4.25)$$

a su vez alcanza su único mínimo en $L_0 < \infty$.

Supongamos que existe $L_{MISE} = \lim_{n \rightarrow \infty} L_{MISE,n}$. Entonces, de existir el límite, éste tiene que ser igual a L_0 . Para probarlo, en el paso (a) aplicamos que la función $\psi(\cdot)$ es continua, el paso (b) se debe a la expresión de $MISE_L(\cdot)$ dada en (4.24), y el paso (c) es directo puesto que, por definición, $L_{MISE,n}$ es el argumento que minimiza $MISE_L(\cdot)$:

$$\begin{aligned} \frac{5}{4^{4/5}}\psi(L_{MISE}) &\stackrel{(a)}{=} \lim_{n \rightarrow \infty} \frac{5}{4^{4/5}}\psi(L_{MISE,n}) \stackrel{(b)}{=} \lim_{n \rightarrow \infty} n^{4/5} MISE_L(L_{MISE,n}) \\ &\stackrel{(c)}{\leq} \lim_{n \rightarrow \infty} n^{4/5} MISE_L(L_0) = \frac{5}{4^{4/5}}\psi(L_0). \end{aligned} \quad (4.26)$$

Por tanto, $\psi(L_{MISE}) \leq \psi(L_0) < \infty$. Puesto que L_0 es el único argumento que minimiza $\psi(\cdot)$, la anterior desigualdad permite concluir que

$$L_{MISE} = \lim_{n \rightarrow \infty} L_{MISE,n} = L_0. \quad (4.27)$$

Veamos ahora que existe $\lim_{n \rightarrow \infty} L_{MISE,n}$. Esta sucesión está acotada inferiormente por el cero, de modo que existe $L_1 = \liminf_{n \rightarrow \infty} L_{MISE,n}$. Definimos entonces $L_2 = \limsup_{n \rightarrow \infty} L_{MISE,n}$. Comprobaremos que $L_2 < \infty$ y, posteriormente, que $L_1 = L_2$ lo que implica finalmente (4.27).

Para demostrar que $L_2 < \infty$ en primer lugar (paso (a)) consideramos la forma de la función $c_1(\cdot)$ dada en (4.20), en el paso (b) aplicamos la definición de $\psi(\cdot)$ dada en (4.14), a continuación tenemos en cuenta la continuidad de $\psi(\cdot)$ y la definición de L_2 (paso (c)), la igualdad (4.24) en el paso (d), el hecho de que, por definición, $L_{MISE,n}$ es el argumento que minimiza la función $MISE_L(\cdot)$ (paso (e)) y, por último, en el paso (f), la relación (4.25):

$$\begin{aligned} \left(L_2^2 \times \int M(v) N(v) \omega(v) dv \times c_2^4(L_2) \right)^{1/5} &\stackrel{(a)}{\leq} (c_1(L_2) c_2^4(L_2))^{1/5} \stackrel{(b)}{=} \psi(L_2) \\ &\stackrel{(c)}{\leq} \limsup_{n \rightarrow \infty} \frac{5}{4^{4/5}} \psi(L_{MISE,n}) \stackrel{(d)}{\leq} \limsup_{n \rightarrow \infty} n^{4/5} MISE_L(L_{MISE,n}) \\ &\stackrel{(e)}{\leq} \limsup_{n \rightarrow \infty} n^{4/5} MISE_L(L_0) \stackrel{(f)}{=} \frac{5}{4^{4/5}} \psi(L_0) < \infty. \end{aligned}$$

Para ver ahora que $L_1 = L_2 = L_0$, usamos el mismo razonamiento que en (4.26) con L_1 y L_2 respectivamente en lugar de L_{MISE} para concluir

$$\psi(L_1) \leq \psi(L_0) \quad \text{y} \quad \psi(L_2) \leq \psi(L_0).$$

Pero dado que, por definición, L_0 es el único mínimo de la función $\psi(\cdot)$, entonces $L_1 = L_2 = L_0$, demostrando así (4.27).

A continuación, demostraremos que el término dominante de s_{MISE} es precisamente de la forma $c_0(L_0) n^{-1/5}$ donde $c_0(\cdot)$ viene dado por (4.22), es decir,

$$\lim_{n \rightarrow \infty} n^{1/5} s_{MISE} = c_0(L_0).$$

La ventana s_{MISE} es, por definición:

$$s_{MISE} = s_{MISE}(L_{MISE,n}) = \arg \min_{s>0} MISE_{sL}(s, L_{MISE,n}).$$

Evaluamos la función $MISE_{sL}(\cdot, \cdot)$ en una ventana s elegida de la forma $kn^{-1/5}$ y en $L_{MISE,n}$:

$$MISE(kn^{-1/5}, L_{MISE,n}) = (c_1(L_{MISE,n}) k^4 + c_2(L_{MISE,n}) k^{-1}) n^{-4/5} + o(n^{-4/5}).$$

Por la continuidad de las funciones $c_1(\cdot)$ y $c_2(\cdot)$, y puesto que $\lim_{n \rightarrow \infty} L_{MISE,n} = L_0$, entonces

$$\lim_{n \rightarrow \infty} n^{4/5} MISE_{sL}(kn^{-1/5}, L_{MISE,n}) = c_1(L_0)k^4 + c_2(L_0)k^{-1}. \quad (4.28)$$

Pero la función $c_1(L_0)x^4 + c_2(L_0)x^{-1}$ alcanza su mínimo en el punto $x = (c_2(L_0)/4c_1(L_0))^{1/5}$ de modo que la mejor elección para la constante k es $k = c_0(L_0)$.

Se ha probado que, de existir el límite $\lim_{n \rightarrow \infty} n^{1/5} s_{MISE}$, éste tiene que ser $c_0(L_0)$. Veamos ahora que ese límite existe. Para ello, en primer lugar veremos que

$$0 < U_1 = \liminf_{n \rightarrow \infty} n^{1/5} s_{MISE} \leq \limsup_{n \rightarrow \infty} n^{1/5} s_{MISE} = U_2 < \infty,$$

y en segundo lugar que $U_1 = U_2 = c_0(L_0)$, lo que implicaría que

$$\lim_{n \rightarrow \infty} n^{1/5} s_{MISE} = c_0(L_0).$$

Así, en el paso (a) aplicamos la definición de U_2 , que la función $c_1(\cdot)$ es continua en L y que $\lim_{n \rightarrow \infty} L_{MISE,n} = L_0$, en el paso (b) tenemos en cuenta que la parte dominante de $MISE_{sL}(s, L_{MISE,n})$ es igual a $c_1(L_{MISE,n})s^4$ más términos positivos que provienen de la integral de la varianza, el paso (c) es directo sin más que recordar que s_{MISE} es por definición la ventana que minimiza $MISE_{sL}(\cdot, L_{MISE,n})$ y en el paso (d) consideramos la igualdad (4.28):

$$\begin{aligned} c_1(L_0)U_2^4 &\stackrel{(a)}{=} \limsup_{n \rightarrow \infty} n^{4/5} c_1(L_{MISE,n}) s_{MISE}^4 \\ &\stackrel{(b)}{\leq} \limsup_{n \rightarrow \infty} n^{4/5} MISE_{sL}(s_{MISE}, L_{MISE,n}) \\ &\stackrel{(c)}{\leq} \limsup_{n \rightarrow \infty} n^{4/5} MISE_{sL}(c_0(L_0)n^{-1/5}, L_{MISE,n}) \\ &\stackrel{(d)}{=} c_1(L_0)c_0^4(L_0) + c_2(L_0)c_0^{-1}(L_0) = \frac{5}{4^{4/5}} (c_1(L_0)c_2^4(L_0))^{1/5}, \end{aligned}$$

lo que implica que $U_2 < \infty$.

A continuación, por un argumento similar, se obtiene que $U_1 > 0$. En el paso (a) usamos la definición de U_1 , en el paso (b) consideramos la igualdad $\liminf_{n \rightarrow \infty} n^{1/5} s_{MISE} = (\limsup_{n \rightarrow \infty} n^{-1/5} s_{MISE}^{-1})^{-1}$, en el paso (c) tenemos en cuenta que la parte dominante de $MISE_{sL}(s, L_{MISE,n})$ es igual a $c_2(L_{MISE,n})n^{-1}s^{-1}$ más términos positivos que provienen de la integral del sesgo al cuadrado, el paso (d) es directo puesto que s_{MISE} y $L_{MISE,n}$ son, por definición, los valores que minimizan $MISE_{sL}(\cdot, \cdot)$ y en el paso (e) usamos la igualdad (4.28):

$$\begin{aligned}
c_2(L_0) U_1^{-1} &\stackrel{(a)}{=} c_2(L_0) \left(\liminf_{n \rightarrow \infty} n^{1/5} s_{MISE} \right)^{-1} \stackrel{(b)}{=} c_2(L_0) \limsup_{n \rightarrow \infty} n^{-1/5} s_{MISE}^{-1} \\
&= \limsup_{n \rightarrow \infty} n^{4/5} c_2(L_{MISE,n}) n^{-1} s_{MISE}^{-1} \\
&\stackrel{(c)}{\leq} \limsup_{n \rightarrow \infty} n^{4/5} MISE_{sL}(s_{MISE}, L_{MISE,n}) \\
&\stackrel{(d)}{\leq} \limsup_{n \rightarrow \infty} n^{4/5} MISE_{sL}(c_0(L_0) n^{-1/5}, L_{MISE,n}) \\
&\stackrel{(e)}{=} \frac{5}{4^{4/5}} (c_1(L_0) c_2^4(L_0))^{1/5}.
\end{aligned}$$

Una vez probado que la sucesión $n^{1/5} s_{MISE}$ tiene límite inferior y superior, veamos que ambos son iguales a $c_0(L_0)$. Para ello, consideramos ahora dos subsucesiones $\{s_{MISE n_{1(k)}}\}$ y $\{s_{MISE n_{2(k)}}\}$ verificando

$$\lim_{k \rightarrow \infty} n_{1(k)}^{1/5} s_{MISE} = U_1 \quad \text{y} \quad \lim_{k \rightarrow \infty} n_{2(k)}^{1/5} s_{MISE} = U_2.$$

Entonces

$$\begin{aligned}
c_1(L_0) U_2^4 + c_2(L_0) U_2^{-1} &= \lim_{k \rightarrow \infty} n_{2(k)}^{4/5} MISE_{sL}(s_{MISE n_{2(k)}}, L_{MISE,n}) \\
&\leq \lim_{k \rightarrow \infty} n_{2(k)}^{4/5} MISE_{sL}(c_0(L_0) n_{2(k)}^{-1/5}, L_{MISE,n}) \\
&= c_1(L_0) c_0^4(L_0) + c_2(L_0) c_0^{-1}(L_0).
\end{aligned}$$

Del mismo modo

$$c_1(L_0) U_1^4 + c_2(L_0) U_1^{-1} \leq c_1(L_0) c_0^4(L_0) + c_2(L_0) c_0^{-1}(L_0).$$

Pero como $c_0(L_0)$ es el único real positivo donde se alcanza el mínimo de la función $c_1(L_0) x^4 + c_2(L_0) x^{-1}$ entonces $c_0(L_0) = U_1 = U_2$, es decir

$$\lim_{n \rightarrow \infty} n^{1/5} s_{MISE} = c_0(L_0) = \left(\frac{c_2(L_0)}{4c_1(L_0)} \right)^{1/5}. \quad (4.29)$$

Por último, de (4.27) y (4.29) se obtiene que

$$\lim_{n \rightarrow \infty} n^{1/5} b_{MISE} = \lim_{n \rightarrow \infty} L_{MISE,n} \times n^{1/5} b_{MISE} = L_0 \times c_0(L_0).$$

Queda así demostrado el teorema. ■

4.2. Comparación de las ventanas MISE y AMISE por simulación

Se ha realizado un estudio de simulación para ilustrar el comportamiento en la práctica de las ventanas *MISE* y *AMISE*, y comprobar cómo las ventanas *AMISE* se aproximan a las correspondientes *MISE* cuando aumenta el tamaño muestral. De nuevo, hemos considerado los cuatro modelos simulados en los capítulos anteriores. En los modelos 1, 2 y 4 la variable observable sigue la distribución $Z \stackrel{d}{=} U[0, 1]$, mientras que en el modelo 3 se tiene que $Z \stackrel{d}{=} W(1.75, 5)$.

En la tabla 4.1 se muestran las funciones de densidad teóricas a estimar $f(\cdot)$ y las probabilidades condicionales de no censura $p(\cdot)$ para cada uno de los modelos. Además, en la figura 4.1 se representan la función de densidad $f(\cdot)$ de la variable de interés (línea con puntos), la función de densidad de la variable observable $h(\cdot)$ (línea continua gruesa) y la función de probabilidad condicional de no censura $p(\cdot)$ (línea continua fina) para cada uno de los modelos. Para todas ellas, la función peso $\omega(\cdot)$ es la función indicadora del intervalo cuyos extremos vienen dados por las líneas verticales. Estos intervalos son los mismos que los tomados en el estudio de simulación de los capítulos anteriores.

Como se vio al comienzo del presente capítulo, la expresión de las ventanas *AMISE* depende de la relación asintótica entre las ventanas s y b . Por eso, en primer lugar estudiaremos para cada modelo si, en el óptimo, el cociente b/s tiende, cuando n tiende a infinito, a un número positivo $L > 0$ ó a cero. Esta relación asintótica entre s y b también indicará, para cada modelo, si la eficiencia del estimador presuavizado $f_n^P(\cdot)$ con respecto a $f_n^{KM}(\cdot)$ es de primer orden ($L > 0$) o de tercer orden ($L = 0$).

Tabla 4.1. Funciones de densidad $f(\cdot)$ y probabilidad condicional de no censura $p(\cdot)$ en los modelos 1-4.

Función de densidad	Función $p(t)$
$f(t) = (1-t)^{3/8} e^{-1/4t^2} \left(\frac{3}{8}(1-t)^{-1} + \frac{1}{2}t \right)$	$p(t) = \frac{1}{2} \left[1 - \left(t - \frac{1}{2} \right)^2 \right]$
$f(t) = 3te^{-3/2t^2}$	$p(t) = \left[\frac{3}{4} - 3 \left(t - \frac{1}{2} \right)^2 \right] \mathbf{1}_{\{0 \leq t \leq 1\}}$
$f(t) = \frac{5}{2}t^4 e^{-t^5} + 80t^4 e^{-32t^5}$	$p(t) = \frac{\frac{5}{2}e^{-t^5} + 80e^{-32t^5}}{41.0327(e^{-t^5} + e^{-32t^5})}$
$f(t) = (20t^3 - 12t^2 + 1) e^{-5t^4 + 4t^3 - t}$	$p(t) = (20t^3 - 12t^2 + 1)(1-t)$

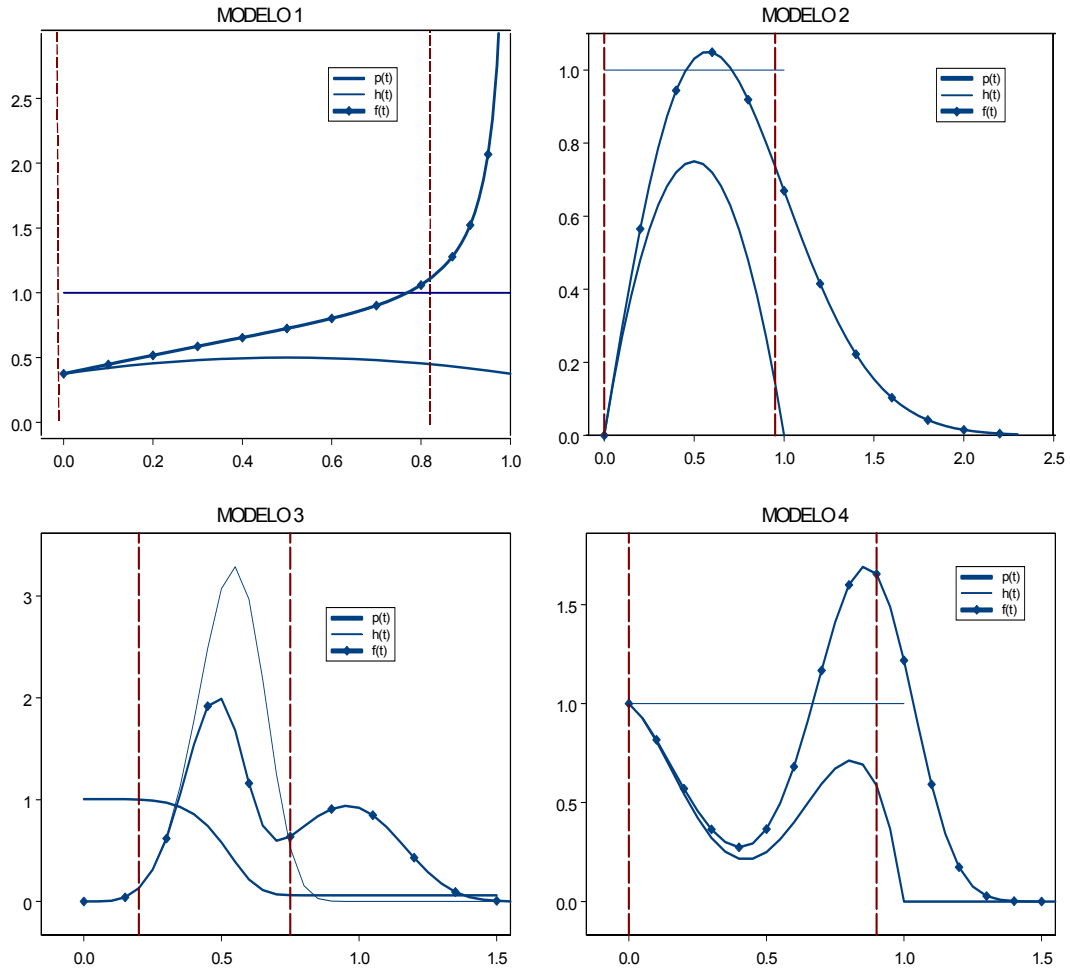


Figura 4.1. Funciones $p(\cdot)$, $h(\cdot)$ y $f(\cdot)$ para los modelos 1-4. La función peso $\omega(\cdot)$ es la función indicadora del intervalo marcado por las líneas verticales.

El valor de la constante L_0 es aquel que minimiza la función

$$\psi^5(L) = \left(\int A(v, L) \omega(v) dv \right)^4 \int \left(\frac{1}{2} M(v) + L^2 N(v) \right)^2 \omega(v) dv,$$

siendo $A(\cdot, \cdot)$, $M(\cdot)$ y $N(\cdot)$ las funciones dadas en (4.5), (4.2) y (4.3) respectivamente. Se ha calculado la expresión exacta de esta función para cada modelo (ver figura 4.2) y se han obtenido los valores $L_0 \geq 0$ que minimizan dichas funciones. Este valor es $L_0 = 3.56$ para el modelo 1 y $L_0 = 0$ para los modelos 2-4. La eficiencia del estimador presuavizado es de primer orden únicamente en el modelo 1.

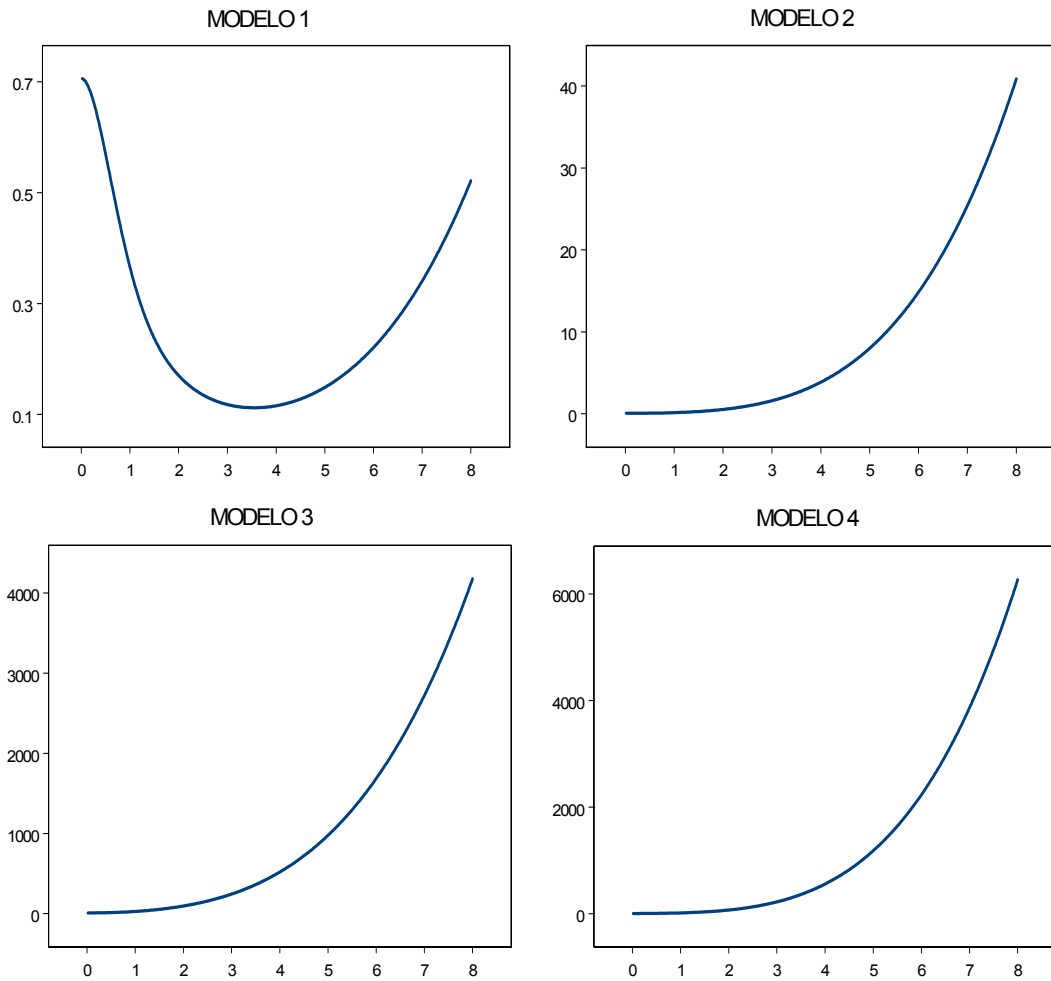


Figura 4.2. Función $\psi^5(\cdot)$ para los modelos 1-4 para valores de $L \in [0, 8]$.

Para los modelos 2-4, en los que la relación asintótica entre las ventanas *AMISE* es $b/s \rightarrow L = 0$, la expresión de las ventanas óptimas \bar{s}_{AMISE} y \bar{b}_{AMISE} viene dada por (4.9) y (4.10) respectivamente. Aunque éstas son las expresiones de las ventanas asintóticamente óptimas, para tamaños muestrales finitos, el valor de la ventana \bar{b}_{AMISE} puede llegar a ser extraordinariamente pequeño. Por esa razón, en la práctica se ha considerado que las ventanas óptimas verifican que $b/s = L_n \rightarrow L = 0$, y se han tomado como ventana *AMISE* de suavizado \bar{s}_{AMISE} dada por (4.9) y

$$\bar{b}_{AMISE} = L_n \times \bar{s}_{AMISE}.$$

Para cada tamaño muestral n , el valor de L_n se ha calculado como aquel que minimiza la función *AMISE*(\cdot, \cdot) dada por (4.8).

Para cada modelo, y para los tamaños muestrales $n = 50, 100, 200$ y 1000 se ha calculado el valor de s_{AMISE}^{KM} , la ventana $AMISE$ para el estimador de Kaplan-Meier suavizado $f_n^{KM}(\cdot)$. Con dicha ventana, y mediante una aproximación de Montecarlo con $m = 500$ muestras, se ha obtenido el error cuadrático medio integrado cometido, es decir, $MISE_{s_{AMISE}^{KM}}(f_n^{KM}(\cdot))$.

Definimos ahora la eficiencia relativa $AMISE$ del estimador de Kaplan-Meier $f_n^{KM}(\cdot)$ con respecto al estimador presuavizado $f_n^P(\cdot)$ como el cociente entre $MISE(f_n^P(\cdot))$ y $MISE_{s_{AMISE}^{KM}}(f_n^{KM}(\cdot))$, que depende de las ventanas s y b :

$$RE_{AMISE}^{KM,P}(s, b) = \frac{MISE(f_n^P(\cdot))}{MISE_{s_{AMISE}^{KM}}(f_n^{KM}(\cdot))}.$$

Valores de $RE_{AMISE}^{KM,P}(s, b)$ menores que 1 indican una mayor eficiencia del estimador presuavizado $f_n^P(\cdot)$ calculado con el par de ventanas (s, b) con respecto al estimador tipo núcleo con pesos Kaplan-Meier $f_n^{KM}(\cdot)$ para la correspondiente ventana $AMISE$. Para ello se ha simulado mediante Montecarlo $MISE(f_n^P(\cdot))$ a partir de $m = 500$ muestras de tamaños $n = 50, 100, 200$ y 1000 .

En las figuras 4.3-4.6 se representa, mediante curvas de nivel, el valor de $RE_{AMISE}^{KM,P}(\cdot, \cdot)$ en una rejilla de valores de s y b para los cuatro modelos considerados y con los cuatro tamaños muestrales $n = 50, 100, 200$ y 1000 . Se marca con un aspa (\times) el punto que indica las ventanas s_{AMISE}^P y b_{AMISE}^P .

En todos los modelos simulados existe una región muy amplia de valores para el par de ventanas (s, b) con las que el estimador presuavizado $f_n^P(\cdot)$ es más eficiente que el estimador de Kaplan-Meier suavizado. Esta ineficiencia del estimador $f_n^{KM}(\cdot)$ es excepcional en el modelo 1 puesto que, dependiendo del tamaño muestral, $RE_{AMISE}^{KM,P}(\cdot, \cdot)$ puede tomar valores incluso menores que 0.5. Podemos ver además en las figuras 4.3-4.6 que el valor teórico de las ventanas s_{AMISE} y b_{AMISE} se encuentran, con holgura, dentro de la región de ventanas para las que $RE_{AMISE}^{KM,P}(\cdot, \cdot) < 1$. Se espera, por tanto, que cualquier selector consistente para las ventanas s y b proporcione estimaciones presuavizadas de $f(\cdot)$ más eficientes que el estimador de Kaplan-Meier suavizado aún con el valor exacto de su correspondiente ventana $AMISE$, independientemente de si los datos proceden de un modelo con eficiencia de primer o tercer orden.

El buen comportamiento de las ventanas $AMISE$ mostrado en estos gráficos motiva la definición y estudio, en la siguiente sección, de un selector de las ventanas de tipo plug-in.

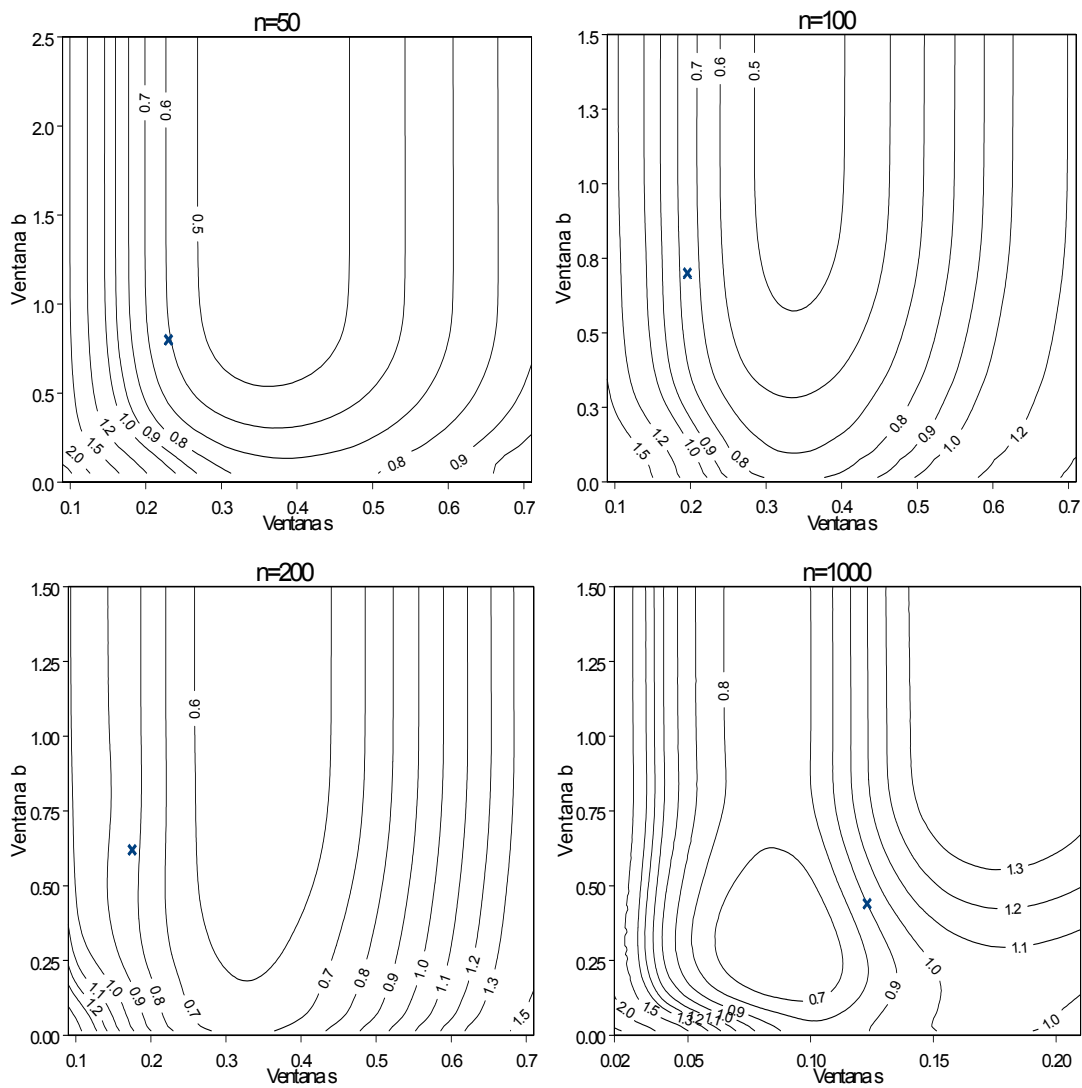


Figura 4.3. Eficiencia relativa $RE_{AMISE}^{KM,P}(s, b)$ para el modelo 1 con tamaños muestrales $n = 50, 100, 200$ y 1000 . El valor de las ventanas s_{AMISE} y b_{AMISE} está marcado con un aspa (\times).

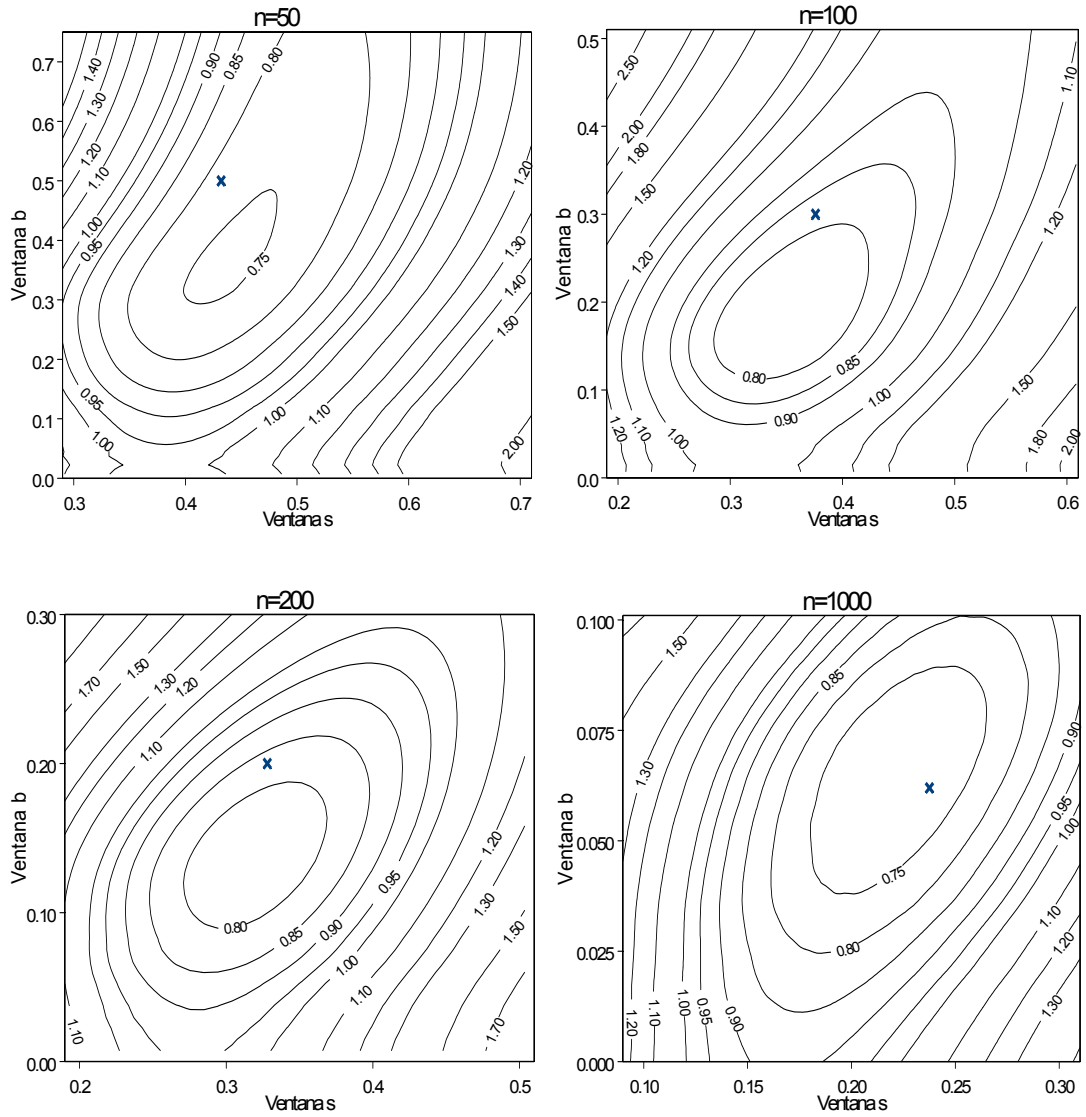


Figura 4.4. Eficiencia relativa $RE_{AMISE}^{KM,P}(s, b)$ para el modelo 2 con tamaños muestrales $n = 50, 100, 200$ y 1000 . El valor de las ventanas s_{AMISE} y b_{AMISE} está marcado con un aspa (\times).

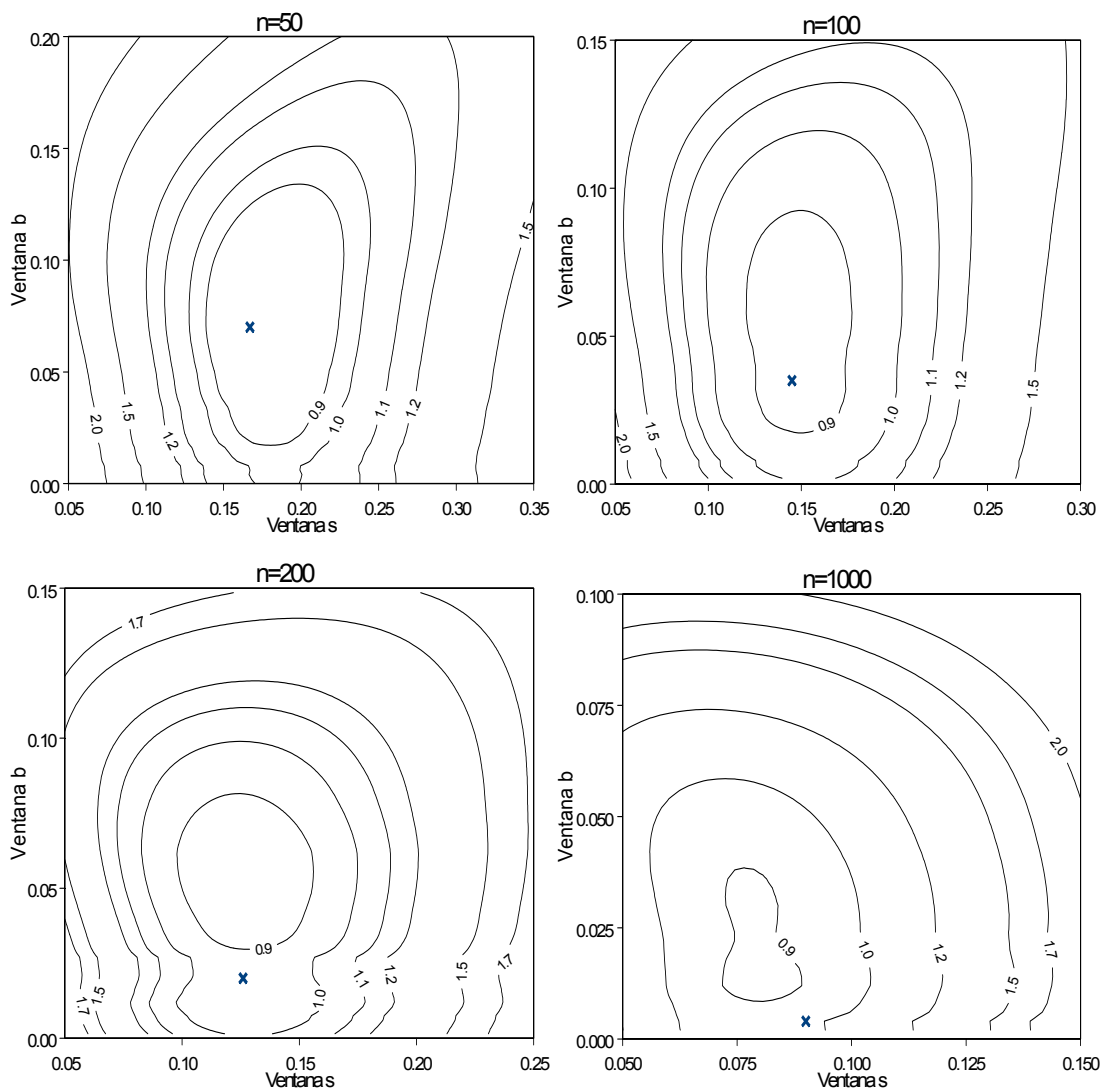


Figura 4.5. Eficiencia relativa $RE_{AMISE}^{KM,P}(s, b)$ para el modelo 3 con tamaños muestrales $n = 50, 100, 200$ y 1000 . El valor de las ventanas s_{AMISE} y b_{AMISE} está marcado con un aspa (\times).

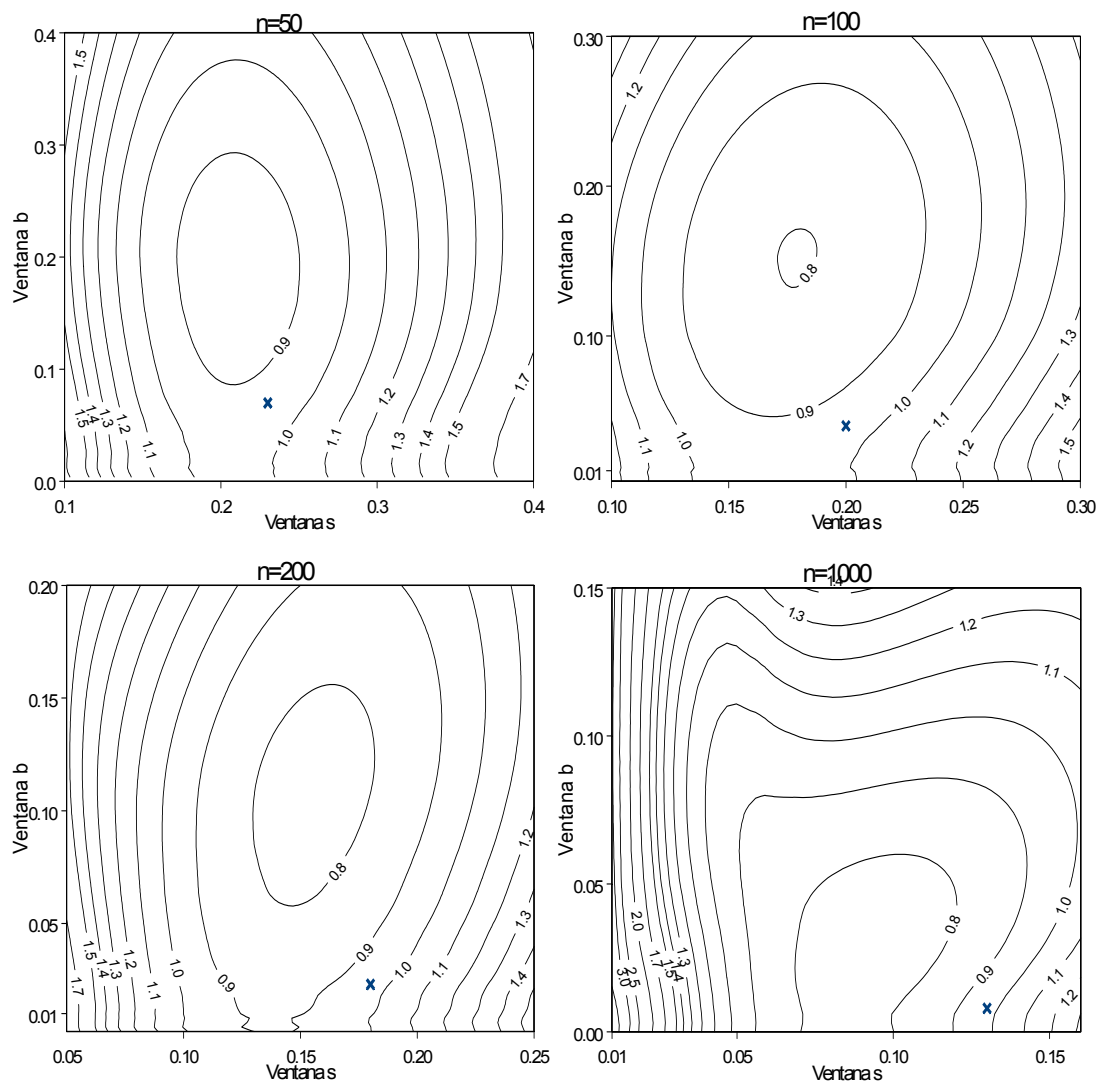


Figura 4.6. Eficiencia relativa $RE_{AMISE}^{KM,P}(s, b)$ para el modelo 4 con tamaños muestrales $n = 50, 100, 200$ y 1000 . El valor de las ventanas s_{AMISE} y b_{AMISE} está marcado con un aspa (\times).

4.3. Selectores de las ventanas de tipo plug-in

En esta sección estudiamos selectores de tipo plug-in de las ventanas s y b en el caso en el que ambas son del mismo orden, es decir, $b/s \rightarrow L > 0$, y por tanto la eficiencia del estimador presuavizado $f_n^P(\cdot)$ es de primer orden con respecto al estimador de Kaplan-Meier $f_n^{KM}(\cdot)$. Este selector se obtiene reemplazando, en las expresiones (4.12) y (4.13) de las ventanas óptimas, las integrales que dependen de funciones teóricas desconocidas por estimaciones adecuadas de las mismas.

Observación 4.3.1

En la definición de las ventanas plug-in se hará uso de la relación que hay entre las funciones razón de fallo de la variable observable $\lambda_H(\cdot)$ y la de interés $\lambda_F(\cdot)$ por medio de la función $p(\cdot)$:

$$\lambda_F(t) = p(t) \lambda_H(t) \Rightarrow p(t) = \frac{f(t)(1-H(t))}{h(t)(1-F(t))}.$$

Esta relación de $p(\cdot)$ con las funciones $f(\cdot)$, $F(\cdot)$, $h(\cdot)$ y $H(\cdot)$ permite escribir la función $A(\cdot, \cdot)$ dada en (4.5) de la forma:

$$\begin{aligned} A(t, L) &= A(t)p(t)c_K + A(t)(1-p(t))A_K(L) \\ &= \frac{f^2(t)}{h(t)}c_K + \left(f(t)\frac{1-F(t)}{1-H(t)} - \frac{f^2(t)}{h(t)} \right) A_K(L). \end{aligned}$$

Las ventanas plug-in propuestas son

$$\hat{s} = \left(\frac{\hat{c}_2(\hat{L})}{4\hat{c}_1(\hat{L})} \right)^{1/5} n^{-1/5} \quad \text{y} \quad \hat{b} = \hat{L} \times \hat{s}, \quad (4.30)$$

donde

$$\begin{aligned} \hat{c}_1(L) &= \int \left(\frac{1}{2}M_n(v) + L^2N_n(v) \right)^2 \omega(v) dv = \frac{1}{4}d_K^2 \int f_n^{P''}(v)^2 \omega(v) dv \quad (4.31) \\ &+ L^4 d_K^2 \int [(1-F_n^P(v))\alpha'_n(v) - f_n^P(v)\alpha_n(v)]^2 \omega(v) dv \\ &+ L^2 d_K^2 \int f_n^{P''}(v) [(1-F_n^P(v))\alpha'_n(v) - f_n^P(v)\alpha_n(v)] \omega(v) dv, \end{aligned}$$

$$\begin{aligned} \hat{c}_2(L) &= \int A_n(v, L) \omega(v) dv = c_K \int \frac{f_n^P(v)^2}{h_n(v)} \omega(v) dv \quad (4.32) \\ &+ A_K(L) \left(\int f_n^P(v) \frac{1-F_n^P(v)}{1-H_n(v) + \frac{1}{n}} \omega(v) dv - \int \frac{f_n^P(v)^2}{h_n(v)} \omega(v) dv \right) \end{aligned}$$

y el valor estimado de L_0 es $\hat{L} = \arg \min_{L>0} \psi_n(L)$, donde $\psi_n(t) = (\hat{c}_1(L)\hat{c}_2^4(L))^{1/5}$.

Estas funciones dependen del estimador presuavizado de la densidad $f_n^P(\cdot)$ con ventanas piloto s_0 y b_0 , y del estimador presuavizado de la función de distribución con ventana piloto b_0 , siendo $p_n(\cdot)$ el estimador de Nadaraya-Watson de $p(\cdot)$:

$$p_n(t) = \frac{\frac{1}{n} \sum_{i=1}^n K_{b_0}(t - Z_i) \delta_i}{\frac{1}{n} \sum_{i=1}^n K_{b_0}(t - Z_i)} = \frac{\psi_n(t)}{h_n(t)},$$

y $h_n(\cdot)$ el estimador de Parzen-Rosenblatt de $h(\cdot)$, ambos calculados con ventana b_0 .

Además, el estimador $f_n^{P''}(\cdot)$ no es más que la derivada segunda de $f_n^P(\cdot)$, $H_n(\cdot)$ es el estimador empírico de $H(\cdot)$, y los estimadores de las funciones $\alpha(\cdot)$, dada en (4.4), y $\alpha'(\cdot)$ son

$$\alpha_n(t) = \int_0^t \frac{\frac{1}{2} p_n''(v) h_n(v) + p_n'(v) h_n'(v)}{1 - H_n(v) + \frac{1}{n}} dv \quad \text{y} \quad \alpha_n'(t) = \frac{\frac{1}{2} p_n''(t) h_n(t) + p_n'(t) h_n'(t)}{1 - H_n(t) + \frac{1}{n}}.$$

Las funciones $p_n'(\cdot)$ y $p_n''(\cdot)$ son las derivadas primera y segunda, respectivamente, del estimador $p_n(\cdot)$:

$$p_n'(t) = \frac{\psi_n'(t) h_n(t) - \psi_n(t) h_n'(t)}{h_n^2(t)},$$

$$p_n''(t) = \frac{\psi_n''(t) h_n^2(t) - \psi_n(t) h_n''(t) h_n(t) - 2\psi_n'(t) h_n'(t) h_n(t) + 2\psi_n(t) h_n'(t)^2}{h_n^3(t)},$$

siendo

$$h_n^{(k)}(t) = \frac{1}{n} \sum_{i=1}^n K_{b_0}^{(k)}(t - Z_i), \quad \psi_n^{(k)}(t) = \frac{1}{n} \sum_{i=1}^n K_{b_0}^{(k)}(t - Z_i) \delta_i$$

y

$$K_{b_0}^{(k)}(t) = \frac{1}{b_0^{k+1}} K^{(k)}\left(\frac{t}{b_0}\right).$$

Las ventanas plug-in dependen entonces de dos ventanas piloto: la ventana de presuavizado b_0 , necesaria para estimar no paramétricamente las funciones $p(\cdot)$, $p'(\cdot)$, $p''(\cdot)$, $F(\cdot)$, $h(\cdot)$, $h'(\cdot)$ y $h''(\cdot)$, y la ventana de suavizado s_0 , que se usará para estimar $f_n^P(\cdot)$ a partir de $F_n^P(\cdot)$.

Observación 4.3.2

En el caso en que $\hat{L} = 0$, la ventanas plug-in propuestas son

$$\hat{s}_0 = \left(\frac{1}{4} \hat{c}_2(0) \hat{c}_1^{-1}(0) \right)^{1/5} n^{-1/5} \quad \text{y} \quad \hat{b}_0 = \tilde{L} \times \hat{s}_0,$$

siendo \tilde{L} el argumento donde alcanza su mínimo la función

$$\tilde{\psi}(L) = \hat{s}_0^4 \times \hat{c}_1(L) + \frac{1}{n \hat{s}_0} \hat{c}_2(L) - 2 \frac{\hat{s}_0 \times L}{n} \int C_n(v) \omega(v) dv,$$

es decir, la estimación tipo plug-in de $\overline{AMISE}(\cdot, \cdot)$ (menos el término en $\frac{1}{n}$), y donde

$$C_n(t) = e_K f_n^P(t)^2 \frac{p_n(t)(1-p_n(t))h_n(t)}{(1-H_n(t) + \frac{1}{n})^2}.$$

4.3.1. Consistencia de las ventanas plug-in

Para comprobar la consistencia de las ventanas dadas en (4.30), probaremos que las ventanas plug-in \hat{s} y \hat{b} se aproximan en probabilidad a las respectivas ventanas teóricas $MISE$, es decir,

$$\frac{\hat{s} - s_{MISE}}{s_{MISE}} \xrightarrow{P} 0 \quad \text{y} \quad \frac{\hat{b} - b_{MISE}}{b_{MISE}} \xrightarrow{P} 0$$

que, en virtud del teorema 4.1.2, equivalen a

$$\frac{\hat{s} - s_{AMISE}}{s_{AMISE}} \xrightarrow{P} 0 \quad \text{y} \quad \frac{\hat{b} - b_{AMISE}}{b_{AMISE}} \xrightarrow{P} 0.$$

Para ello basta demostrar

$$\hat{L} \xrightarrow{P} L_0, \quad \hat{c}_1(\hat{L}) \xrightarrow{P} c_1(L_0) \quad \text{y} \quad \hat{c}_2(\hat{L}) \xrightarrow{P} c_2(L_0). \quad (4.33)$$

En la demostración serán necesarias las hipótesis sobre las funciones $K(\cdot)$, $H(\cdot)$, $p(\cdot)$ y $f(\cdot)$ ya enunciadas en capítulos anteriores, junto con las siguientes adicionales:

(p3) La función $p(\cdot)$ verifica que $p(0) = 1$, y existe una constante positiva $\varepsilon > 0$ tal que $\varepsilon = \sup \{t > 0 : p(x) = 1 \quad \forall x \in [0, t]\} > 0$.

Las ventanas piloto s_0 de suavizado y b_0 de presuavizado verifican

(v5) $s \rightarrow 0$ y $ns^2 \rightarrow \infty$.

(v6) $b \rightarrow 0$ y $nb \rightarrow \infty$.

(v7) $nb^5 \log(1/b)^{-1} \rightarrow \infty$.

(v8) $nbs^3 \rightarrow \infty$ y $b^4 s^{-3} \rightarrow 0$.

(v9) $nbs(\log n)^{-3} \rightarrow \infty$, $b^4 s^{-1}(\log n)^2 \rightarrow 0$ y $nb^{-3} s^2(\log n)^5 \rightarrow \infty$.

La condición (p3) es de tipo técnico, pero no excesivamente restrictiva en la aplicación a datos reales, puesto que implica que un tiempo de vida no puede estar censurado por un número arbitrariamente pequeño.

Por su parte, la hipótesis (v7) implica la consistencia uniforme de las segundas derivadas de los estimadores tipo núcleo $p_n(\cdot)$ y $h_n(\cdot)$. Finalmente, (v9) garantiza la despreciabilidad del término de error $e_n(\cdot)$ dado en la representación del teorema 3.2.1. La hipótesis (v4), o también (v8), implican (v9).

Propiedad 4.3.1

Sea $I \left[\widehat{L}, L_0 \right]$ el intervalo de extremos \widehat{L} y L_0 . Si las funciones $\widehat{c}_1(\cdot)$ y $\widehat{c}_2(\cdot)$ son derivables con derivada uniformemente acotada, en probabilidad, en $L \in I \left[\widehat{L}, L_0 \right]$, entonces para probar las convergencias (4.33) basta demostrar lo siguiente:

- (i) $\widehat{L} \xrightarrow{P} L_0$
- (ii) $\widehat{c}_1(L) \xrightarrow{P} c_1(L) \quad \forall L > 0$
- (iii) $\widehat{c}_2(L) \xrightarrow{P} c_2(L) \quad \forall L > 0$

Demostración.

Por el teorema del valor medio, se tiene

$$\left| \widehat{c}_1(\widehat{L}) - \widehat{c}_1(L_0) \right| \leq \left| \widehat{L} - L_0 \right| \sup_{L \in I[\widehat{L}, L_0]} \left| \frac{\partial}{\partial L} \widehat{c}_1(L) \right|.$$

Dado que, por hipótesis, $\widehat{c}_1(\cdot)$ es derivable con derivada uniformemente acotada en probabilidad en $L \in I \left[\widehat{L}, L_0 \right]$, la condición (i) implica $\widehat{c}_1(\widehat{L}) - \widehat{c}_1(L_0) \xrightarrow{P} 0$.

Por otra parte, de la condición (ii) se obtiene $\widehat{c}_1(L_0) \xrightarrow{P} c_1(L_0)$. De esta convergencia y de la anterior, se deduce la convergencia deseada:

$$\widehat{c}_1(\widehat{L}) \xrightarrow{P} c_1(L_0).$$

La demostración para $\widehat{c}_2(\cdot)$ es análoga. ■

Propiedad 4.3.2

Bajo las condiciones (K1), (H3), (H4), (p2), (p3), (f1), (ω 1), (v5) – (v8) y suponiendo $\widehat{L} \xrightarrow{P} L_0$, las funciones $\widehat{c}_1(\cdot)$ y $\widehat{c}_2(\cdot)$ son derivables con derivada uniformemente acotada en probabilidad.

Demostración.

La función $\widehat{c}_1(\cdot)$ es derivable con derivada uniformemente acotada en probabilidad, puesto que

$$\begin{aligned} \frac{\partial}{\partial L} \widehat{c}_1(L) &= \frac{\partial}{\partial L} \left\{ \frac{1}{4} d_K^2 \int f_n^{P''}(v)^2 \omega(v) dv \right. \\ &\quad + L^4 d_K^2 \int \left[(1 - F_n^P(v)) \alpha'_n(v) - f_n^P(v) \alpha_n(v) \right]^2 \omega(v) dv \\ &\quad \left. + L^2 d_K^2 \int f_n^{P''}(v) \left[(1 - F_n^P(v)) \alpha'_n(v) - f_n^P(v) \alpha_n(v) \right] \omega(v) dv \right\} \\ &= 4L^3 d_K^2 \int \left[(1 - F_n^P(v)) \alpha'_n(v) - f_n^P(v) \alpha_n(v) \right]^2 \omega(v) dv \\ &\quad + 2L d_K^2 \int f_n^{P''}(v) \left[(1 - F_n^P(v)) \alpha'_n(v) - f_n^P(v) \alpha_n(v) \right] \omega(v) dv \end{aligned}$$

está uniformemente acotada en $L \in I[\widehat{L}, L_0]$ por

$$\begin{aligned} \sup_{L \in I[\widehat{L}, L_0]} \left| \frac{\partial}{\partial L} \widehat{c}_1(L) \right| &= 4\widetilde{L}^3 d_K^2 \int [(1 - F_n^P(v)) \alpha'_n(v) - f_n^P(v) \alpha_n(v)]^2 \omega(v) dv \quad (4.34) \\ &\quad + 2\widetilde{L} d_K^2 \int f_n^{P''}(v) |(1 - F_n^P(v)) \alpha'_n(v) - f_n^P(v) \alpha_n(v)| \omega(v) dv \end{aligned}$$

donde $\widetilde{L} = \max\{\widehat{L}, L_0\} = O_P(1)$. Bajo las condiciones (K1), (H3), (p2), (p3), (f1), (ω 1), (v5), (v7) y (v8), y en virtud de la observación 4.3.3, la desigualdad de Cauchy-Schwarz y los lemas 4.3.10, 4.3.11 y 4.3.12 que se demostrarán en la siguiente subsección, las integrales que aparecen en (4.34) son de orden $O_P(1)$.

La función $\widehat{c}_2(\cdot)$ es también derivable con derivada acotada en probabilidad, puesto que su derivada

$$\begin{aligned} \frac{\partial}{\partial L} \widehat{c}_2(L) &= \frac{\partial}{\partial L} \left[c_K \int \frac{f_n^P(v)^2}{h_n(v)} \omega(v) dv \right. \\ &\quad \left. + A_K(L) \left(\int f_n^P(v) \frac{1 - F_n^P(v)}{1 - H_n(v) + \frac{1}{n}} \omega(v) dv - \int \frac{f_n^P(v)^2}{h_n(v)} \omega(v) dv \right) \right] \\ &= \left(\int f_n^P(v) \frac{1 - F_n^P(v)}{1 - H_n(v) + \frac{1}{n}} \omega(v) dv - \int \frac{f_n^P(v)^2}{h_n(v)} \omega(v) dv \right) \frac{\partial}{\partial L} A_K(L) \end{aligned}$$

está uniformemente acotada en $L \in I[\widehat{L}, L_0]$ por

$$\sup_{L \in I[\widehat{L}, L_0]} \left| \frac{\partial}{\partial L} \widehat{c}_2(L) \right| = \left(\int f_n^P(v) \frac{1 - F_n^P(v)}{1 - H_n(v) + \frac{1}{n}} \omega(v) dv + \int \frac{f_n^P(v)^2}{h_n(v)} \omega(v) dv \right) \cdot C \quad (4.35)$$

donde la constante C es igual a

$$\begin{aligned} C &= \sup_{L > 0} \left| \frac{\partial}{\partial L} A_K(L) \right| \\ &= \sup_{L > 0} \left| \frac{\partial}{\partial L} \iiint K(u) K(v) K(w) K(u + L(v - w)) dudvdw \right| \\ &= \sup_{L > 0} \left| \iiint K(u) K(v) K(w) (v - w) K'(u + L(v - w)) dudvdw \right| \\ &\leq 2 \int |u| K(u) du \sup_{u \in [-1, 1]} |K'(u)|. \end{aligned}$$

Bajo las condiciones (K1), (H3), (H4), (p2), (p3), (f1), (ω 1), (v5) y (v6), y por los lemas 4.3.8 y 4.3.9 que establecen la acotación en probabilidad de las integrales en (4.35), el supremo en $L \in I[\widehat{L}, L_0]$ de la derivada de $\widehat{c}_2(\cdot)$ es de orden $O_P(1)$. ■

Propiedad 4.3.3

Si las funciones $\widehat{c}_1(\cdot)$ y $\widehat{c}_2(\cdot)$ verifican

$$\widehat{c}_1(L) \xrightarrow{P} c_1(L) \quad \text{y} \quad \widehat{c}_2(L) \xrightarrow{P} c_2(L) \quad \forall L > 0, \quad (4.36)$$

entonces $\widehat{L} \xrightarrow{P} L_0$.

Demostración.

Usando (4.36) se obtiene

$$\psi_n(L) = (\widehat{c}_1(L) \widehat{c}_2^4(L))^{1/5} \xrightarrow{P} \psi(L) = (c_1(L) c_2^4(L))^{1/5} \quad \forall L > 0. \quad (4.37)$$

Tendremos demostrada la convergencia en probabilidad $\widehat{L} \xrightarrow{P} L_0$ cuando probemos

$$\forall \varepsilon > 0, \quad P\left(\left|\widehat{L} - L_0\right| > \varepsilon\right) \rightarrow 0.$$

Fijemos un $\varepsilon > 0$. Puesto que $L_0 = \arg \min_{L>0} \psi(L)$, existe un $\delta > 0$ y existen dos valores $0 < a < b$ tales que

$$\psi(L) > \psi(L_0) + \delta \quad \forall L \in [a, b] \setminus (L_0 - \varepsilon, L_0 + \varepsilon). \quad (4.38)$$

Sea el suceso $A_n = \{|\psi_n(L_0) - \psi(L_0)| \geq \frac{\delta}{2}\}$. Por la aplicación de (4.37) en L_0 , tenemos que $\lim P(A_n) = 0$. Además,

$$\forall \omega \in \overline{A_n} \Rightarrow \psi_n(L_0) < \psi(L_0) + \frac{\delta}{2}.$$

Definimos ahora el suceso

$$B_n = \left\{ \sup_{L \in [a, b] \setminus (L_0 - \varepsilon, L_0 + \varepsilon)} |\psi_n(L) - \psi(L)| \geq \frac{\delta}{2} \right\},$$

del que también se tiene que $\lim P(B_n) = 0$, puesto que, por (4.37) y la continuidad de las funciones $\psi_n(\cdot)$ y $\psi(\cdot)$, se tiene

$$\sup_{L \in [a, b] \setminus (L_0 - \varepsilon, L_0 + \varepsilon)} |\psi_n(L) - \psi(L)| \xrightarrow{P} 0.$$

Además, considerando la desigualdad (4.38)

$$\omega \in \overline{B_n} \Rightarrow \psi_n(L) \geq \psi(L) - \frac{\delta}{2} > \psi(L_0) + \frac{\delta}{2} \quad \forall L \in [a, b] \setminus (L_0 - \varepsilon, L_0 + \varepsilon).$$

En resumen,

$$\omega \in \overline{A_n} \cap \overline{B_n} \Rightarrow \psi_n(L_0) < \psi(L_0) + \frac{\delta}{2} < \psi_n(L) \quad \forall L \in [a, b] \setminus (L_0 - \varepsilon, L_0 + \varepsilon).$$

Pero teniendo en cuenta que $\widehat{L} = \arg \min_{L>0} \psi_n(L)$, entonces

$$\omega \in \overline{A_n} \cap \overline{B_n} \Rightarrow \widehat{L} \in [L_0 - \varepsilon, L_0 + \varepsilon] \Leftrightarrow \left| \widehat{L} - L_0 \right| \leq \varepsilon,$$

de modo que

$$P\left(\left|\widehat{L} - L_0\right| > \varepsilon\right) = P\left(\overline{\overline{A_n} \cap \overline{B_n}}\right) = P(A_n \cup B_n) \leq P(A_n) + P(B_n) \rightarrow 0.$$

■

Antes de afrontar la demostración de las convergencias en probabilidad:

$$\widehat{c}_1(L) \xrightarrow{P} c_1(L) \quad \text{y} \quad \widehat{c}_2(L) \xrightarrow{P} c_2(L) \quad \forall L > 0,$$

probaremos previamente algunos lemas que se utilizarán en dicha demostración.

Lemas previos

En las demostraciones de la siguiente sección, será necesario aplicar el teorema de la convergencia dominada para calcular el límite cuando $n \rightarrow \infty$ de distintas integrales que dependen de n a través de las ventanas piloto. Por ello, en los siguientes lemas se demuestra que ciertas funciones, que aparecerán en los integrandos de dichas integrales, están acotadas por otra función, y se calculan sus límites cuando $n \rightarrow \infty$.

Lema 4.3.1

Supóngase la condición $(\omega 1)$ y sea $g(\cdot)$ cierta función y $g_n(\cdot)$ un estimador de la misma. Si la función $g(\cdot)$ y los momentos $\mathbb{E}[g_n(\cdot)]$ y $\mathbb{E}[g_n^2(\cdot)]$ son continuos, sus valores absolutos están acotados en t y además

$$\lim_{n \rightarrow \infty} \mathbb{E}[g_n(t)] = g(t) \quad \text{y} \quad \lim_{n \rightarrow \infty} \mathbb{E}[g_n^2(t)] = g^2(t),$$

entonces

$$\int (g_n(v) - g(v))^2 \omega(v) dv \xrightarrow{P} 0 \quad \text{y, por tanto,} \quad \int g_n^2(v) \omega(v) dv \xrightarrow{P} \int g^2(v) \omega(v) dv.$$

Demostración.

Si definimos la norma $\|g\|_2 = \int g^2(v) \omega(v) dv$, entonces por la desigualdad triangular se tiene que

$$\left| \|g_n\|_2 - \|g\|_2 \right| \leq \|g_n - g\|_2,$$

y por tanto

$$\int (g_n(v) - g(v))^2 \omega(v) dv \xrightarrow{P} 0 \Rightarrow \int g_n^2(v) \omega(v) dv \xrightarrow{P} \int g^2(v) \omega(v) dv.$$

Pero al ser $\int (g_n(v) - g(v))^2 \omega(v) dv$ una sucesión de variables aleatorias positivas, para demostrar su convergencia en probabilidad a cero basta probar que

$$\lim_{n \rightarrow \infty} \mathbb{E} \left[\int (g_n(v) - g(v))^2 \omega(v) dv \right] = 0,$$

donde

$$\begin{aligned} & \lim_{n \rightarrow \infty} \mathbb{E} \left[\int (g_n(v) - g(v))^2 \omega(v) dv \right] \\ &= \lim_{n \rightarrow \infty} \int \mathbb{E} \left[(g_n(v) - g(v))^2 \right] \omega(v) dv \\ &= \lim_{n \rightarrow \infty} \int \left(\text{Sesgo}(g_n(v))^2 + \text{Var}(g_n(v)) \right) \omega(v) dv \\ &= \lim_{n \rightarrow \infty} \int \left(\mathbb{E}[g_n(v)] - g(v) \right)^2 \omega(v) dv + \lim_{n \rightarrow \infty} \int \left(\mathbb{E}[g_n^2(v)] - \mathbb{E}[g_n(v)]^2 \right) \omega(v) dv. \end{aligned}$$

Si la función $g(\cdot)$ y los momentos $\mathbb{E}[g_n(\cdot)]$ y $\mathbb{E}[g_n^2(\cdot)]$ son continuos, su valor absoluto está acotado en t y si además

$$\lim_{n \rightarrow \infty} \mathbb{E}[g_n(t)] = g(t) \quad y \quad \lim_{n \rightarrow \infty} \mathbb{E}[g_n^2(t)] = g^2(t),$$

entonces, gracias a $(\omega 1)$, el teorema de la convergencia dominada garantiza que

$$\begin{aligned} \lim_{n \rightarrow \infty} \mathbb{E} \left[\int (g_n(v) - g(v))^2 \omega(v) dv \right] &= \int \lim_{n \rightarrow \infty} \left(\mathbb{E}[g_n(v)] - g(v) \right)^2 \omega(v) dv \\ &+ \int \lim_{n \rightarrow \infty} \left(\mathbb{E}[g_n^2(v)] - \mathbb{E}[g_n(v)]^2 \right) \omega(v) dv = 0. \end{aligned}$$

■

Lema 4.3.2

Sea $\varepsilon_1(\cdot)$ la función dada en la representación i.i.d. del estimador presuavizado $f_n^P(\cdot)$ en el teorema 3.2.1 del capítulo anterior:

$$\varepsilon_1(t) = (1 - F(t)) (g_1(t, Z_1) - g_2(t, Z_1) + g_3(t, Z_1, \delta_1)), \quad (4.39)$$

donde

$$g_1(t, Z) = \frac{p(t)}{1 - H(t)} (\mathbf{1}_{\{Z \leq t\}} - H(t)), \quad (4.40)$$

$$g_2(t, Z) = \int_0^t \frac{\mathbf{1}_{\{Z \leq v\}} - H(v)}{1 - H(v)} p'(v) dv, \quad (4.41)$$

$$g_3(t, Z, \delta) = \int_0^t K_b(v - Z) \frac{\delta - p(v)}{1 - H(v)} dv. \quad (4.42)$$

Entonces, bajo las hipótesis $(K1)$, $(H2)$, $(p2)$ y $(p3)$:

a) La función $\mathbb{E}[\varepsilon_1(\cdot)]$ es continua, acotada por una constante y verifica

$$\lim_{n \rightarrow \infty} \frac{1}{b^2} \mathbb{E}(\varepsilon_1(t)) = \alpha(t) d_K,$$

siendo $\alpha(\cdot)$ la función dada en (4.4).

b) La función $m(t_1, \dots, t_{n_0}) = \mathbb{E}[\varepsilon_1(t_1) \dots \varepsilon_1(t_{n_0})]$ es medible, acotada por una constante y su límite cuando $n \rightarrow \infty$ es finito.

Demostración.

Las funciones $g_1(\cdot, Z)$, $g_2(\cdot, Z)$ y $g_3(\cdot, Z, \delta)$ son continuas para todo $t > 0$, excepto la función $g_1(\cdot, Z)$ que lo es en todo $t > 0$ salvo en un conjunto de medida nula. Además, dado que $\mathbb{E}[g_1(t, Z)] = \mathbb{E}[g_2(t, Z)] = 0$, se tiene que

$$\mathbb{E}[\varepsilon_1(t)] = (1 - F(t)) \iint_0^t K_b(v - u) \frac{p(u) - p(v)}{1 - H(v)} h(u) dv du. \quad (4.43)$$

Entonces, $\mathbb{E}[\varepsilon_1(\cdot)]$ es continua y la función $m(\cdot, \dots, \cdot)$ es medible. Además, bajo la condición (p2), el valor absoluto de ambas funciones está acotado por una constante, puesto que para todo $t \in [0, t_H]$ y para todo par (Z, δ) , se tiene

$$\begin{aligned} |g_1(t, Z)| &\leq \frac{1}{1 - H(t_H)}, \\ |g_2(t, Z)| &\leq \frac{1}{1 - H(t_H)} \int_0^{t_H} |p'(v)| dv, \\ |g_3(t, Z, \delta)| &\leq \frac{1}{1 - H(t_H)}. \end{aligned}$$

Por tanto, también $\lim_{n \rightarrow \infty} |m(t_1, \dots, t_{n_0})| < \infty$.

Con respecto al límite cuando $n \rightarrow \infty$ de $\mathbb{E}[\varepsilon_1(\cdot)]$, realizamos un cambio de variable (paso (a)) y deshacemos la integral en dos sumandos según el intervalo de integración (paso (b)):

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{1}{b^2} \mathbb{E}[\varepsilon_1(t)] &= (1 - F(t)) \lim_{n \rightarrow \infty} \frac{1}{b^2} \iint_0^t K_b(v - u) \frac{p(u) - p(v)}{1 - H(v)} h(u) dv du \\ &\stackrel{(a)}{=} (1 - F(t)) \lim_{n \rightarrow \infty} \frac{1}{b^2} \iint_0^t K(w) \frac{p(v - bw) - p(v)}{1 - H(v)} h(v - bw) \mathbf{1}_{\{v - bw \geq 0\}} dv dw \\ &\stackrel{(b)}{=} (1 - F(t)) \lim_{n \rightarrow \infty} \frac{1}{b^2} \int_0^t \int_{-1}^1 K(w) \frac{p(v - bw) - p(v)}{1 - H(v)} h(v - bw) dw dv \\ &\quad - (1 - F(t)) \lim_{n \rightarrow \infty} \frac{1}{b^2} \int_0^b \int_{v/b}^1 K(w) \frac{p(v - bw) - p(v)}{1 - H(v)} h(v - bw) dw dv \\ &= (1 - F(t)) \lim_{n \rightarrow \infty} \left(\frac{1}{b^2} I_1(t, b) - \frac{1}{b^2} I_2(b) \right). \end{aligned}$$

Gracias a las condiciones (H2) y (p2), realizamos un desarrollo de Taylor de orden dos, con resto en forma integral, de las funciones $p(\cdot)$ y $h(\cdot)$ que aparecen en la integral $I_1(\cdot, \cdot)$, y por las propiedades de simetría del núcleo, se tiene

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{1}{b^2} I_1(t, b) &= \lim_{n \rightarrow \infty} \frac{1}{b^2} \iint_0^t K(w) \\ &\times \frac{-bw p'(v) + \frac{1}{2} b^2 w^2 p''(v) - \frac{1}{2} b^3 w^3 \int_0^1 (1-x_1)^2 p'''(v - bw x_1) dx_1}{1 - H(v)} \\ &\times \left(h(v) - bwh'(v) + b^2 w^2 \int_0^1 (1-x_2) h''(v - bw x_2) dx_2 \right) dv dw = \alpha(t) d_K. \end{aligned}$$

En el segundo término, $I_2(\cdot)$, llevamos a cabo el cambio de variable $u = \frac{v}{b}$ (paso (a)) y un desarrollo de Taylor de las funciones $p(\cdot)$, $H(\cdot)$ y $h(\cdot)$ en torno al cero (paso (b)), junto con la condición (p3) que implica $p'(0) = 0$, se tiene

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{1}{b^2} I_2(b) &\stackrel{(a)}{=} \lim_{n \rightarrow \infty} \frac{1}{b} \int_0^1 \int_u^1 K(w) \frac{p(bu - bw) - p(bu)}{1 - H(bu)} h(bu - bw) dudw \\ &\stackrel{(b)}{=} \lim_{n \rightarrow \infty} \int_0^1 \int_u^1 K(w) \left[(u-w) \int_0^1 p'(bux_1 - bwx_1) dx_1 - u \int_0^1 p'(byu) dy \right] \\ &\times \left(\frac{1}{1 - H(0)} + bu \int_0^1 \frac{h(byu)}{(1 - H(byu))^2} dy \right) \\ &\times \left(h(0) + b(u-w) \int_0^1 h'(bux_2 - bwx_2) dx_2 \right) dudw = 0. \end{aligned}$$

Se demuestra así lo establecido en el lema. ■

Lema 4.3.3

Sea $\sigma_n(\cdot)$ el tercer término en la representación i.i.d. del estimador presuavizado $f_n^P(\cdot)$ dada en el teorema 3.2.1 del capítulo anterior:

$$\sigma_n(t) = \frac{1}{ns} \sum_{i=1}^n \int \varepsilon_i(t - sv) K'(v) dv. \quad (4.44)$$

Entonces, bajo las hipótesis (K1), (H2), (p2), (p3) y (f1), su esperanza es continua, su valor absoluto está acotado y

$$\lim_{n \rightarrow \infty} \mathbb{E}[\sigma_n(t)] = 0.$$

Demostración.

La esperanza de la función $\sigma_n(\cdot)$ es igual a

$$\mathbb{E}[\sigma_n(t)] = \frac{1}{s} \int \mathbb{E}[\varepsilon_1(t - sv)] K'(v) dv, \quad (4.45)$$

siendo $\varepsilon_1(\cdot)$ la función dada en (4.39). Por el lema 4.3.2 y teniendo en cuenta las hipótesis (K1) y (p1), entonces para todo $t \in [0, t_H]$ la esperanza $\mathbb{E}[\sigma_n(\cdot)]$ está acotada por una constante. Además, es continua por serlo la función $\mathbb{E}[\varepsilon_1(\cdot)]$ dada en (4.43). Para calcular el límite, cuando $n \rightarrow \infty$, de $\mathbb{E}[\sigma_n(\cdot)]$ realizamos un desarrollo de Taylor de orden 1, con resto en forma integral, de la función $\mathbb{E}[\varepsilon_1(t - sv)]$:

$$\begin{aligned}\mathbb{E}[\sigma_n(t)] &= \frac{1}{s} \int \left(\mathbb{E}[\varepsilon_1(t)] - vs \int_0^1 \mathbb{E}[\varepsilon_1(t - vsx)]' dx \right) K'(v) dv \\ &= \int v K'(v) \int_0^1 \mathbb{E}[\varepsilon_1(t - vsx)]' dx dv,\end{aligned}$$

donde, aplicando el cambio un variable,

$$\begin{aligned}\frac{\partial}{\partial t} \mathbb{E}[\varepsilon_1(t)] &= -f(t) \iint_0^t K_b(v-u) \frac{p(u) - p(v)}{1 - H(v)} h(u) dv du \\ &\quad + \frac{1 - F(t)}{1 - H(t)} \int K_b(t-u) (p(u) - p(t)) h(u) du \\ &= -f(t) \iint_0^t K(w) \frac{p(v-wb) - p(v)}{1 - H(v)} h(v-wb) dw du \\ &\quad + \frac{1 - F(t)}{1 - H(t)} \int K(v) (p(t-vb) - p(t)) h(t-vb) dv\end{aligned}$$

es, por las hipótesis (H2), (p2) y (f1) continua y acotada en $t \in [0, t_H]$. Aplicamos entonces el teorema de la convergencia dominada:

$$\lim_{n \rightarrow \infty} \mathbb{E}[\sigma_n(t)] = \int v K'(v) \int_0^1 \lim_{n \rightarrow \infty} \mathbb{E}[\varepsilon_1(t - vsx)]' dx dv = - \lim_{n \rightarrow \infty} \mathbb{E}[\varepsilon_1(t)]'.$$

Para demostrar que $\lim_{n \rightarrow \infty} \mathbb{E}[\sigma_n(t)] = 0$ probaremos que el límite cuando $n \rightarrow \infty$ de cada uno de los sumandos en los que se descompone $\frac{\partial}{\partial t} \mathbb{E}[\varepsilon_1(t)]$ es cero.

Dada una función genérica $m(\cdot)$ dos veces continuamente diferenciable en el intervalo compacto $[0, t_H]$, usando desarrollos de Taylor se tiene que

$$\int_0^{t_H} K_b(t-u) m(u) du = m(t) + b^2 \iint_0^1 v^2 K(v) (1-x) m''(t-vbx) dx dv,$$

de modo que, si $m(t) = 0$ entonces se tiene que

$$\lim_{n \rightarrow \infty} \frac{1}{b^2} \int K_b(t-u) m(u) du = \frac{1}{2} d_K m''(t) < \infty.$$

Aplicamos este resultado al segundo sumando de $\frac{\partial}{\partial t} \mathbb{E}[\varepsilon_1(t)]$, de manera que

$$\lim_{n \rightarrow \infty} \frac{1}{b^2} \frac{1 - F(t)}{1 - H(t)} \int K_b(t-u) (p(u) - p(t)) h(u) du < \infty.$$

El primer sumando es también despreciable, puesto que para todo $t \in [0, t_H]$, y aplicando de nuevo el teorema de la convergencia dominada, se tiene que

$$\begin{aligned} & \lim_{n \rightarrow \infty} f(t) \int_0^t \int_0^t K_b(v-u) \frac{p(u) - p(v)}{1 - H(v)} h(u) dudv \\ &= f(t) \lim_{n \rightarrow \infty} \int_0^t \int_0^t K(w) \frac{p(v-bw) - p(v)}{1 - H(v)} h(v-bw) \mathbf{1}_{\{v-bw \geq 0\}} dw dv = 0. \end{aligned}$$

De esta forma, se obtiene $\mathbb{E}[\sigma_n(t)] = o(1)$. ■

Lema 4.3.4

Sea $\sigma_n(\cdot)$ la función dada en (4.44). Bajo las hipótesis (K1), (H2), (p2), (p3), (f1) y (v5) su momento de orden dos es continuo, su valor absoluto está acotado y

$$\lim_{n \rightarrow \infty} \mathbb{E}[\sigma_n^2(t)] = 0.$$

Demostración.

El momento de orden dos de la función $\sigma_n(\cdot)$ es igual a

$$\begin{aligned} \mathbb{E}[\sigma_n^2(t)] &= \frac{1}{n^2 s^2} \sum_{i,j=1}^n \iint \mathbb{E}[\varepsilon_i(t-us) \varepsilon_j(t-vs)] K'(u) K'(v) dudv \\ &= \frac{n(n-1)}{n^2 s^2} \iint \mathbb{E}[\varepsilon_1(t-us) \varepsilon_2(t-vs)] K'(u) K'(v) dudv \\ &\quad + \frac{1}{n s^2} \iint \mathbb{E}[\varepsilon_1(t-us) \varepsilon_1(t-vs)] K'(u) K'(v) dudv. \end{aligned}$$

Puesto que las variables Z_1 y Z_2 son independientes también lo son $\varepsilon_1(\cdot)$ y $\varepsilon_2(\cdot)$, de modo que, aplicando (4.45), el primer sumando es igual a

$$\frac{n-1}{n s^2} \left(\int \mathbb{E}[\varepsilon_1(t-us)] K'(u) du \right)^2 = \frac{n-1}{n s^2} (s \mathbb{E}[\sigma_n(t)])^2 = \frac{n-1}{n} \mathbb{E}[\sigma_n(t)]^2,$$

que es, por el lema 4.3.3, continuo, acotado y con límite cero.

Con respecto al segundo sumando de $\mathbb{E}[\sigma_n^2(\cdot)]$, aplicamos el lema 4.3.2 de modo que también es continuo, acotado y, bajo la condición (v5), su límite es cero. ■

Lema 4.3.5

Sea $\overline{f_n^P}(\cdot)$ la representación i.i.d. del estimador presuavizado $f_n^P(\cdot)$, dada en el teorema 3.2.1 del capítulo anterior. Entonces, bajo las hipótesis (K1), (H2), (p1) y (f1) su esperanza $\mathbb{E}[\overline{f_n^P}(\cdot)]$ es continua, su valor absoluto está acotado y

$$\lim_{n \rightarrow \infty} \mathbb{E}[\overline{f_n^P}(t)] = f(t).$$

Demostación.

La esperanza de la función $\overline{f_n^P}(\cdot)$ es igual a

$$\mathbb{E} \left[\overline{f_n^P}(t) \right] = f(t) + \beta_n(t) + \mathbb{E}[\sigma_n(t)],$$

siendo

$$\beta_n(t) = \int f(t - sv) K(v) dv - f(t),$$

y $\sigma_n(\cdot)$ dada en (4.44). Bajo las hipótesis (K1) y (f1), los dos primeros sumandos en la expresión de $\mathbb{E} \left[\overline{f_n^P}(\cdot) \right]$ son continuos y acotados en $t \in [0, t_H]$. Además,

$$\begin{aligned} \lim_{n \rightarrow \infty} (\beta_n(t) + f(t)) &= \lim_{n \rightarrow \infty} \int K_s(t - v) f(v) dv \\ &= \lim_{n \rightarrow \infty} \int_{-1}^{t/s} K(u) f(t - su) du = f(t). \end{aligned}$$

La demostración concluye teniendo en cuenta el lema 4.3.3. ■

Lema 4.3.6

Sea $\overline{f_n^P}(\cdot)$ la representación i.i.d. del estimador presuavizado $f_n^P(\cdot)$, dada en el teorema 3.2.1 del capítulo anterior. Entonces, bajo las hipótesis (K1), (H2), (p2), (p3), (f1) y (v5) su momento de orden dos $\mathbb{E} \left[\overline{f_n^P}(\cdot)^2 \right]$ es continuo, su valor absoluto está acotado y

$$\lim_{n \rightarrow \infty} \mathbb{E} \left[\overline{f_n^P}(t)^2 \right] = f^2(t).$$

Demostación.

El momento de orden dos de $\overline{f_n^P}(\cdot)$ es

$$\mathbb{E} \left[\overline{f_n^P}(t)^2 \right] = (f(t) + \beta_n(t))^2 + \mathbb{E}[\sigma_n^2(t)] + 2(f(t) + \beta_n(t)) \mathbb{E}[\sigma_n(t)].$$

El primer sumando es continuo, acotado y su límite es

$$\lim_{n \rightarrow \infty} (f(t) + \beta_n(t))^2 = \lim_{n \rightarrow \infty} \left(\int K_s(t - v) f(v) dv \right)^2 = f^2(t),$$

mientras que, a partir de los resultados de los lemas 4.3.3 y 4.3.4, se tiene que también los dos últimos términos de $\mathbb{E} \left[\overline{f_n^P}(\cdot)^2 \right]$ son continuos, acotados y de límite cero. ■

Lema 4.3.7

Bajo las condiciones (K1), (H3), (p2) y (f1), el momento de orden dos de $\sigma_n''(\cdot)$ es una función continua, acotada para todo $t \in [0, t_H]$ y

$$\lim_{n \rightarrow \infty} \mathbb{E} \left[\sigma_n''(t)^2 \right] = 0.$$

Demostración.

La derivada segunda de la función $\sigma_n(\cdot)$ es igual a

$$\sigma_n''(t) = \frac{1}{ns^3} \sum_{i=1}^n \int \varepsilon_i(t - vs) K'''(v) dv,$$

siendo $\varepsilon_1(\cdot)$ la función dada en (4.39). Entonces, el momento de orden dos de $\sigma_n''(\cdot)$ es

$$\begin{aligned} \mathbb{E} \left[\sigma_n''(t)^2 \right] &= \frac{1}{n^2 s^6} \sum_{i=1}^n \sum_{j=1}^n \iint \mathbb{E} [\varepsilon_i(t - us) \varepsilon_j(t - vs)] K'''(u) K'''(v) dudv \\ &= \frac{n(n-1)}{n^2} \left(\frac{1}{s^3} \int \mathbb{E} [\varepsilon_1(t - vs)] K'''(v) dv \right)^2 \\ &\quad + \frac{1}{ns^6} \iint \mathbb{E} [\varepsilon_1(t - us) \varepsilon_1(t - vs)] K'''(u) K'''(v) dudv. \end{aligned} \quad (4.46)$$

Probaremos que cada uno de estos sumandos es continuo, acotado y que su límite cuando $n \rightarrow \infty$ es cero.

Bajo las hipótesis (H3), (p2) y (f1), la esperanza de la función $\varepsilon_1(\cdot)$ dada en (4.43) es una función tres veces derivable en $t \in [0, t_H]$, de modo que para calcular la integral del primer sumando en (4.46), realizamos una integración por partes:

$$\frac{1}{s^3} \int \mathbb{E} [\varepsilon_1(t - vs)] K'''(v) dv = \int \mathbb{E} [\varepsilon_1(t - vs)]''' K(v) dv.$$

Basta probar que la derivada tercera de $\mathbb{E}[\varepsilon_1(\cdot)]$ es continua, acotada para todo $t \in [0, t_H]$ y $\lim_{n \rightarrow \infty} \mathbb{E}[\varepsilon_1(t)]''' = 0$. Calculamos ahora la derivada tercera de $\mathbb{E}[\varepsilon_1(\cdot)]$:

$$\begin{aligned} \frac{\partial^3}{\partial t^3} \mathbb{E} [\varepsilon_1(t)] &= -f''(t) \mathbb{E} [g_3(t, Z_1, \delta_1)] - 3f'(t) \frac{\partial}{\partial t} \mathbb{E} [g_3(t, Z_1, \delta_1)] \\ &\quad - 3f(t) \frac{\partial^2}{\partial t^2} \mathbb{E} [g_3(t, Z_1, \delta_1)] + (1 - F(t)) \frac{\partial^3}{\partial t^3} \mathbb{E} [g_3(t, Z_1, \delta_1)]. \end{aligned}$$

donde

$$\begin{aligned} \mathbb{E} [g_3(t, Z, \delta)] &= \iint_0^t K_b(v - u) \frac{p(u) - p(v)}{1 - H(v)} h(u) dvdu, \\ \frac{\partial}{\partial t} \mathbb{E} [g_3(t, Z, \delta)] &= \frac{1}{1 - H(t)} \int K_b(t - u) (p(u) - p(t)) h(u) du, \\ \frac{\partial^2}{\partial t^2} \mathbb{E} [g_3(t, Z, \delta)] &= \frac{h(t)}{(1 - H(t))^2} \int K_b(t - u) p(u) h(u) du \\ &\quad + \frac{1}{1 - H(t)} \int K_b'(t - u) p(u) h(u) du - \left(\frac{p(t)}{1 - H(t)} \right)' \int K_b(t - u) h(u) du \\ &\quad - \frac{p(t)}{1 - H(t)} \int K_b'(t - u) h(u) du, \end{aligned}$$

$$\begin{aligned} \frac{\partial^3}{\partial t^3} \mathbb{E}[g_3(t, Z, \delta)] &= \left(\frac{h(t)}{(1-H(t))^2} \right)' \int K_b(t-u) p(u) h(u) du \\ &+ 2 \frac{h(t)}{(1-H(t))^2} \int K_b'(t-u) p(u) h(u) du - \left(\frac{p(t)}{1-H(t)} \right)'' \int K_b(t-u) h(u) du \\ &- 2 \left(\frac{p(t)}{1-H(t)} \right)' \int K_b'(t-u) h(u) du \\ &+ \frac{1}{1-H(t)} \int K_b''(t-u) p(u) h(u) du - \frac{p(t)}{1-H(t)} \int K_b''(t-u) h(u) du. \end{aligned}$$

Aplicando un cambio de variable (a) e integración por partes (b), la siguiente integral dependiente del núcleo, $K(\cdot)$, y de una función genérica, $m(\cdot)$, es igual a:

$$\begin{aligned} \int K_b^{(k)}(t-u) m(u) du &\stackrel{(a)}{=} \frac{1}{b^k} \int K^{(k)}(v) m(t-vb) dv \stackrel{(b)}{=} \int K(v) m^{(k)}(t-vb) dv \\ &= m^{(k)}(t) + b^2 \int \int_0^1 (1-x) v^2 K(v) m^{(k+2)}(t-vbx) dx dv, \end{aligned}$$

de modo que es continua y está acotada si $m(\cdot)$ es $k+2$ veces derivable con $m^{(k+2)}(\cdot)$ continua y acotada. Además,

$$\lim_{n \rightarrow \infty} \int K_b^{(k)}(t-u) m(u) du = m^{(k)}(t).$$

Teniendo esta propiedad en cuenta, y bajo las hipótesis (H3), (p2) y (f1), la derivada tercera de la esperanza $\mathbb{E}[\varepsilon_1(\cdot)]$ es continua, acotada y $\lim_{n \rightarrow \infty} \mathbb{E}[\varepsilon_1(t)]''' = 0$.

Por tanto, el primer sumando en (4.46) también es continuo, acotado y su límite cuando $n \rightarrow \infty$ es cero. Demostraremos que el segundo sumando verifica las mismas propiedades.

Definamos ahora la función $m(\cdot, \cdot)$ de la siguiente manera:

$$m(x, y) =: \mathbb{E}[\varepsilon_1(x) \varepsilon_1(y)],$$

de modo que el segundo sumando en (4.46) es

$$\frac{1}{ns^6} \int \left(\int m(t-us, t-vs) K'''(u) du \right) K'''(v) dv.$$

La expresión desarrollada de la función $m(\cdot, \cdot)$ es

$$\begin{aligned} m(x, y) &= (1-F(x))(1-F(y)) \{ \mathbb{E}[g_1(x, Z) g_1(y, Z)] - \mathbb{E}[g_1(x, Z) g_2(y, Z)] \\ &+ \mathbb{E}[g_1(x, Z) g_3(y, Z, \delta)] + \mathbb{E}[g_2(x, Z) g_2(y, Z)] - \mathbb{E}[g_2(x, Z) g_1(y, Z)] \\ &- \mathbb{E}[g_2(x, Z) g_3(y, Z, \delta)] + \mathbb{E}[g_3(x, Z, \delta) g_3(y, Z, \delta)] \\ &+ \mathbb{E}[g_3(x, Z, \delta) g_1(y, Z)] - \mathbb{E}[g_3(x, Z, \delta) g_2(y, Z)] \}, \end{aligned}$$

siendo $g_1(\cdot, \cdot)$, $g_2(\cdot, \cdot)$ y $g_3(\cdot, \cdot, \cdot)$ las funciones dadas en (4.40)-(4.42). Calculamos ahora cada una de las esperanzas en las que se descompone la función $m(\cdot, \cdot)$:

$$\begin{aligned}\mathbb{E}[g_1(x, Z)g_1(y, Z)] &= \frac{p(x)}{1-H(x)}\frac{p(y)}{1-H(y)}(H(x \wedge y) - H(x)H(y)). \\ \mathbb{E}[g_2(x, Z)g_2(y, Z)] &= \int_0^x \int_0^y \frac{H(u \wedge v) - H(u)H(v)}{(1-H(u))(1-H(v))} p'(u)p'(v) dudv. \\ \mathbb{E}[g_3(x, Z, \delta)g_3(y, Z, \delta)] &= \int_0^x \int_0^y \int K_b(v-u)K_b(w-u) \\ &\quad \times \frac{p(u) - p(v)(p(v) + p(w)) + p(v)p(w)}{(1-H(v))(1-H(w))} h(u) dudvdw. \\ \mathbb{E}[g_1(x, Z)g_3(y, Z, \delta)] &= \frac{p(x)}{1-H(x)} \int_0^x \int_0^y K_b(v-u) \frac{p(u) - p(v)}{1-H(v)} h(u) dvdu \\ &\quad - \frac{p(x)H(x)}{1-H(x)} \int_0^y \int_0^y K_b(v-u) \frac{p(u) - p(v)}{1-H(v)} h(u) dvdu.\end{aligned}$$

Para calcular la esperanza de $g_1(x, Z)g_2(y, Z)$ y $g_2(x, Z)g_3(y, Z, \delta)$ usamos las propiedades 3.2.1 y 3.2.3 del capítulo anterior:

$$\begin{aligned}\mathbb{E}[g_1(x, Z)g_2(y, Z)] &= \frac{p(x)}{1-H(x)} \mathbb{E} \left[(\mathbf{1}_{\{Z \leq x\}} - H(x)) \int_0^y \frac{\mathbf{1}_{\{Z \leq v\}} - H(v)}{1-H(v)} p'(v) dv \right] \\ &= \frac{p(x)H(x)}{1-H(x)} \left(\frac{p(y)}{1-H(y)} - \gamma(y) \right) - \gamma(x)p(x) \\ &\quad - \frac{p(x)H(x)}{1-H(x)} \int_0^y \frac{H(v)p'(v)}{1-H(v)} dv. \\ \mathbb{E}[g_2(x, Z)g_3(y, Z, \delta)] &= \mathbb{E} \left[\int_0^x \frac{\mathbf{1}_{\{Z \leq v\}} - H(v)}{1-H(v)} p'(v) dv \int_0^y K_b(v-Z) \frac{\delta - p(v)}{1-H(v)} dv \right] \\ &= \left(\frac{p(x)}{1-H(x)} - \gamma(x) \right) \int_0^y K_b(v-u) \frac{p(u) - p(v)}{1-H(v)} h(u) dvdu \\ &\quad - \int_0^y \int_0^y K_b(v-u) \left(\frac{p(u)}{1-H(u)} - \gamma(u) \right) \frac{p(u) - p(v)}{1-H(v)} h(u) dvdu \\ &\quad - \int_0^x \frac{H(v)p'(v)}{1-H(v)} dv \int_0^y K_b(v-u) \frac{p(u) - p(v)}{1-H(v)} h(u) dvdu.\end{aligned}$$

siendo

$$\gamma(t) = \int_0^t \frac{p(v)h(v)}{(1-H(v))^2} dv.$$

Bajo las hipótesis (H3), (p2) y (f1), todas estas esperanzas son funciones continuas (salvo en una región de medida nula) y derivables hasta, al menos, el tercer orden. Por lo tanto, mediante sucesivas integraciones por partes, se tiene

$$\begin{aligned} & \frac{1}{ns^6} \int \left(\int m(t-us, t-vs) K'''(u) du \right) K'''(v) dv \\ &= \frac{1}{ns^3} \int \left(\int \frac{\partial^3}{\partial x^3} m(t-us, t-vs) K(u) du \right) K'''(v) dv \\ &= \frac{1}{n} \iint \frac{\partial^6}{\partial x^3 \partial y^3} m(t-us, t-vs) K(u) K(v) dudv. \end{aligned}$$

Puesto que $m(\cdot, \cdot)$ tiene las derivadas parciales terceras continuas, acotadas y con límite finito, el segundo sumando en (4.46) también es continuo, acotado y su límite cuando $n \rightarrow \infty$ es cero. De todo ello se concluye la demostración del lema. ■

Demostración de la consistencia

Para demostrar la consistencia de las ventanas plug-in, dadas por (4.30), comenzaremos demostrando que

$$\widehat{c}_2(L) \xrightarrow{P} c_2(L) \quad \forall L > 0,$$

siendo $c_2(\cdot)$ y $\widehat{c}_2(\cdot)$ las funciones (4.21) y (4.32) respectivamente. Para ello probaremos, en los lemas 4.3.8. y 4.3.9, la convergencia en probabilidad de cada una de las integrales en las que se descompone $\widehat{c}_2(\cdot)$.

Lema 4.3.8

Bajo las hipótesis (K1), (H2), (H4), (p2), (p3), (f1), (ω 1), (v5), (v6) y (v9) se tiene

$$\int_{\frac{\varepsilon}{2}}^{t_H} \frac{f_n^P(v)^2}{h_n(v)} \omega(v) dv \xrightarrow{P} \int_{\frac{\varepsilon}{2}}^{t_H} \frac{f^2(v)}{h(v)} \omega(v) dv.$$

Demostración.

Paso 1. El primer paso consiste en sustituir el denominador aleatorio $h_n(\cdot)$ en la integral por la densidad teórica $h(\cdot)$ más un término despreciable en probabilidad:

$$\int_{\frac{\varepsilon}{2}}^{t_H} \frac{f_n^P(v)^2}{h_n(v)} \omega(v) dv = \int_{\frac{\varepsilon}{2}}^{t_H} \frac{f_n^P(v)^2}{h(v)} \omega(v) dv + o_P(1).$$

Para ello descomponemos la integral de la siguiente manera:

$$\begin{aligned} \int_{\frac{\varepsilon}{2}}^{t_H} \frac{f_n^P(v)^2}{h_n(v)} \omega(v) dv &= \int_{\frac{\varepsilon}{2}}^{t_H} \frac{f_n^P(v)^2}{h(v)} \omega(v) dv + \int_{\frac{\varepsilon}{2}}^{t_H} \frac{f_n^P(v)^2}{h(v)} \frac{h(v) - h_n(v)}{h(v)} \omega(v) dv \\ &\quad + \int_{\frac{\varepsilon}{2}}^{t_H} \frac{f_n^P(v)^2}{h^2(v)} \frac{(h(v) - h_n(v))^2}{h_n(v)} \omega(v) dv. \end{aligned} \quad (4.47)$$

Veremos que los dos últimos sumandos son $o_P\left(\int f_n^P(v)^2 h^{-1}(v)\omega(v)dv\right)$ y por tanto, aplicando el lema 4.3.6 por el cual $\int f_n^P(v)^2 dv = O_P(1)$, son de orden $o_P(1)$.

Empezamos acotando el segundo sumando. Para ello, en el primer paso (paso (a)) extraemos los supremos de las funciones $h(\cdot)$ y $h(\cdot) - h_n(\cdot)$ y usamos la condición (H4), mientras que en el segundo (paso (b)) aplicamos el lema 1 de Mack y Silverman (1982) para el estimador de Parzen-Rosenblatt, $h_n(\cdot)$. Finalmente, en el paso (c), consideramos la hipótesis (v6):

$$\begin{aligned} \left| \int_{\frac{\varepsilon}{2}}^{t_H} \frac{f_n^P(v)^2}{h(v)} \frac{h(v) - h_n(v)}{h(v)} \omega(v) dv \right| &\stackrel{(a)}{\leq} \frac{1}{\mu} \sup_{0 \leq t \leq t_H} |h_n(t) - h(t)| \int_{\frac{\varepsilon}{2}}^{t_H} \frac{f_n^P(v)^2}{h(v)} \omega(v) dv \\ &\stackrel{(b)}{=} O_P\left(b^2 + (nb)^{-1/2} (\log \frac{1}{b})^{1/2}\right) \int_{\frac{\varepsilon}{2}}^{t_H} \frac{f_n^P(v)^2}{h(v)} \omega(v) dv \\ &\stackrel{(c)}{=} o_P\left(\int_{\frac{\varepsilon}{2}}^{t_H} \frac{f_n^P(v)^2}{h(v)} \omega(v) dv\right). \end{aligned}$$

Con respecto al último sumando en (4.47), en primer lugar, por la hipótesis (H4), se tiene para todo $t \in [0, t_H]$,

$$h_n(t) \geq h_n(t) - h(t) + \mu.$$

Se utiliza esta acotación (paso (a)), en el paso (b) extraemos el supremo de la función $h(\cdot) - h_n(\cdot)$ y aplicamos que es positivo, y el paso (c) es consecuencia del lema 1 de Mack y Silverman (1982) y de la hipótesis (v6):

$$\begin{aligned} \left| \int_{\frac{\varepsilon}{2}}^{t_H} \frac{f_n^P(v)^2}{h^2(v)} \frac{(h_n(v) - h(v))^2}{h_n(v)} \omega(v) dv \right| &\leq \left| \int_{\frac{\varepsilon}{2}}^{t_H} \frac{f_n^P(v)^2}{h^2(v)} \frac{(h_n(v) - h(v))^2}{h_n(v) - h(v) + \mu} \omega(v) dv \right| \\ &\stackrel{(b)}{\leq} \frac{1}{\mu} \sup_{0 \leq t \leq t_H} |h_n(t) - h(t)| \int_{\frac{\varepsilon}{2}}^{t_H} \frac{f_n^P(v)^2}{h(v)} \omega(v) dv \\ &\stackrel{(c)}{=} o_P\left(\int_{\frac{\varepsilon}{2}}^{t_H} \frac{f_n^P(v)^2}{h(v)} \omega(v) dv\right). \end{aligned}$$

Paso 2. A continuación, se va a aproximar la integral $\int f_n^P(v)^2 h^{-1}(v)\omega(v)dv$ por otra resultante de escribir la representación i.i.d. del estimador presuavizado $\overline{f_n^P}(\cdot)$ en lugar del propio estimador $f_n^P(\cdot)$:

$$\int_{\frac{\varepsilon}{2}}^{t_H} \frac{f_n^P(v)^2}{h(v)} \omega(v) dv = \int_{\frac{\varepsilon}{2}}^{t_H} \frac{\overline{f_n^P}(v)^2}{h(v)} \omega(v) dv + o_P(1).$$

Descomponemos esa integral en tres sumandos de la forma

$$\begin{aligned} \int_{\frac{\varepsilon}{2}}^{t_H} \frac{f_n^P(v)^2}{h(v)} \omega(v) dv &= \int_{\frac{\varepsilon}{2}}^{t_H} \frac{\overline{f_n^P}(v)^2}{h(v)} \omega(v) dv + \int_{\frac{\varepsilon}{2}}^{t_H} \frac{e_n^2(v)}{h(v)} \omega(v) dv \\ &\quad + 2 \int_{\frac{\varepsilon}{2}}^{t_H} \frac{\overline{f_n^P}(v) e_n(v)}{h(v)} \omega(v) dv, \end{aligned} \quad (4.48)$$

de los cuales los dos últimos son de orden $o_P(1)$.

Teniendo en cuenta la hipótesis (H4), el valor absoluto del segundo sumando en (4.48) está acotado por

$$\frac{1}{\mu} \left(\sup_{t \in [\frac{\varepsilon}{2}, t_H]} |e_n(t)|^2 \right) \int_{\frac{\varepsilon}{2}}^{t_H} \omega(v) dv,$$

y por tanto, por el teorema 3.2.1 y la hipótesis (v9), es de orden $o_P(1)$. Bajo las mismas hipótesis, el tercer sumando en (4.48) está acotado por

$$\begin{aligned} \frac{2}{\mu} \sup_{t \in [\frac{\varepsilon}{2}, t_H]} |e_n(t)| \int_{\frac{\varepsilon}{2}}^{t_H} \overline{f_n^P}(v) \omega(v) dv \\ \leq \frac{2}{\mu} \sup_{t \in [\frac{\varepsilon}{2}, t_H]} |e_n(t)| \left[\int_{\frac{\varepsilon}{2}}^{t_H} \overline{f_n^P}(v)^2 \omega(v) dv \right]^{1/2} \left[\int_{\frac{\varepsilon}{2}}^{t_H} \omega(v) dv \right]^{1/2}, \end{aligned}$$

de modo que, por el lema 4.3.6, también es una $o_P(1)$.

Paso 3. Por último, demostraremos

$$\int_{\frac{\varepsilon}{2}}^{t_H} \frac{\overline{f_n^P}(v)^2}{h(v)} \omega(v) dv = \int_{\frac{\varepsilon}{2}}^{t_H} \frac{f^2(v)}{h(v)} \omega(v) dv + o_P(1).$$

La convergencia se obtiene aplicando a las funciones $g_n(\cdot) = \overline{f_n^P}(\cdot) h(\cdot)^{-1/2}$ y $g(\cdot) = f(\cdot) h^{-1/2}(\cdot)$ el lema 4.3.1, y considerando los resultados de los lemas 4.3.5 y 4.3.6. ■

Lema 4.3.9

Bajo las condiciones (K1), (H2), (H4), (p2), (p3), (f1), (v1), (v5) y (v9) se tiene

$$\int_{\frac{\varepsilon}{2}}^{t_H} f_n^P(v) \frac{1 - F_n^P(v)}{1 - H_n(v) + \frac{1}{n}} \omega(v) dv \xrightarrow{P} \int_{\frac{\varepsilon}{2}}^{t_H} f(v) \frac{1 - F(v)}{1 - H(v)} \omega(v) dv. \quad (4.49)$$

Demostración.

Paso 1. El primer paso consiste en sustituir, en la integral, el denominador aleatorio $1 - H_n(\cdot) + n^{-1}$ por la función $1 - H(\cdot)$ más un término despreciable en probabilidad:

$$\int_{\frac{\varepsilon}{2}}^{t_H} f_n^P(v) \frac{1 - F_n^P(v)}{1 - H_n(v) + \frac{1}{n}} \omega(v) dv = \int_{\frac{\varepsilon}{2}}^{t_H} f_n^P(v) \frac{1 - F_n^P(v)}{1 - H(v)} \omega(v) dv \quad (4.50)$$

$$+ o_P \left(\int_{\frac{\varepsilon}{2}}^{t_H} f_n^P(v) \frac{1 - F_n^P(v)}{1 - H(v)} \omega(v) dv \right).$$

Para demostrarlo, descomponemos la integral en dos sumandos:

$$\int_{\frac{\varepsilon}{2}}^{t_H} f_n^P(v) \frac{1 - F_n^P(v)}{1 - H_n(v) + \frac{1}{n}} \omega(v) dv$$

$$= \int_{\frac{\varepsilon}{2}}^{t_H} f_n^P(v) \frac{1 - F_n^P(v)}{1 - H(v)} \omega(v) dv$$

$$+ \int_{\frac{\varepsilon}{2}}^{t_H} f_n^P(v) \frac{1 - F_n^P(v)}{1 - H_n(v) + \frac{1}{n}} \left(1 - \frac{1 - H_n(v) + \frac{1}{n}}{1 - H(v)} \right) \omega(v) dv.$$

El segundo término es

$$o_P \left(\int_{\frac{\varepsilon}{2}}^{t_H} f_n^P(v) \frac{1 - F_n^P(v)}{1 - H(v)} \omega(v) dv \right),$$

puesto que, teniendo en cuenta la siguiente desigualdad

$$1 - H_n(t) + \frac{1}{n} \geq 1 - H(t_H) + \frac{1}{n} - \sup_{t \in [0, t_H]} |H_n(t) - H(t)| \quad (4.51)$$

se tiene

$$\left| \int_{\frac{\varepsilon}{2}}^{t_H} f_n^P(v) \frac{1 - F_n^P(v)}{1 - H(v)} \left(\frac{H_n(v) - H(v) - \frac{1}{n}}{1 - H_n(v) + \frac{1}{n}} \right) \omega(v) dv \right|$$

$$\leq \frac{\sup_{t \in [0, t_H]} |H_n(t) - H(t)| + \frac{1}{n}}{1 - H(t_H) + \frac{1}{n} - \sup_{t \in [0, t_H]} |H_n(t) - H(t)|} \int_{\frac{\varepsilon}{2}}^{t_H} f_n^P(v) \frac{1 - F_n^P(v)}{1 - H(v)} \omega(v) dv$$

con

$$\sup_{0 \leq t \leq t_H} |H_n(t) - H(t)| = O_P(n^{-1/2}) = o_P(1).$$

Paso 2. El segundo paso consiste en sustituir los estimadores presuavizados $f_n^P(\cdot)$ y $F_n^P(\cdot)$ por sus representaciones i.i.d. $\overline{f}_n^P(\cdot)$ y $\overline{F}_n^P(\cdot)$:

$$\int_{\frac{\varepsilon}{2}}^{t_H} f_n^P(v) \frac{1 - F_n^P(v)}{1 - H(v)} \omega(v) dv = \int_{\frac{\varepsilon}{2}}^{t_H} \overline{f}_n^P(v) \frac{1 - \overline{F}_n^P(v)}{1 - H(v)} \omega(v) dv + o_P(1). \quad (4.52)$$

Desarrollamos el producto de los estimadores $f_n^P(\cdot)$ y $1 - F_n^P(\cdot)$:

$$\begin{aligned} f_n^P(t)(1 - F_n^P(t)) &= \overline{f_n^P}(t) \left(1 - \overline{F_n^P}(t)\right) + \left(1 - \overline{F_n^P}(t)\right) e_n(t) \\ &\quad - \overline{f_n^P}(t) R_n(t) - e_n(t) R_n(t). \end{aligned}$$

Entonces,

$$\begin{aligned} \int_{\frac{\varepsilon}{2}}^{t_H} f_n^P(v) \frac{1 - F_n^P(v)}{1 - H(v)} \omega(v) dv &= \int_{\frac{\varepsilon}{2}}^{t_H} \overline{f_n^P}(v) \frac{1 - \overline{F_n^P}(v)}{1 - H(v)} \omega(v) dv + \int_{\frac{\varepsilon}{2}}^{t_H} \frac{1 - \overline{F_n^P}(v)}{1 - H(v)} e_n(v) \omega(v) dv \\ &\quad - \int_{\frac{\varepsilon}{2}}^{t_H} \frac{\overline{f_n^P}(v)}{1 - H(v)} R_n(v) \omega(v) dv - \int_{\frac{\varepsilon}{2}}^{t_H} \frac{e_n(v) R_n(v)}{1 - H(v)} \omega(v) dv \\ &= \int_{\frac{\varepsilon}{2}}^{t_H} \overline{f_n^P}(v) \frac{1 - \overline{F_n^P}(v)}{1 - H(v)} \omega(v) dv + I_1 - I_2 - I_3. \end{aligned}$$

Veamos que cada uno de los términos I_1 , I_2 e I_3 es despreciable en probabilidad.

La integral I_1 es de orden $o_P(1)$, ya que lo es $e_n(\cdot)$ de forma casi segura, por el teorema 3.2.1, bajo la hipótesis (v9), el valor absoluto de la diferencia $F_n^P(\cdot) - \overline{F_n^P}(\cdot)$ por el teorema 2.4.3, y el valor absoluto de I_1 está acotado por

$$\begin{aligned} |I_1| &\leq \int \frac{1 - \overline{F_n^P}(v)}{1 - H(v)} |e_n(v)| \omega(v) dv \\ &\leq \frac{1}{1 - H(t_H)} \left(1 + \sup_{t \in [0, t_H]} |F_n^P(t) - \overline{F_n^P}(t)|\right) \sup_{t \in [0, t_H]} |e_n(t)| \int_0^{t_H} \omega(v) dv. \end{aligned}$$

La integral I_2 es también de orden $o_P(1)$ puesto que, por el teorema 2.4.2 y (v9), lo es $R_n(\cdot)$ de forma casi segura, y por el lema 4.3.4:

$$|I_2| \leq \int \frac{\overline{f_n^P}(v)}{1 - H(v)} |R_n(v)| \omega(v) dv \leq \frac{1}{1 - H(t_H)} \sup_{t \in [0, t_H]} |R_n(t)| \int_0^{t_H} \overline{f_n^P}(v) \omega(v) dv.$$

Análogamente, el término I_3 es de orden $o_P(1)$, dado que

$$|I_3| \leq \int \frac{|e_n(v) R_n(v)|}{1 - H(v)} \omega(v) dv \leq \frac{1}{1 - H(t_H)} \sup_{t \in [0, t_H]} |e_n(t)| \sup_{t \in [0, t_H]} |R_n(t)| \int_0^{t_H} \omega(v) dv.$$

Paso 3. Por último, veremos que

$$\int_{\frac{\varepsilon}{2}}^{t_H} \overline{f_n^P}(v) \frac{1 - \overline{F_n^P}(v)}{1 - H(v)} \omega(v) dv = \int_{\frac{\varepsilon}{2}}^{t_H} f(v) \frac{1 - F(v)}{1 - H(v)} \omega(v) dv + o_P(1). \quad (4.53)$$

Para probar esto, comprobaremos que

$$\lim_{n \rightarrow \infty} \mathbb{E} \left(\int_{\frac{\varepsilon}{2}}^{t_H} \frac{\overline{f_n^P}(v)}{f_n^P(v)} \frac{1 - \overline{F_n^P}(v)}{1 - H(v)} \omega(v) dv \right) = \int_{\frac{\varepsilon}{2}}^{t_H} f(v) \frac{1 - F(v)}{1 - H(v)} \omega(v) dv \quad (4.54)$$

y

$$\lim_{n \rightarrow \infty} Var \left(\int_{\frac{\varepsilon}{2}}^{t_H} \frac{\overline{f_n^P}(v)}{f_n^P(v)} \frac{1 - \overline{F_n^P}(v)}{1 - H(v)} \omega(v) dv \right) = 0. \quad (4.55)$$

En lo que respecta al primer límite, por el teorema de la convergencia dominada, es suficiente demostrar que la esperanza de $\frac{\overline{f_n^P}(\cdot)}{f_n^P(\cdot)} \left(1 - \frac{\overline{F_n^P}(\cdot)}{F(\cdot)}\right)$ es continua, está acotada en $t \in [\frac{\varepsilon}{2}, t_H]$ y que su límite cuando $n \rightarrow \infty$ es $f(\cdot)(1 - F(\cdot))$. Para ello descomponemos el producto de ambos estimadores de la siguiente manera:

$$\begin{aligned} \frac{\overline{f_n^P}(t)}{f_n^P(t)} \left(1 - \frac{\overline{F_n^P}(t)}{F(t)}\right) &= \left(\int K_s(t-v) f(v) dv + \sigma_n(t) \right) \left(1 - F(t) - \frac{1}{n} \sum_{i=1}^n \varepsilon_i(t)\right) \\ &= (1 - F(t)) \int K_s(t-v) f(v) dv + (1 - F(t)) \sigma_n(t) \\ &\quad - \int K_s(t-v) f(v) dv \frac{1}{n} \sum_{i=1}^n \varepsilon_i(t) - \sigma_n(t) \frac{1}{n} \sum_{i=1}^n \varepsilon_i(t). \end{aligned} \quad (4.56)$$

Bajo las hipótesis (K1) y (f1), el primer sumando en (4.56) es continuo y acotado, y

$$\lim_{n \rightarrow \infty} (1 - F(t)) \int K_s(t-v) f(v) dv = (1 - F(t)) f(t).$$

Probaremos que la esperanza de los demás términos en la descomposición (4.56) son funciones continuas, acotadas y que sus límites cuando $n \rightarrow \infty$ son cero.

Por la aplicación del lema 4.3.3, y bajo las hipótesis (K1), (H2), (p2) y (f1), la esperanza del segundo sumando en (4.56) es continua, acotada y su límite cuando $n \rightarrow \infty$ es cero, al igual que la esperanza del tercer sumando, por la aplicación del lema 4.3.2 con $n_0 = 1$. Por último, la esperanza del cuarto sumando en (4.56) es

$$\begin{aligned} \mathbb{E} \left[\sigma_n(t) \frac{1}{n} \sum_{i=1}^n \varepsilon_i(t) \right] &= \mathbb{E} [\sigma_n(t) \varepsilon_1(t)] \\ &= \frac{n-1}{ns} \mathbb{E} [\varepsilon_1(t)] \int \mathbb{E} [\varepsilon_2(t-sv)] K'(v) dv \\ &\quad + \frac{1}{ns} \int \mathbb{E} [\varepsilon_1(t) \varepsilon_1(t-sv)] K'(v) dv \\ &= \frac{n-1}{n} \mathbb{E} [\varepsilon_1(t)] \mathbb{E} [\sigma_n(t)] + \frac{1}{ns} \int \mathbb{E} [\varepsilon_1(t) \varepsilon_1(t-sv)] K'(v) dv, \end{aligned}$$

y, por los lemas 4.3.2 y 4.3.3, es continua, acotada y, como consecuencia de la hipótesis (v5), su límite es cero. Queda por tanto probado (4.54).

Pasamos ahora a demostrar (4.55):

$$\begin{aligned} & \text{Var} \left(\int_{\frac{\varepsilon}{2}}^{t_H} \overline{f_n^P}(v) \frac{1 - \overline{F_n^P}(v)}{1 - H(v)} \omega(v) dv \right) \\ &= \mathbb{E} \left[\left(\int_{\frac{\varepsilon}{2}}^{t_H} \overline{f_n^P}(v) \frac{1 - \overline{F_n^P}(v)}{1 - H(v)} \omega(v) dv \right)^2 \right] - \mathbb{E} \left(\int_{\frac{\varepsilon}{2}}^{t_H} \overline{f_n^P}(v) \frac{1 - \overline{F_n^P}(v)}{1 - H(v)} \omega(v) dv \right)^2 \\ &= \mathbb{E} \left[\left(\int_{\frac{\varepsilon}{2}}^{t_H} \overline{f_n^P}(v) \frac{1 - \overline{F_n^P}(v)}{1 - H(v)} \omega(v) dv \right)^2 \right] - \left(\int_{\frac{\varepsilon}{2}}^{t_H} f(v) \frac{1 - F(v)}{1 - H(v)} \omega(v) dv \right)^2 + o(1). \end{aligned}$$

Tenemos, por tanto, que probar que

$$\lim_{n \rightarrow \infty} \mathbb{E} \left[\left(\int_{\frac{\varepsilon}{2}}^{t_H} \overline{f_n^P}(v) \frac{1 - \overline{F_n^P}(v)}{1 - H(v)} \omega(v) dv \right)^2 \right] = \left(\int_{\frac{\varepsilon}{2}}^{t_H} f(v) \frac{1 - F(v)}{1 - H(v)} \omega(v) dv \right)^2,$$

donde

$$\begin{aligned} & \mathbb{E} \left[\left(\int_{\frac{\varepsilon}{2}}^{t_H} \overline{f_n^P}(v) \frac{1 - \overline{F_n^P}(v)}{1 - H(v)} \omega(v) dv \right)^2 \right] \tag{4.57} \\ &= \int_{\frac{\varepsilon}{2}}^{t_H} \int_{\frac{\varepsilon}{2}}^{t_H} \frac{\mathbb{E} \left[\overline{f_n^P}(v_1) \overline{f_n^P}(v_2) \left(1 - \overline{F_n^P}(v_1) \right) \left(1 - \overline{F_n^P}(v_2) \right) \right]}{(1 - H(v_1))(1 - H(v_2))} \omega(v_1) \omega(v_2) dv_1 dv_2. \end{aligned}$$

La esperanza que aparece en el integrando es igual a

$$\begin{aligned} & \mathbb{E} \left[\left(1 - F(t_1) - \frac{1}{n} \sum_{i=1}^n \varepsilon_i(t_1) \right) \left(1 - F(t_2) - \frac{1}{n} \sum_{j=1}^n \varepsilon_j(t_2) \right) \right. \\ & \quad \times \left(\int K_s(t_1 - u) f(u) du + \sigma_n(t_1) \right) \\ & \quad \times \left. \left(\int K_s(t_2 - v) f(v) dv + \sigma_n(t_2) \right) \right] \\ &= \mathbb{E} \left[\left((1 - F(t_1))(1 - F(t_2)) - (1 - F(t_2)) \frac{1}{n} \sum_{i=1}^n \varepsilon_i(t_1) - (1 - F(t_1)) \frac{1}{n} \sum_{j=1}^n \varepsilon_j(t_2) \right. \right. \\ & \quad + \frac{1}{n^2} \sum_{i,j=1}^n \varepsilon_i(t_1) \varepsilon_j(t_2) \times \left(\int K_s(t_1 - u) f(u) du \int K_s(t_2 - v) f(v) dv \right. \\ & \quad + \sigma_n(t_1) \int K_s(t_2 - v) f(v) dv \\ & \quad \left. \left. + \sigma_n(t_2) \int K_s(t_1 - u) f(u) du + \sigma_n(t_1) \sigma_n(t_2) \right) \right] \tag{4.58} \end{aligned}$$

Por simetría en las variables v_1 y v_2 de integración y teniendo en cuenta (4.58), la esperanza (4.57) es igual a la siguiente integral:

$$\int_{\frac{\varepsilon}{2}}^{t_H} \int_{\frac{\varepsilon}{2}}^{t_H} \frac{m(v_1, v_2)}{(1 - H(v_1))(1 - H(v_2))} \omega(v_1) \omega(v_2) dv_1 dv_2$$

donde

$$\begin{aligned}
& m(t_1, t_2) \\
&= (1 - F(t_1))(1 - F(t_2)) \int K_s(t_1 - u) f(u) du \int K_s(t_2 - v) f(v) dv \\
&\quad - 2(1 - F(t_1)) \int K_s(t_1 - u) f(u) du \int K_s(t_2 - v) f(v) dv \mathbb{E}[\varepsilon_1(t_2)] \\
&\quad + \int K_s(t_1 - u) f(u) du \int K_s(t_2 - v) f(v) dv \frac{1}{n^2} \sum_{i,j=1}^n \mathbb{E}[\varepsilon_i(t_1) \varepsilon_j(t_2)] \\
&\quad + 2(1 - F(t_1))(1 - F(t_2)) \left(\int K_s(t_1 - u) f(u) du \right) \mathbb{E}[\sigma_n(t_2)] \\
&\quad - 2(1 - F(t_1)) \left(\int K_s(t_1 - u) f(u) du \right) \mathbb{E}[\sigma_n(t_2) \varepsilon_1(t_2)] \\
&\quad - 2(1 - F(t_1)) \left(\int K_s(t_2 - u) f(u) du \right) \mathbb{E}[\sigma_n(t_1) \varepsilon_1(t_2)] \\
&\quad + 2 \left(\int K_s(t_1 - u) f(u) du \right) \frac{1}{n^2} \sum_{i,j=1}^n \mathbb{E}[\sigma_n(t_2) \varepsilon_i(t_1) \varepsilon_j(t_2)] \\
&\quad + (1 - F(t_1))(1 - F(t_2)) \mathbb{E}[\sigma_n(t_1) \sigma_n(t_2)] - 2(1 - F(t_1)) \mathbb{E}[\sigma_n(t_1) \sigma_n(t_2) \varepsilon_1(t_2)] \\
&\quad + \frac{1}{n^2} \sum_{i,j=1}^n \mathbb{E}[\sigma_n(t_1) \sigma_n(t_2) \varepsilon_i(t_1) \varepsilon_j(t_2)] \\
&= I_1(t_1, t_2) - 2I_2(t_1, t_2) + I_3(t_1, t_2) + 2I_4(t_1, t_2) \\
&\quad - 2I_5(t_1, t_2) - 2I_6(t_1, t_2) + 2I_7(t_1, t_2) + I_8(t_1, t_2) - 2I_9(t_1, t_2) + I_{10}(t_1, t_2).
\end{aligned}$$

Comprobaremos que todos estos sumandos son continuos, acotados y su límite, cuando $n \rightarrow \infty$, excepto para $I_1(\cdot, \cdot)$, es cero.

De acuerdo con los lemas 4.3.2 y 4.3.3, las cuatro primeras funciones $I_1(\cdot, \cdot)$, $I_2(\cdot, \cdot)$, $I_3(\cdot, \cdot)$ e $I_4(\cdot, \cdot)$ junto con $I_8(\cdot, \cdot)$ son continuas, acotadas, y además

$$\lim_{n \rightarrow \infty} I_1(t_1, t_2) = (1 - F(t_1))(1 - F(t_2)) f(t_1) f(t_2)$$

y

$$\lim_{n \rightarrow \infty} I_2(t_1, t_2) = \lim_{n \rightarrow \infty} I_3(t_1, t_2) = \lim_{n \rightarrow \infty} I_4(t_1, t_2) = \lim_{n \rightarrow \infty} I_8(t_1, t_2) = 0.$$

La esperanza $\mathbb{E}[\sigma_n(\cdot) \varepsilon_1(\cdot)]$ en las funciones $I_5(\cdot, \cdot)$ e $I_6(\cdot, \cdot)$ coincide con la esperanza del cuarto sumando en (4.56), para el que se demostró que es continua, acotada y su límite es cero, de modo que

$$\lim_{n \rightarrow \infty} I_5(t_1, t_2) = \lim_{n \rightarrow \infty} I_6(t_1, t_2) = 0.$$

El término $I_7(\cdot, \cdot)$ es igual a

$$I_7(t_1, t_2) = 2 \int K_s(t_1 - v) f(v) dv \times \frac{1}{n^2} \sum_{i,j=1}^n \mathbb{E}[\sigma_n(t_2) \varepsilon_i(t_1) \varepsilon_j(t_2)],$$

cuyo sumatorio se puede separar, a su vez, como sigue

$$\frac{n-1}{n} \mathbb{E}[\sigma_n(t_2) \varepsilon_1(t_1) \varepsilon_2(t_2)] + \frac{1}{n} \mathbb{E}[\sigma_n(t_2) \varepsilon_1(t_1) \varepsilon_1(t_2)]. \quad (4.59)$$

Veamos que ambas esperanzas son acotadas y que su límite es cero.

La primera esperanza en (4.59) es, por la independencia entre $\varepsilon_i(\cdot)$ y $\varepsilon_j(\cdot)$ con $i \neq j$, igual a

$$\begin{aligned} \mathbb{E}[\sigma_n(t_2) \varepsilon_1(t_1) \varepsilon_2(t_2)] &= \frac{1}{ns} \sum_{i=1}^n \int \mathbb{E}[\varepsilon_1(t_1) \varepsilon_2(t_2) \varepsilon_i(t_2 - vs)] K'(v) dv \\ &= \frac{1}{ns} \mathbb{E}[\varepsilon_2(t_2)] \int \mathbb{E}[\varepsilon_1(t_1) \varepsilon_1(t_2 - vs)] K'(v) dv \\ &\quad + \frac{1}{ns} \mathbb{E}[\varepsilon_1(t_1)] \int \mathbb{E}[\varepsilon_2(t_2) \varepsilon_2(t_2 - vs)] K'(v) dv \\ &\quad + \frac{n-2}{ns} \mathbb{E}[\varepsilon_1(t_1)] \mathbb{E}[\varepsilon_2(t_2)] \int \mathbb{E}[\varepsilon_3(t_2 - vs)] K'(v) dv. \end{aligned}$$

En virtud de los lemas 4.3.2. y 4.3.3, esta esperanza es continua, acotada y su límite es cero.

Mediante un razonamiento análogo, la segunda esperanza en (4.59) es,

$$\begin{aligned} \frac{1}{n} \mathbb{E}[\sigma_n(t_2) \varepsilon_1(t_1) \varepsilon_1(t_2)] &= \frac{1}{n} \frac{n-1}{ns} \mathbb{E}[\varepsilon_1(t_1) \varepsilon_1(t_2)] \int \mathbb{E}[\varepsilon_2(t_2 - vs)] K'(v) dv \\ &\quad + \frac{1}{n} \frac{1}{ns} \int \mathbb{E}[\varepsilon_1(t_1) \varepsilon_1(t_2) \varepsilon_1(t_2 - vs)] K'(v) dv \\ &= \frac{n-1}{n^2} \mathbb{E}[\varepsilon_1(t_1) \varepsilon_1(t_2)] \mathbb{E}[\sigma_n(t_2)] + \frac{1}{n^2 s} \int \mathbb{E}[\varepsilon_1(t_1) \varepsilon_1(t_2) \varepsilon_1(t_2 - vs)] K'(v) dv, \end{aligned}$$

es decir, acotada y, gracias a (v5) y a la aplicación de los lemas 4.3.2 y 4.3.3, con límite cero.

Para la función $I_9(\cdot, \cdot)$ desarrollamos la siguiente esperanza:

$$\begin{aligned} &\mathbb{E}[\varepsilon_1(t_2) \sigma_n(t_1) \sigma_n(t_2)] \\ &= \mathbb{E} \left[\frac{1}{n^2 s^2} \sum_{i,j=1}^n \iint \varepsilon_1(t_2) \varepsilon_i(t_1 - sv_1) \varepsilon_j(t_2 - sv_2) K'(v_1) K'(v_2) dv_1 dv_2 \right] \end{aligned}$$

$$\begin{aligned}
&= \frac{(n-1)(n-2)}{n^2 s^2} \mathbb{E}[\varepsilon_1(t_2)] \iint \mathbb{E}[\varepsilon_2(t_1 - sv_1)] \mathbb{E}[\varepsilon_3(t_2 - sv_2)] K'(v_1) K'(v_2) dv_1 dv_2 \\
&\quad + \frac{n-1}{n^2 s^2} \mathbb{E}[\varepsilon_1(t_2)] \iint \mathbb{E}[\varepsilon_2(t_1 - sv_1) \varepsilon_2(t_2 - sv_2)] K'(v_1) K'(v_2) dv_1 dv_2 \\
&\quad + \frac{n-1}{n^2 s^2} \iint \mathbb{E}[\varepsilon_1(t_2) \varepsilon_1(t_1 - sv_1)] \mathbb{E}[\varepsilon_2(t_2 - sv_2)] K'(v_1) K'(v_2) dv_1 dv_2 \\
&\quad + \frac{n-1}{n^2 s^2} \iint \mathbb{E}[\varepsilon_1(t_2) \varepsilon_1(t_2 - sv_2)] \mathbb{E}[\varepsilon_2(t_1 - sv_1)] K'(v_1) K'(v_2) dv_1 dv_2 \\
&\quad + \frac{1}{n^2 s^2} \iint \mathbb{E}[\varepsilon_1(t_2) \varepsilon_1(t_1 - sv_1) \varepsilon_1(t_2 - sv_2)] K'(v_1) K'(v_2) dv_1 dv_2 \\
&= \frac{(n-1)(n-2)}{n^2} \mathbb{E}[\varepsilon_1(t_2)] \mathbb{E}[\sigma_n(t_1)] \mathbb{E}[\sigma_n(t_2)] \\
&\quad + \frac{n-1}{n^2 s^2} \mathbb{E}[\varepsilon_1(t_2)] \iint \mathbb{E}[\varepsilon_2(t_1 - sv_1) \varepsilon_2(t_1 - sv_2)] K'(v_1) K'(v_2) dv_1 dv_2 \\
&\quad + \frac{n-1}{n^2 s} \mathbb{E}[\sigma_n(t_2)] \int \mathbb{E}[\varepsilon_1(t_2) \varepsilon_1(t_1 - sv_1)] K'(v_1) dv_1 \\
&\quad + \frac{n-1}{n^2 s} \mathbb{E}[\sigma_n(t_1)] \int \mathbb{E}[\varepsilon_1(t_2) \varepsilon_1(t_2 - sv_2)] K'(v_2) dv_2 \\
&\quad + \frac{1}{n^2 s^2} \iint \mathbb{E}[\varepsilon_1(t_2) \varepsilon_1(t_1 - sv_1) \varepsilon_1(t_2 - sv_2)] K'(v_1) K'(v_2) dv_1 dv_2.
\end{aligned}$$

A partir de los resultados de los lemas 4.3.2 y 4.3.3 y como consecuencia de la hipótesis (v5), es inmediato concluir que

$$\lim_{n \rightarrow \infty} I_9(t_1, t_2) = 0.$$

Por último, para calcular el límite de la función $I_{10}(\cdot, \cdot)$ operamos de la siguiente manera:

$$\begin{aligned}
&\frac{1}{n^2} \sum_{i,j=1}^n \mathbb{E}[\sigma_n(t_1) \sigma_n(t_2) \varepsilon_i(t_1) \varepsilon_j(t_2)] \\
&= \frac{1}{n^4 s^2} \sum_{i,j=1}^n \sum_{k,l=1}^n \mathbb{E} \left[\iint \varepsilon_i(t_1) \varepsilon_j(t_2) \varepsilon_k(t_1 - sv_1) \varepsilon_l(t_2 - sv_2) K'(v_1) K'(v_2) dv_1 dv_2 \right]
\end{aligned}$$

Mediante razonamientos análogos a los usados para $I_9(\cdot, \cdot)$ se concluye que, bajo (v5),

$$\lim_{n \rightarrow \infty} I_{10}(t_1, t_2) = 0.$$

Recopilando (4.50), (4.52) y (4.53) se obtiene (4.49). ■

Los lemas 4.3.8 y 4.3.9 implican que

$$\widehat{c}_2(L) \xrightarrow{P} c_2(L) \quad \forall L > 0,$$

siendo $c_2(\cdot)$ y $\widehat{c}_2(\cdot)$ las funciones (4.21) y (4.32) respectivamente. Para demostrar la consistencia de las ventanas plug-in necesitamos ahora probar que

$$\widehat{c}_1(L) \xrightarrow{P} c_1(L) \quad \forall L > 0,$$

siendo $c_1(\cdot)$ la función dada en (4.20), y $\widehat{c}_1(\cdot)$ la función dada en (4.31). Si definimos las siguientes funciones:

$$\widehat{c}_{11}(t) =: \frac{1}{2} d_K f_n^{P''}(t) \quad \text{y} \quad \widehat{c}_{12}(t, L) =: L^2 d_K [(1 - F_n^P(t)) \alpha'_n(t) - f_n^P(t) \alpha_n(t)],$$

entonces la función $\widehat{c}_1(\cdot)$ se puede expresar como

$$\widehat{c}_1(L) = \int (\widehat{c}_{11}(v) + \widehat{c}_{12}(v, L))^2 \omega(v) dv,$$

mientras que

$$c_1(L) = \int (c_{11}(v) + c_{12}(v, L))^2 \omega(v) dv,$$

donde $c_{11}(\cdot)$ y $c_{12}(\cdot, \cdot)$ vendrían dadas, respectivamente, por

$$c_{11}(t) = \frac{1}{2} d_K f''(t) \quad \text{y} \quad c_{12}(t, L) = L^2 d_K [(1 - F(t)) \alpha'(t) - f(t) \alpha(t)].$$

Observación 4.3.3

Para demostrar la convergencia en probabilidad de $\widehat{c}_1(\cdot)$ a $c_1(\cdot)$ basta probar

$$\int (\widehat{c}_{11}(v) - c_{11}(v))^2 \omega(v) dv \xrightarrow{P} 0 \quad \text{y} \quad \int (\widehat{c}_{12}(v, L) - c_{12}(v, L))^2 \omega(v) dv \xrightarrow{P} 0 \quad (4.60)$$

puesto que, por la desigualdad triangular, tenemos que

$$\begin{aligned} \int (\widehat{c}_{11}(v) - c_{11}(v))^2 \omega(v) dv &\xrightarrow{P} 0 \Rightarrow \int \widehat{c}_{11}^2(v) \omega(v) dv \xrightarrow{P} \int c_{11}^2(v) \omega(v) dv \\ \int (\widehat{c}_{12}(v, L) - c_{12}(v, L))^2 \omega(v) dv &\xrightarrow{P} 0 \\ &\Rightarrow \int \widehat{c}_{12}^2(v, L) \omega(v) dv \xrightarrow{P} \int c_{12}^2(v, L) \omega(v) dv. \end{aligned}$$

Falta probar entonces que

$$\int \widehat{c}_{11}(v) \widehat{c}_{12}(v, L) \omega(v) dv \xrightarrow{P} \int c_{11}(v) c_{12}(v, L) \omega(v) dv,$$

donde

$$\begin{aligned}
\int \widehat{c}_{11}(v) \widehat{c}_{12}(v, L) \omega(v) dv &= \int (\widehat{c}_{11}(v) - c_{11}(v)) (\widehat{c}_{12}(v, L) - c_{12}(v, L)) \omega(v) dv \\
&+ \int c_{11}(v) (\widehat{c}_{12}(v, L) - c_{12}(v, L)) \omega(v) dv \\
&+ \int (\widehat{c}_{11}(v) - c_{11}(v)) c_{12}(v, L) \omega(v) dv + \int c_{11}(v) c_{12}(v, L) \omega(v) dv.
\end{aligned} \tag{4.61}$$

Para ello, probaremos que las tres primeras integrales en (4.61) convergen en probabilidad a cero. Pero, por la desigualdad de Cauchy-Schwarz, estas tres integrales pueden acotarse fácilmente por

$$\begin{aligned}
&\left| \int (\widehat{c}_{11}(v) - c_{11}(v)) (\widehat{c}_{12}(v, L) - c_{12}(v, L)) \omega(v) dv \right| \\
&\leq \left[\int (\widehat{c}_{11}(v) - c_{11}(v))^2 \omega(v) dv \right]^{1/2} \left[\int (\widehat{c}_{12}(v, L) - c_{12}(v, L))^2 \omega(v) dv \right]^{1/2}, \\
&\left| \int c_{11}(v) (\widehat{c}_{12}(v, L) - c_{12}(v, L)) \omega(v) dv \right| \\
&\leq \left[\int c_{11}^2(v) \omega(v) dv \right]^{1/2} \left[\int (\widehat{c}_{12}(v, L) - c_{12}(v, L))^2 \omega(v) dv \right]^{1/2}, \\
&\left| \int (\widehat{c}_{11}(v) - c_{11}(v)) c_{12}(v, L) \omega(v) dv \right| \\
&\leq \left[\int (\widehat{c}_{11}(v) - c_{11}(v))^2 \omega(v) dv \right]^{1/2} \left[\int c_{12}^2(v, L) \omega(v) dv \right]^{1/2},
\end{aligned}$$

de modo que, si se verifica (4.60) entonces

$$\int \widehat{c}_{11}(v) \widehat{c}_{12}(v, L) \omega(v) dv \xrightarrow{P} \int c_{11}(v) c_{12}(v, L) \omega(v) dv.$$

A continuación demostraremos el primer resultado en (4.60).

Lema 4.3.10

Bajo las condiciones (K1), (H3), (p2), (p3), (f1), (v7) y (v8) se tiene

$$\int (f_n^{P''}(v) - f''(v))^2 \omega(v) dv \xrightarrow{P} 0. \tag{4.62}$$

Demostración.

Paso 1. El primer paso consiste en sustituir en la integral la segunda derivada del estimador presuavizado $f_n^P(\cdot)$ por la derivada segunda de su representación i.i.d. $\overline{f}_n^P(\cdot)$, y probar que la nueva integral es equivalente en probabilidad a la primera:

$$\begin{aligned}
&\int (f_n^{P''}(v) - f''(v))^2 \omega(v) dv \\
&= \int (\overline{f}_n^{P''}(v) - f''(v))^2 \omega(v) dv + o_P \left(\int (\overline{f}_n^{P''}(v) - f''(v))^2 \omega(v) dv \right).
\end{aligned} \tag{4.63}$$

El estimador presuavizado es igual a $f_n^P(\cdot) = \overline{f_n^P}(\cdot) + e_n(\cdot)$, de modo que

$$(f_n^{P''}(t) - f''(t))^2 = \left(\overline{f_n^{P''}}(t) - f''(t)\right)^2 + e_n''(t)^2 + 2\left(\overline{f_n^{P''}}(t) - f''(t)\right)e_n''(t).$$

Vamos a comprobar que las integrales de los dos últimos sumandos son de orden $o_P(1)$. El término del error $e_n(\cdot)$ del estimador presuavizado $f_n^P(\cdot)$ viene dado por

$$e_n(t) = \int K_s(t-v) dR_n(v),$$

donde $R_n(\cdot)$ es el resto de la representación asintótica del estimador presuavizado de la función de distribución. Para calcular la segunda derivada de $e_n(\cdot)$, en el paso (a) efectuamos una integración por partes, y en el paso (b) un cambio de variable:

$$\begin{aligned} e_n''(t) &= \frac{1}{s^3} \int K''\left(\frac{t-v}{s}\right) dR_n(v) \stackrel{(a)}{=} \frac{1}{s^4} \int K''' \left(\frac{t-v}{s}\right) R_n(v) dv \\ &\stackrel{(b)}{=} \frac{1}{s^3} \int K'''(z) R_n(t-sz) dz. \end{aligned}$$

Para calcular la integral de la esperanza de $e_n''(\cdot)^2$, en el paso (a) aplicamos el teorema de Fubini, en el paso (b) usamos la desigualdad de Cauchy-Schwarz, y en el paso (c) extraemos el supremo:

$$\begin{aligned} \int \mathbb{E} \left[e_n''(v)^2 \right] \omega(v) dv &= \frac{1}{s^6} \mathbb{E} \left[\int \left(\int K'''(z) R_n(v-sz) dz \right)^2 \omega(v) dv \right] \\ &\stackrel{(a)}{=} \frac{1}{s^6} \iiint \mathbb{E} [R_n(v-sz_1) R_n(v-sz_2)] |K'''(z_1)| |K'''(z_2)| \omega(v) dz_1 dz_2 dv \\ &\stackrel{(b)}{\leq} \frac{1}{s^6} \iint \mathbb{E} [R_n^2(v-sz_1)]^{1/2} |K'''(z_1)| dz_1 \int \mathbb{E} [R_n^2(v-sz_2)]^{1/2} |K'''(z_2)| \omega(v) dz_2 dv \\ &\stackrel{(c)}{\leq} \frac{1}{s^6} \sup_{t \in [0, t_H]} \mathbb{E} [R_n^2(t)] \left(\int |K'''(z)| dz \right)^2 \int \omega(v) dv = O \left(\left(\frac{b^4}{s^3} + \frac{1}{nbs^3} \right)^2 \right), \end{aligned}$$

donde el orden de la acotación resulta del teorema 2.4.2 del capítulo 2. Por tanto, por la desigualdad de Markov y bajo la hipótesis (v8), la integral $\int e_n''(v)^2 \omega(v) dv$ es de orden $o_P(1)$.

Con respecto a la integral $\int \left(\overline{f_n^P}(v)'' - f''(v)\right) e_n''(v) \omega(v) dv$, aplicamos ahora la desigualdad de Cauchy-Schwarz con respecto a la medida en $[0, t_H]$ que tiene como derivada respecto de la medida de Lebesgue la función $\omega(\cdot)$, de modo que

$$\begin{aligned} \int \left(\overline{f_n^P}(v)'' - f''(v)\right) e_n''(v) \omega(v) dv &\leq \left[\int \left(\overline{f_n^P}(v)'' - f''(v)\right)^2 \omega(v) dv \right]^{1/2} \\ &\quad \times \left(\int e_n''(v)^2 \omega(v) dv \right)^{1/2}. \end{aligned}$$

Esta integral es, por tanto, de orden o_P $\left(\int \left(\overline{f_n^P}''(v) - f''(v) \right)^2 \omega(v) dv \right)$ puesto que, como se comprobó en el paso anterior, $\int e_n''(v)^2 \omega(v) dv = o_P(1)$.

Paso 2. Probamos ahora que

$$\int \left(\overline{f_n^P}''(v) - f''(v) \right)^2 \omega(v) dv = o_P(1). \quad (4.64)$$

Por el lema 4.3.1, para probar esta convergencia en probabilidad a cero basta demostrar que las esperanzas $\mathbb{E} \left[\overline{f_n^P}''(\cdot) \right]$ y $\mathbb{E} \left[\overline{f_n^P}''(\cdot)^2 \right]$ son continuas, su valor absoluto está acotado, y

$$\lim_{n \rightarrow \infty} \mathbb{E} \left[\overline{f_n^P}''(t) \right] = f''(t) \quad \text{y} \quad \lim_{n \rightarrow \infty} \mathbb{E} \left[\overline{f_n^P}''(t)^2 \right] = f''(t)^2.$$

La esperanza de la segunda derivada del estimador presuavizado es

$$\mathbb{E} \left[\overline{f_n^P}''(t) \right] = \int K_s''(t-v) f(v) dv + \mathbb{E} \left[\sigma_n''(t) \right].$$

A continuación comprobaremos que ambos términos son continuos y acotados, y que el límite cuando $n \rightarrow \infty$ del primer sumando es $f''(\cdot)$ y del segundo es cero.

En el primer sumando aplicamos un cambio de variable (paso (a)) y sucesivas integraciones por partes (paso (b)) para un n suficientemente grande:

$$\begin{aligned} \int K_s''(t-v) f(v) dv &\stackrel{(a)}{=} \frac{1}{s^2} \int_{-1}^{t/s} K''(u) f(t-us) du \\ &\stackrel{(b)}{=} \int_{-1}^{t/s} K(u) f''(t-us) du \end{aligned}$$

de modo que, bajo la hipótesis (f1), entonces es una función continua, acotada y

$$\lim_{n \rightarrow \infty} \int K_s''(t-v) f(v) dv = f''(t).$$

Por las hipótesis (K1), (H3) y (p2), la esperanza de $\sigma_n''(\cdot)$, que es igual a

$$\mathbb{E} \left[\sigma_n''(t) \right] = \frac{1}{ns^3} \sum_{i=1}^n \int \mathbb{E} [\varepsilon_i(t-vs)] K'''(v) dv = \frac{1}{s^3} \int \mathbb{E} [\varepsilon_1(t-vs)] K'''(v) dv,$$

es una función continua. Además, se puede acotar por

$$|\mathbb{E} [\sigma_n''(t)]| \leq \mathbb{E} \left[\sigma_n''(t)^2 \right]^{1/2},$$

de modo que, por el lema 4.3.7, es acotada y además

$$\lim_{n \rightarrow \infty} \mathbb{E} \left[\sigma_n''(t)^2 \right] = 0.$$

Todo esto permite concluir que la esperanza $\mathbb{E} \left[\overline{f_n^P}''(\cdot) \right]$ es continua, acotada y su límite cuando $n \rightarrow \infty$ es $f''(\cdot)$.

La esperanza $\mathbb{E} \left[\overline{f_n^P}''(\cdot)^2 \right]$ se puede descomponer en los siguientes sumandos:

$$\begin{aligned} \mathbb{E} \left(\overline{f_n^P}''(t)^2 \right) &= \left(\int K_s''(t-v) f(v) dv \right)^2 \\ &+ 2 \left(\int K_s''(t-v) f(v) dv \right) \mathbb{E} [\sigma_n''(t)] + \mathbb{E} [\sigma_n''(t)^2], \end{aligned}$$

donde $\int K_s''(t-v) f(v) dv$ es continua y acotada por la hipótesis (f1). Además,

$$\lim_{n \rightarrow \infty} \int \left(\int K_s''(v-w) f(w) dw \right)^2 \omega(v) dv = \int f''(v)^2 \omega(v) dv.$$

De este resultado y del lema 4.3.7 se obtiene (4.64), que junto con (4.63), permite concluir (4.62). ■

Pasamos ahora a demostrar la convergencia a cero de la integral

$$\int (\widehat{c}_{12}(v, L) - c_{12}(v, L))^2 \omega(v) dv,$$

es decir,

$$\begin{aligned} \int [((1 - F_n^P(v)) \alpha_n'(v) - f_n^P(v) \alpha_n(v)) \\ - ((1 - F(v)) \alpha'(v) - f(v) \alpha(v))]^2 \omega(v) dv \xrightarrow{P} 0. \end{aligned}$$

Observación 4.3.4

Si definimos los siguientes estimadores

$$a_n(t) =: (1 - F_n^P(t)) \alpha_n'(t) \quad y \quad b_n(t) =: f_n^P(t) \alpha_n(t)$$

de $a(t) =: (1 - F(t)) \alpha'(t)$ y $b(t) =: f(t) \alpha(t)$ respectivamente, entonces lo que queremos demostrar es

$$\int [(a_n(v) - b_n(v)) - (a(v) - b(v))]^2 \omega(v) dv \xrightarrow{P} 0.$$

Desarrollando este cuadrado tenemos que

$$\begin{aligned} \int [(a_n(v) - b_n(v)) - (a(v) - b(v))]^2 \omega(v) dv \\ = \int (a_n(v) - a(v))^2 \omega(v) dv + \int (b_n(v) - b(v))^2 \omega(v) dv \\ - 2 \int (a_n(v) - a(v))(b_n(v) - b(v)) \omega(v) dv, \end{aligned}$$

de modo que, en virtud de la desigualdad de Cauchy-Schwarz, basta con demostrar

$$\int (a_n(v) - a(v))^2 \omega(v) dv = o_P(1) \quad y \quad \int (b_n(v) - b(v))^2 \omega(v) dv = o_P(1),$$

es decir,

$$\int [(1 - F_n^P(v)) \alpha'_n(v) - (1 - F(v)) \alpha'(v)]^2 \omega(v) dv \xrightarrow{P} 0$$

y

$$\int (f_n^P(v) \alpha_n(v) - f(v) \alpha(v))^2 \omega(v) dv \xrightarrow{P} 0.$$

Lema 4.3.11

Bajo las condiciones (K1), (H3), (H4), (p2), (p3), (f1), (v5), (v7) y (v8) se tiene

$$\int_{\frac{\varepsilon}{2}}^{t_H} [(1 - F_n^P(v)) \alpha'_n(v) - (1 - F(v)) \alpha'(v)]^2 \omega(v) dv \xrightarrow{P} 0. \quad (4.65)$$

Demostración.

Paso 1. Empezamos eliminando el denominador aleatorio del estimador $\alpha'_n(\cdot)$:

$$\begin{aligned} & \int [(1 - F_n^P(v)) \alpha'_n(v) - (1 - F(v)) \alpha'(v)]^2 \omega(v) dv \\ = & \int \left[\frac{1}{2} (1 - F_n^P(v)) \frac{\psi_n''(v) - p_n(v) h_n''(v)}{1 - H(v)} - (1 - F(v)) \alpha'(v) \right]^2 \omega(v) dv \\ & + o_P \left(\int \left[\frac{1}{2} (1 - F_n^P(v)) \frac{\psi_n''(v) - p_n(v) h_n''(v)}{1 - H(v)} - (1 - F(v)) \alpha'(v) \right]^2 \omega(v) dv \right). \end{aligned}$$

Las funciones $\alpha'_n(\cdot)$ y $\alpha'(\cdot)$ vienen dadas por

$$\alpha'_n(t) = \frac{\frac{1}{2} p_n''(t) h_n(t) + p_n'(t) h_n'(t)}{1 - H_n(t) + \frac{1}{n}} \quad y \quad \alpha'(t) = \frac{\frac{1}{2} p''(t) h(t) + p'(t) h'(t)}{1 - H(t)}$$

respectivamente, y además $\psi_n(t) = p_n(t) h_n(t)$. Teniendo en cuenta las relaciones

$$\begin{aligned} p_n'(t) &= \frac{\psi_n'(t) h_n(t) - \psi_n(t) h_n'(t)}{h_n^2(t)}, \\ p_n''(t) &= \frac{\psi_n''(t) h_n^2(t) - \psi_n(t) h_n''(t) h_n(t) - 2\psi_n'(t) h_n'(t) h_n(t) + 2\psi_n(t) h_n'(t)^2}{h_n^3(t)}, \end{aligned}$$

entonces las funciones $\alpha'_n(\cdot)$ y $\alpha'(\cdot)$ se pueden escribir de la siguiente manera:

$$\alpha'_n(t) = \frac{1}{2} \frac{\psi_n''(t) - p_n(t) h_n''(t)}{1 - H_n(t) + \frac{1}{n}} \quad y \quad \alpha'(t) = \frac{1}{2} \frac{\psi''(t) - p(t) h''(t)}{1 - H(t)}. \quad (4.66)$$

Este primer paso consiste en sustituir el denominador aleatorio $1 - H_n(\cdot) + 1/n$ de $\alpha'_n(\cdot)$ por la función no aleatoria $1 - H(\cdot)$. Para ello, sustituimos $\alpha'_n(\cdot)$ y $\alpha'(\cdot)$ por

sus expresiones dadas en (4.66) (paso (a)), sumamos y restamos un término adecuado (paso (b)) y desarrollamos el cuadrado (paso (c)):

$$\begin{aligned}
& \int [(1 - F_n^P(v)) \alpha'_n(v) - (1 - F(v)) \alpha'(v)]^2 \omega(v) dv \\
\stackrel{(a)}{=} & \frac{1}{4} \int \left[(1 - F_n^P(v)) \frac{\psi''_n(v) - p_n(v) h''_n(v)}{1 - H_n(v) + \frac{1}{n}} \right. \\
& \left. - (1 - F(v)) \frac{\psi''(v) - p(v) h''(v)}{1 - H(v)} \right]^2 \omega(v) dv \\
\stackrel{(b)}{=} & \frac{1}{4} \int \left[(1 - F_n^P(v)) \frac{\psi''_n(v) - p_n(v) h''_n(v)}{1 - H_n(v) + \frac{1}{n}} \left(\frac{1 - H_n(v) + \frac{1}{n}}{1 - H(v)} + 1 - \frac{1 - H_n(v) + \frac{1}{n}}{1 - H(v)} \right) \right. \\
& \left. - (1 - F(v)) \frac{\psi''(v) - p(v) h''(v)}{1 - H(v)} \right]^2 \omega(v) dv \\
\stackrel{(c)}{=} & \frac{1}{4} \int \left[(1 - F_n^P(v)) \frac{\psi''_n(v) - p_n(v) h''_n(v)}{1 - H(v)} \right. \\
& \left. - (1 - F(v)) \frac{\psi''(v) - p(v) h''(v)}{1 - H(v)} \right]^2 \omega(v) dv \\
& + \frac{1}{4} \int (1 - F_n^P(v))^2 \left(\frac{\psi''_n(v) - p_n(v) h''_n(v)}{1 - H(v)} \right)^2 \left(\frac{H_n(v) - H(v) - \frac{1}{n}}{1 - H_n(v) + \frac{1}{n}} \right)^2 \omega(v) dv \\
& + \frac{1}{2} \int \left[(1 - F_n^P(v)) \frac{\psi''_n(v) - p_n(v) h''_n(v)}{1 - H(v)} - (1 - F(v)) \frac{\psi''(v) - p(v) h''(v)}{1 - H(v)} \right] \\
& \times (1 - F_n^P(v)) \frac{\psi''_n(v) - p_n(v) h''_n(v)}{1 - H(v)} \left(\frac{H_n(v) - H(v) - \frac{1}{n}}{1 - H_n(v) + \frac{1}{n}} \right) \omega(v) dv \\
= & I_1 + I_2 + I_3.
\end{aligned} \tag{4.67}$$

Por la desigualdad de Cauchy-Schwarz, la última integral en (4.67) I_3 es despreciable si lo son las dos primeras. La despreciableidad de la segunda integral, I_2 , se demuestra a continuación, y la de la primera, I_1 , se probará más adelante (pasos 2-3).

Gracias a la desigualdad (4.51) tenemos

$$\left| \frac{H_n(v) - H(v) - \frac{1}{n}}{1 - H_n(v) + \frac{1}{n}} \right| \leq \frac{\sup_{t \in [0, t_H]} |H_n(t) - H(t)| + \frac{1}{n}}{1 - H(t_H) + \frac{1}{n} - \sup_{t \in [0, t_H]} |H_n(t) - H(t)|}$$

y por tanto,

$$\sup_{0 \leq t \leq t_H} \left| \frac{H_n(t) - H(t) - \frac{1}{n}}{1 - H_n(t) + \frac{1}{n}} \right| = O_P(n^{-1/2}) = o_P(1). \tag{4.68}$$

Esto permite acotar la segunda integral en (4.67) de la siguiente manera:

$$\begin{aligned} I_2 &= \frac{1}{4} \int_{\frac{\varepsilon}{2}}^{t_H} (1 - F_n^P(v))^2 \left(\frac{\psi_n''(v) - p_n(v) h_n''(v)}{1 - H(v)} \right)^2 \left(\frac{H_n(v) - H(v) - \frac{1}{n}}{1 - H_n(v) + \frac{1}{n}} \right)^2 \omega(v) dv \\ &\leq \frac{1}{4} \left(\sup_{0 \leq t \leq t_H} \left| \frac{H_n(t) - H(t) - \frac{1}{n}}{1 - H_n(t) + \frac{1}{n}} \right| \right)^2 \int_{\frac{\varepsilon}{2}}^{t_H} \left(\frac{\psi_n''(v) - p_n(v) h_n''(v)}{1 - H(v)} \right)^2 \omega(v) dv, \end{aligned}$$

de modo que es de orden

$$O_P \left(\frac{1}{n} \int_{\frac{\varepsilon}{2}}^{t_H} \left(\frac{\psi_n''(v) - p_n(v) h_n''(v)}{1 - H(v)} \right)^2 \omega(v) dv \right).$$

Mediante cálculos análogos a los usados en los lemas 2.2.24 y 2.2.25 de López de Ullibarri (2004), esta integral verifica

$$\int_{\frac{\varepsilon}{2}}^{t_H} \left(\frac{\psi_n''(v) - p_n(v) h_n''(v)}{1 - H(v)} \right)^2 \omega(v) dv = O_P(1),$$

y por tanto $I_2 = o_P(1)$.

La despreciabilidad de I_1 se hará en dos etapas (pasos 2 y 3); en la primera de las cuales se aproxima la expresión $\psi_n''(\cdot) - p_n(\cdot) h_n''(\cdot)$ que aparece en el integrando por su análoga poblacional $\psi''(\cdot) - p(\cdot) h''(\cdot)$, y en la segunda se aproxima el estimador de la función de supervivencia $1 - F_n^P(\cdot)$ por la función teórica $1 - F(\cdot)$.

Paso 2. En este paso se demostrará que

$$\begin{aligned} &\frac{1}{4} \int \left[(1 - F_n^P(v)) \frac{\psi_n''(v) - p_n(v) h_n''(v)}{1 - H(v)} - (1 - F(v)) \frac{\psi''(v) - p(v) h''(v)}{1 - H(v)} \right]^2 \omega(v) dv \\ &= \frac{1}{4} \int \left([(1 - F_n^P(v)) - (1 - F(v))] \frac{\psi''(v) - p(v) h''(v)}{1 - H(v)} \right)^2 \omega(v) dv \\ &+ o_P \left(\int \left([(1 - F_n^P(v)) - (1 - F(v))] \frac{\psi''(v) - p(v) h''(v)}{1 - H(v)} \right)^2 \omega(v) dv \right). \end{aligned}$$

Puesto que el estimador $p_n(\cdot)$ tiene un denominador aleatorio, a continuación se descompone la expresión $\psi_n''(\cdot) - p_n(\cdot) h_n''(\cdot)$ de la siguiente manera:

$$\begin{aligned} \psi_n''(t) - p_n(t) h_n''(t) &= \psi''(t) - p(t) h''(t) + (\psi_n''(t) - \psi''(t)) - p(t) (h_n''(t) - h''(t)) \\ &\quad - (\psi_n(t) - \psi(t)) h''(t) h^{-1}(t) + p(t) (h_n(t) - h(t)) h''(t) h^{-1}(t) \\ &\quad + (p_n(t) - p(t)) [(h_n(t) - h(t)) h''(t) h^{-1}(t) - (h_n''(t) - h''(t))]. \end{aligned}$$

Así pues, podemos escribir lo siguiente:

$$\begin{aligned}
& \frac{1}{4} \int \left[(1 - F_n^P(v)) \frac{\psi_n''(v) - p_n(v) h_n''(v)}{1 - H(v)} - (1 - F(v)) \frac{\psi''(v) - p(v) h''(v)}{1 - H(v)} \right]^2 \omega(v) dv \\
&= \frac{1}{4} \int \left[(1 - F_n^P(v)) \frac{\psi''(v) - p(v) h''(v)}{1 - H(v)} + (1 - F_n^P(v)) \frac{\psi_n''(v) - \psi''(v)}{1 - H(v)} \right. \\
&\quad - (1 - F_n^P(v)) p(v) \frac{h_n''(v) - h''(v)}{1 - H(v)} - (1 - F_n^P(v)) \frac{\psi_n(v) - \psi(v)}{h(v)(1 - H(v))} h''(v) \\
&\quad + (1 - F_n^P(v)) p(v) \frac{h_n(v) - h(v)}{h(v)(1 - H(v))} h''(v) - (1 - F(v)) \frac{\psi''(v) - p(v) h''(v)}{1 - H(v)} \\
&\quad \left. + (1 - F_n^P(v)) (p_n(v) - p(v)) \left[\frac{h_n(v) - h(v)}{h(v)(1 - H(v))} h''(v) - \frac{h_n''(v) - h''(v)}{1 - H(v)} \right] \right]^2 \omega(v) dv.
\end{aligned}$$

Desarrollamos el cuadrado, de modo que tenemos las tres siguientes integrales

$$\begin{aligned}
& \frac{1}{4} \int \left[(1 - F_n^P(v)) \frac{\psi''(v) - p(v) h''(v)}{1 - H(v)} - (1 - F(v)) \frac{\psi''(v) - p(v) h''(v)}{1 - H(v)} \right]^2 \omega(v) dv \\
&+ \frac{1}{4} \int \left[(1 - F_n^P(v)) \frac{\psi_n''(v) - \psi''(v)}{1 - H(v)} - (1 - F_n^P(v)) p(v) \frac{h_n''(v) - h''(v)}{1 - H(v)} \right. \\
&\quad - (1 - F_n^P(v)) \frac{\psi_n(v) - \psi(v)}{h(v)(1 - H(v))} h''(v) + (1 - F_n^P(v)) p(v) \frac{h_n(v) - h(v)}{h(v)(1 - H(v))} h''(v) \\
&\quad \left. + (1 - F_n^P(v)) (p_n(v) - p(v)) \left(\frac{h_n(v) - h(v)}{h(v)(1 - H(v))} h''(v) - \frac{h_n''(v) - h''(v)}{1 - H(v)} \right) \right]^2 \omega(v) dv \\
&+ \frac{1}{2} \int \left[(1 - F_n^P(v)) \frac{\psi''(v) - p(v) h''(v)}{1 - H(v)} - (1 - F(v)) \frac{\psi''(v) - p(v) h''(v)}{1 - H(v)} \right] \\
&\quad \times \left[(1 - F_n^P(v)) \frac{\psi_n''(v) - \psi''(v)}{1 - H(v)} - (1 - F_n^P(v)) p(v) \frac{h_n''(v) - h''(v)}{1 - H(v)} \right. \\
&\quad - (1 - F_n^P(v)) \frac{\psi_n(v) - \psi(v)}{h(v)(1 - H(v))} h''(v) + (1 - F_n^P(v)) p(v) \frac{h_n(v) - h(v)}{h(v)(1 - H(v))} h''(v) \\
&\quad \left. + (1 - F_n^P(v)) (p_n(v) - p(v)) \left(\frac{h_n(v) - h(v)}{h(v)(1 - H(v))} h''(v) - \frac{h_n''(v) - h''(v)}{1 - H(v)} \right) \right] \omega(v) dv \\
&= \frac{1}{4} J_1 + \frac{1}{4} J_2 + \frac{1}{2} J_3.
\end{aligned}$$

Por la desigualdad de Cauchy-Schwarz, la tercera integral, J_3 , es despreciable si lo son las dos primeras. La despreciableidad de la integral J_2 se va a demostrar a continuación, mientras que la de la primera, J_1 , se verá en el paso 3.

Para demostrar que J_2 es despreciable, de nuevo por la desigualdad de Cauchy-Schwarz basta demostrar que la integral del cuadrado de cada uno de los sumandos es de orden $o_P(1)$, es decir, demostrar la despreciableidad de las siguientes integrales:

$$\begin{aligned}
& \int (1 - F_n^P(v))^2 \left(\frac{\psi_n''(v) - \psi''(v)}{1 - H(v)} \right)^2 \omega(v) dv \\
& - \int (1 - F_n^P(v))^2 p^2(v) \left(\frac{h_n''(v) - h''(v)}{1 - H(v)} \right)^2 \omega(v) dv \\
& - \int (1 - F_n^P(v))^2 \left(\frac{\psi_n(v) - \psi(v)}{h(v)(1 - H(v))} \right)^2 h''(v)^2 \omega(v) dv \\
& + \int (1 - F_n^P(v))^2 p^2(v) \left(\frac{h_n(v) - h(v)}{h(v)(1 - H(v))} \right)^2 h''(v)^2 \omega(v) dv \\
& + \int (1 - F_n^P(v))^2 (p_n(v) - p(v))^2 \left(\frac{h_n(v) - h(v)}{h(v)(1 - H(v))} \right)^2 h''(v)^2 \omega(v) dv \\
& - \int (1 - F_n^P(v))^2 (p_n(v) - p(v))^2 \left(\frac{h_n''(v) - h''(v)}{1 - H(v)} \right)^2 \omega(v) dv
\end{aligned}$$

La idea consiste en aplicar la generalización de los resultados de Silverman (1978) y Mack y Silverman (1982) al caso de las derivadas de orden k de las funciones de regresión y densidad:

$$\sup_{t \in \mathbb{R}} \left| \psi_n^{(k)}(t) - \psi^{(k)}(t) \right| = O_P \left(b^2 + n^{-1/2} b^{-1/2-k} (\log \frac{1}{b})^{1/2} \right) \quad (4.69)$$

$$\sup_{t \in \mathbb{R}} \left| h_n^{(k)}(t) - h^{(k)}(t) \right| = O_P \left(b^2 + n^{-1/2} b^{-1/2-k} (\log \frac{1}{b})^{1/2} \right) \quad (4.70)$$

y

$$\sup_{t \in [0, t_H]} \left| p_n^{(k)}(t) - p^{(k)}(t) \right| = O_P \left(n^{-1/2} b^{-1/2-k} (\log \frac{1}{b})^{1/2} \right). \quad (4.71)$$

Bajo las condiciones (H4) y (p2), las funciones $h''(\cdot)$ y $p''(\cdot)$ están acotadas, y por (v7) se tiene $nb^5 (\log \frac{1}{b})^{-1} \rightarrow \infty$, de modo que las integrales del desarrollo de J_2 son despreciables, y por tanto también J_2 .

Paso 3. En esta última etapa de la demostración se probará que

$$\frac{1}{4} \int \left([(1 - F_n^P(v)) - (1 - F(v))] \frac{\psi''(v) - p(v)h''(v)}{1 - H(v)} \right)^2 \omega(v) dv = o_P(1).$$

Esta integral es igual a $\int \alpha'(v)^2 (F_n^P(v) - F(v))^2 \omega(v) dv$, y dado que

$$F_n^P(t) = \overline{F_n^P}(t) + R_n(t) = F(t) + \frac{1}{n} \sum_{i=1}^n \varepsilon_i(t) + R_n(t),$$

entonces

$$\begin{aligned}
\int \alpha'(v)^2 (F_n^P(v) - F(v))^2 \omega(v) dv &= \int \alpha'(v)^2 \left(\frac{1}{n} \sum_{i=1}^n \varepsilon_i(v) + R_n(v) \right)^2 \omega(v) dv \\
&\leq \sup_{t \in [0, t_H]} |\alpha'(t)|^2 \int \left(\frac{1}{n} \sum_{i=1}^n \varepsilon_i(v) + R_n(v) \right)^2 \omega(v) dv,
\end{aligned}$$

donde

$$\begin{aligned} & \int \left(\frac{1}{n} \sum_{i=1}^n \varepsilon_i(v) + R_n(v) \right)^2 \omega(v) dv \\ &= \int \left(\frac{1}{n} \sum_{i=1}^n \varepsilon_i(v) \right)^2 \omega(v) dv + \int R_n^2(v) \omega(v) dv + 2 \frac{1}{n} \sum_{i=1}^n \int \varepsilon_i(v) R_n(v) \omega(v) dv. \end{aligned}$$

La segunda integral es de orden $o_P(1)$ al ser, bajo (v8), $\sup_{t \in [0, t_H]} |R_n^2(t)| = o(1)$ de forma casi segura, mientras que por la desigualdad de Cauchy-Schwarz, la tercera es también despreciable si lo son las dos primeras. Pasamos a demostrar ahora que la primera integral es despreciable.

Al tratarse de una sucesión de variables aleatorias positivas, basta demostrar que su esperanza tiende a cero. Descomponemos su esperanza en dos sumandos, según los índices i y j sean iguales o no:

$$\begin{aligned} \mathbb{E} \left[\int \left(\frac{1}{n} \sum_{i=1}^n \varepsilon_i(v) \right)^2 \omega(v) dv \right] &= \frac{n(n-1)}{n^2} \int \mathbb{E}[\varepsilon_1(v)] \mathbb{E}[\varepsilon_2(v)] \omega(v) dv \\ &\quad + \frac{1}{n} \int \mathbb{E}[\varepsilon_1^2(v)] \omega(v) dv. \end{aligned}$$

Por el lema 4.3.2, podemos concluir que la integral $\int \left(\frac{1}{n} \sum_{i=1}^n \varepsilon_i(v) \right)^2 \omega(v) dv$ es de orden $o_P(1)$. De todo ello se sigue el resultado del paso 3. Dicha expresión, junto con la del paso 2 y (4.67), permite concluir (4.65). ■

Para demostrar la convergencia en probabilidad de $\hat{c}_1(\cdot)$ a $c_1(\cdot)$, falta probar que

$$\int (f_n^P(v) \alpha_n(v) - f(v) \alpha(v))^2 \omega(v) dv \xrightarrow{P} 0.$$

En el primer paso, aproximaremos en la integral el estimador presuavizado $f_n^P(\cdot)$ por su representación asintótica $\overline{f_n^P}(\cdot)$, y en el segundo aproximaremos $\alpha_n(\cdot)$ por $\alpha(\cdot)$.

Lema 4.3.12

Bajo las condiciones (K1), (H3), (p2), (p3), (f1), (v5), (v7) y (v8) se tiene

$$\int (f_n^P(v) \alpha_n(v) - f(v) \alpha(v))^2 \omega(v) dv \xrightarrow{P} 0. \quad (4.72)$$

Demostración.

La demostración es totalmente análoga a la del lema 4.3.10.

Paso 1. Empezamos eliminando el denominador aleatorio del estimador $\alpha_n(\cdot)$:

$$\begin{aligned} & \int (f_n^P(v) \alpha_n(v) - f(v) \alpha(v))^2 \omega(v) dv \\ &= \int \left(\frac{1}{2} f_n^P(v) \int_0^v \frac{\psi_n''(w) - p_n(w) h_n''(w)}{1 - H(w)} dw - f(v) \alpha(v) \right)^2 \omega(v) dv \\ & \quad + o_P \left(\int \left(\frac{1}{2} f_n^P(v) \int_0^v \frac{\psi_n''(w) - p_n(w) h_n''(w)}{1 - H(w)} dw - f(v) \alpha(v) \right)^2 \omega(v) dv \right). \end{aligned}$$

Esta igualdad es inmediata al aplicar (4.66) y (4.68).

Paso 2. En este paso se demostrará que

$$\begin{aligned} & \int \left(\frac{1}{2} f_n^P(v) \int_0^v \frac{\psi_n''(w) - p_n(w) h_n''(w)}{1 - H(w)} dw - f(v) \alpha(v) \right)^2 \omega(v) dv \\ &= \frac{1}{4} \int \left[(f_n^P(v) - f(v)) \int_0^v \frac{\psi''(w) - p(w) h''(w)}{1 - H(w)} dw \right]^2 \omega(v) dv \\ & \quad + o_P \left(\int \left[(f_n^P(v) - f(v)) \int_0^v \frac{\psi''(w) - p(w) h''(w)}{1 - H(w)} dw \right]^2 \omega(v) dv \right). \end{aligned}$$

Bajo las condiciones (H3) y (p2), realizamos las mismas descomposiciones que en la demostración del lema anterior, y aplicamos los resultados (4.69)-(4.71).

Paso 3. Por último, probaremos que

$$\frac{1}{4} \int \left[(f_n^P(v) - f(v)) \int_0^v \frac{\psi''(w) - p(w) h''(w)}{1 - H(w)} dw \right]^2 \omega(v) dv = o_P(1). \quad (4.73)$$

Por (4.66), esta integral es igual a

$$\begin{aligned} & \int (f_n^P(v) - f(v))^2 \alpha^2(v) \omega(v) dv = \int (\overline{f_n^P}(v) - f(v))^2 \alpha^2(v) \omega(v) dv \\ & \quad + \int e_n^2(v) \alpha^2(v) \omega(v) dv + 2 \int (\overline{f_n^P}(v) - f(v)) e_n(v) \alpha^2(v) \omega(v) dv. \end{aligned}$$

Aplicando los resultados de los lemas 4.3.1, 4.3.5 y 4.3.6 y el teorema 3.2.1 se tiene (4.73), y por tanto, (4.72). ■

4.3.2. Estudio de simulación

En este estudio de simulación, en el que hemos considerado los cuatro modelos simulados en los capítulos anteriores, se investiga el comportamiento práctico de los selectores de las ventanas plug-in del estimador presuavizado, $f_n^P(\cdot)$, en comparación con el selector plug-in del estimador tipo núcleo con pesos de Kaplan-Meier, $f_n^{KM}(\cdot)$.

El selector tipo plug-in de la ventana considerado para el estimador $f_n^{KM}(\cdot)$ ha sido el propuesto por Sánchez Sellero, González Manteiga y Cao (1999) en el contexto más general de censura aleatoria por la derecha y truncamiento por la izquierda:

$$\widehat{s}^{KM} = \left(\frac{c_K \int \left(\frac{1 - F_n^{KM}(v)}{1 - H_n(v)} \right)^2 \omega(v) dH_n^1(v)}{d_K^2 \int \widehat{f}_g''(v)^2 \omega(v) dv} \right)^{1/5} n^{-1/5}, \quad (4.74)$$

donde $H_n^1(t) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{\{Z_i \leq t, \delta_i=1\}}$ es el estimador empírico de $H^1(t) = P(Z \leq t, \delta = 1)$, y $\widehat{f}_g''(\cdot)$ es la segunda derivada del estimador tipo núcleo de la función de densidad $f(\cdot)$ con ventana g .

Obsérvese que este selector necesita únicamente una ventana piloto g para estimar la integral que aparece en el denominador de \widehat{s}^{KM} , puesto que las estimaciones que figuran en el numerador pueden hacerse sin suavización. La parte dominante de la ventana piloto g que minimiza el error cuadrático medio

$$\mathbb{E} \left[\left(\int \widehat{f}_g''(v)^2 \omega(v) dv - \int f''(v)^2 \omega(v) dv \right)^2 \right]$$

depende del signo de la integral $\int f''(v) f^{(4)}(v) \omega(v) dv$ (ver teorema 2.3 de Sánchez Sellero, González Manteiga y Cao (1999)). Si dicha integral es negativa se obtiene un orden mejor de convergencia. En este estudio de simulación, hemos considerado las mismas funciones de peso $\omega(\cdot)$ para cada modelo que en capítulos anteriores. En todos los casos dicha integral es negativa. La parte dominante de la ventana piloto g es, en este caso,

$$g = \left(\frac{i_{K_1} \int \left(\frac{1 - F(v)}{1 - H(v)} \right)^2 \omega(v) dH^1(v)}{d_{K_1} \int f'''(v)^2 \omega(v) dv} \right)^{1/7} n^{-1/7}. \quad (4.75)$$

con $i_{K_1} = \int K_1''(v)^2 dv$, siendo $K_1(\cdot)$ una función núcleo que puede ser, en principio, diferente a $K(\cdot)$.

La expresión de la ventana piloto, g , dada en (4.75) ha sido estimada sustituyendo las integrales por estimaciones de ellas:

$$\widehat{g} = \left(\frac{i_{K_1} \int \left(\frac{1 - F_n^{KM}(v)}{1 - H_n(v)} \right)^2 \omega(v) dH_n^1(v)}{d_{K_1} \int \widehat{f}_g'''(v)^2 \omega(v) dv} \right)^{1/7} n^{-1/7},$$

donde $f(\cdot)$ ha sido estimada suponiendo un modelo paramétrico, en concreto, una mixtura de tres lognormales. El núcleo $K_1(\cdot)$ usado ha sido el gaussiano.

Para obtener el valor de las ventanas plug-in del estimador presuavizado $f_n^P(\cdot)$ se han de estimar las funciones $F(\cdot)$, $f(\cdot)$, $f''(\cdot)$, $H(\cdot)$, $h(\cdot)$, $h'(\cdot)$, $h''(\cdot)$, $p(\cdot)$, $p'(\cdot)$ y $p''(\cdot)$. Aunque para cada estimación sería necesario una ventana piloto diferente, excepto para la función de distribución $H(\cdot)$, que se ha estimado por su versión empírica $H_n(\cdot)$, se han considerado únicamente cuatro ventanas piloto.

La primera ventana piloto, b_0 , se ha estimado por validación cruzada, y se ha usado para estimar la función $p(\cdot)$ mediante el estimador de Nadaraya-Watson $p_n(\cdot)$, sus derivadas primera y segunda y, a partir de $p_n(\cdot)$, el estimador presuavizado $F_n^P(\cdot)$. Con la segunda ventana piloto, llamada b_1 , se han estimado la función de densidad $h(\cdot)$ y su derivada primera $h'(\cdot)$. La expresión de esta ventana piloto es

$$b_1 = \left(\frac{i_K}{d_K \int \widehat{h}'''(v)^2 \omega(v) dv} \right)^{1/7} n^{-1/7}, \quad (4.76)$$

que corresponde con la ventana asintóticamente óptima para estimar la curvatura de la función de densidad $h(\cdot)$ (ver Cao *et al.* (1994)). La integral del denominador ha sido estimada paraméricamente suponiendo que $h(\cdot)$ sigue una distribución de Weibull. La tercera ventana piloto, s_0 , se usó en la estimación presuavizada de la función de densidad, $f_n^P(\cdot)$, a partir de $F_n^P(\cdot)$. Su expresión es análoga a (4.76), donde la derivada tercera de la función de densidad $f(\cdot)$, que aparece en el denominador, ha sido estimada suponiendo el mismo modelo paramétrico que en la estimación de la ventana g , es decir, una mixtura de tres lognormales. Los parámetros del modelo supuesto para $f(\cdot)$ han sido estimados por el método de máxima verosimilitud ponderada por los pesos presuavizados con ventana b_0 . Finalmente, la función $f''(\cdot)$ ha sido estimada con el método núcleo y la misma ventana ventana piloto, g , que se usó en el cálculo de \widehat{s}^{KM} .

Se han simulado $m = 500$ muestras diferentes de los modelos 1-4 con tamaños muestrales $n = 50, 100$ y 200 con el núcleo de Epanechnikov. Una vez estimadas las funciones involucradas, se calculó $\psi_n(L) = (\widehat{c}_1(L) \widehat{c}_2^4(L))^{1/5}$, donde $\widehat{c}_1(\cdot)$ y $\widehat{c}_2(\cdot)$ vienen dadas por (4.31) y (4.32) respectivamente, y para cada muestra se ha calculado el valor \widehat{L} que minimiza $\psi_n(\cdot)$. Aunque los modelos 2-4 presentan una eficiencia teórica de tercer orden, y para ellos el valor asintóticamente óptimo es $L_0 = 0$, para cada muestra de los tamaños simulados se obtuvieron siempre valores $\widehat{L} > 0$.

A partir de dicho valor para la constante L , las ventanas plug-in consideradas han sido:

$$\widehat{s}^P = \left(\frac{\widehat{c}_2(\widehat{L})}{4\widehat{c}_1(\widehat{L})} \right)^{1/5} n^{-1/5} \quad \text{y} \quad \widehat{b}^P = \widehat{L} \times \widehat{s}^P.$$

Cuando no se presuaviza en absoluto, es decir, cuando $b = 0$, y teniendo en cuenta la observación 4.1.4, las ventanas de suavizado *AMISE* para el estimador presuavizado $f_n^P(\cdot)$ y para el estimador con pesos de Kaplan-Meier $f_n^{KM}(\cdot)$ coinciden.

En la figura 4.7 pueden verse los diagramas de caja de las ventanas de presuavizado, s y b , plug-in, calculadas para $m = 500$ muestras diferentes de los modelos 1-4 con tamaño muestral $n = 100$. Las líneas horizontales señalan el valor de las ventanas *MISE* (línea continua) y *AMISE* (línea discontinua). En todos los modelos, excepto en el modelo 4, la ventana plug-in de suavizado, s , es sesgada con respecto a las ventanas teóricas *MISE* y *AMISE*. Las ventanas plug-in de presuavizado, b , son también sesgadas con respecto a b_{AMISE} , aunque se aproximan más a las ventanas óptimas b_{MISE} .

Para estudiar el efecto de este sesgo en la estimación final de la densidad se ha calculado, para cada una de las $m = 500$ muestras, el error cuadrático integrado (*ISE*) del estimador presuavizado $f_n^P(\cdot)$ con ventanas plug-in, \widehat{s}^P y \widehat{b}^P :

$$ISE \left(f_{n, \widehat{s}^P, \widehat{b}^P}^P(\cdot) \right) = \int \left(f_{n, \widehat{s}^P, \widehat{b}^P}^P(v) - f(v) \right)^2 \omega(v) dv,$$

y se ha comparado con el del estimador suavizado de Kaplan-Meier $f_n^{KM}(\cdot)$ con ventana plug-in \widehat{s}^{KM} .

En la tabla 4.2 se pueden ver, para cada modelo y para cada tamaño muestral, la eficiencia relativa del estimador de Kaplan-Meier, $f_n^{KM}(\cdot)$, con respecto al estimador presuavizado $f_n^P(\cdot)$, calculados ambos con sus respectivas ventanas plug-in:

$$RE^{KM,P} \left(\widehat{s}^P, \widehat{b}^P; \widehat{s}^{KM} \right) = \frac{MISE \left(f_{n, \widehat{s}^P, \widehat{b}^P}^P(\cdot) \right)}{MISE \left(f_{n, \widehat{s}^{KM}}^{KM}(\cdot) \right)},$$

donde el *MISE* se ha aproximado por la media muestral de los valores de *ISE* en las $m = 500$ muestras:

$$MISE \left(f_{n, \widehat{s}^P, \widehat{b}^P}^P(\cdot) \right) = \mathbb{E} \left[ISE \left(f_{n, \widehat{s}^P, \widehat{b}^P}^P(\cdot) \right) \right] \simeq \frac{1}{m} \sum_{i=1}^m ISE \left(f_{n, \widehat{s}^P, \widehat{b}^P}^P(\cdot) \right).$$

Valores de $RE^{KM,P}(\cdot, \cdot, \cdot)$ menores que 1 indican una mayor eficiencia del estimador presuavizado $f_n^P(\cdot)$ con respecto a $f_n^{KM}(\cdot)$ cuando para su cálculo se usan las estimaciones de las ventanas óptimas mediante los selectores de tipo plug-in.

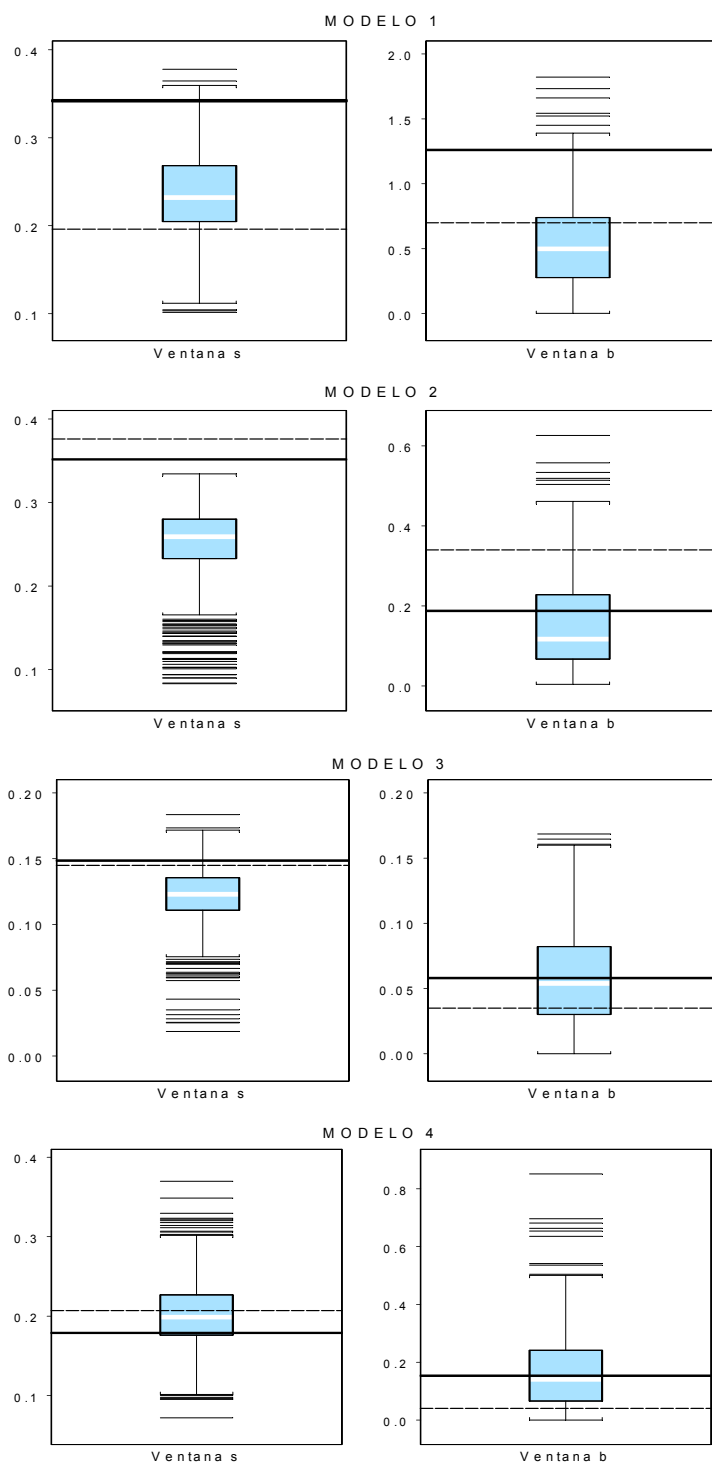


Figura 4.7. Diagrama de caja de las ventanas s y b plug-in para los modelos 1-4 con tamaño muestral $n = 100$. Las líneas horizontales señalan el valor de las ventanas $MISE$ (línea continua) y $AMISE$ (línea discontinua).

Tabla 4.2 Eficiencia relativa del estimador con pesos de Kaplan-Meier $f_n^{KM}(\cdot)$ con respecto al estimador presuavizado $f_n^P(\cdot)$, calculados ambos con sus respectivas ventanas plug-in.

MODELO	$n = 50$	$n = 100$	$n = 200$
1	0.674	0.617	0.563
2	0.901	0.879	0.876
3	1.220	1.130	1.028
4	0.946	0.972	0.996

A pesar de los sesgos observados en la figura 4.7, esta eficiencia relativa es, para cualquier tamaño muestral, menor que 1 en los modelos 1, 2 y 4. Esta mayor eficiencia de $f_n^P(\cdot)$ es llamativa, sobre todo, en el modelo 1, para el que la eficiencia relativa (en tanto por ciento) de $f_n^P(\cdot)$ es, aproximadamente, de un 150% para $n = 50$, y aumenta con el tamaño muestral. Esto, sin embargo, no es de extrañar, puesto que en el modelo 1 el estimador presuavizado presenta una eficiencia teórica de primer orden con respecto al estimador $f_n^{KM}(\cdot)$.

Para el resto de los modelos, en los que la eficiencia teórica de $f_n^P(\cdot)$ es de tercer orden, la eficiencia relativa $RE^{KM,P}(\cdot, \cdot, \cdot)$ tiende a 1 a medida que aumenta el tamaño muestral. Aunque el estimador presuavizado $f_n^P(\cdot)$ es más eficiente que $f_n^{KM}(\cdot)$ en los modelos 1, 2 y 4, no es así en el modelo 3 (aunque este efecto se mitiga cuando aumenta n). Esto puede ser debido al sesgo del selector de la ventana de suavizado s tipo plug-in (ver figura 4.7).

Para cada muestra se ha calculado la eficiencia relativa del estimador de Kaplan-Meier, $f_n^{KM}(\cdot)$, con respecto al estimador presuavizado, $f_n^P(\cdot)$, en términos del ISE cometido, es decir,

$$RE_{ISE}^{KM,P}(\hat{s}^P, \hat{b}^P; \hat{s}^{KM}) = \frac{ISE\left(f_{n, \hat{s}^P, \hat{b}^P}^P(\cdot)\right)}{ISE\left(f_{n, \hat{s}^{KM}}^{KM}(\cdot)\right)}.$$

Valores de $RE_{ISE}^{KM,P}(\cdot, \cdot, \cdot)$ menores que 1 indican que, para cada muestra, el estimador presuavizado comete un menor error ISE que el estimador con pesos de Kaplan-Meier. La figura 4.8 muestra las estimaciones de Parzen-Rosenblatt de la función de densidad de $RE_{ISE}^{KM,P}(\cdot, \cdot, \cdot)$ a partir de los $m = 500$ valores observados, con núcleo gaussiano y selector de ventana de Sheather-Jones. Se puede ver, claramente, que el estimador presuavizado, $f_n^P(\cdot)$, es más eficiente en términos de ISE que el estimador $f_n^{KM}(\cdot)$.

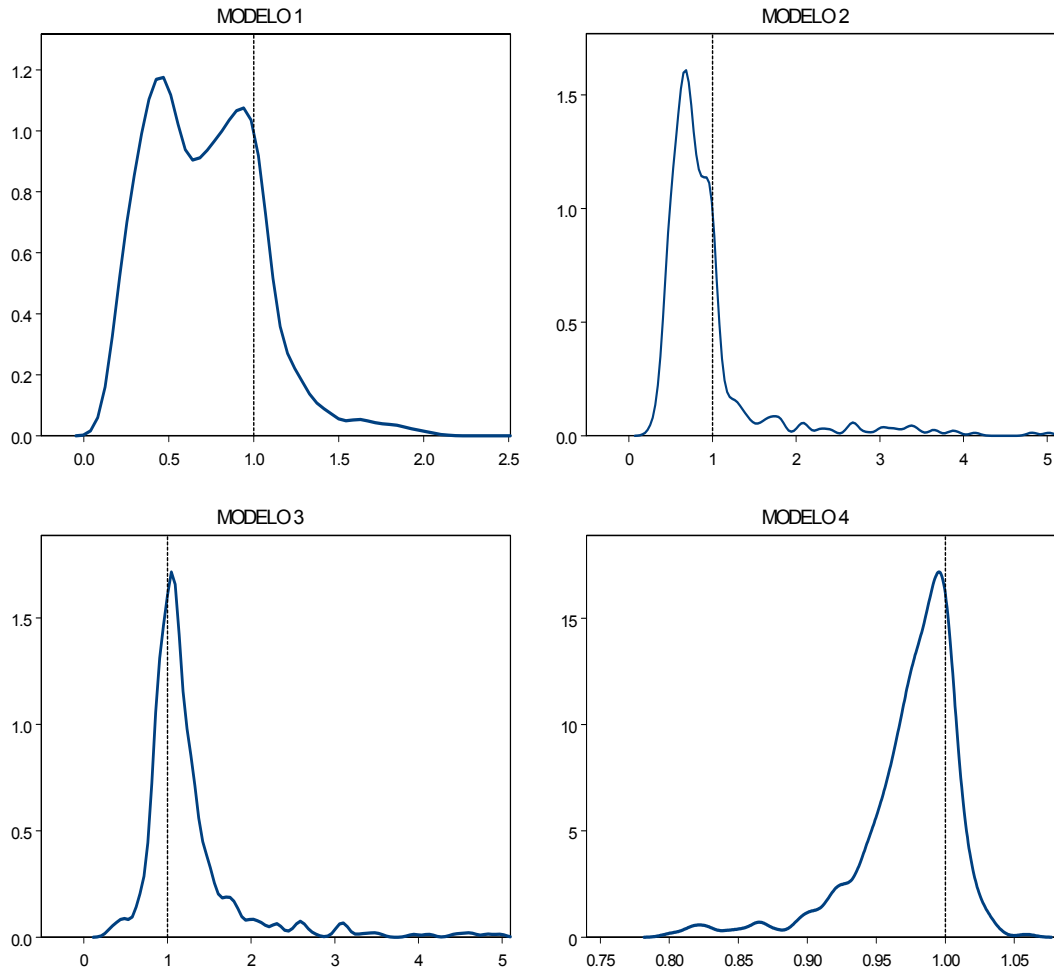


Figura 4.8. Estimaciones de Parzen-Rosenblatt de las funciones de densidad de la eficiencia relativa $RE_{ISE}^{KM,P}(\cdot, \cdot, \cdot)$ para $m = 500$ muestras de tamaño $n = 100$ de los modelos 1-4.

Aunque, como se vio en esta sección, el selector de las ventanas de tipo plug-in es consistente en probabilidad, para tamaños muestrales relativamente pequeños se hace necesario estimar adecuadamente todas las funciones implicadas en la expresión de las ventanas $AMISE$ con el fin de aproximar suficientemente bien las ventanas plug-in a las ventanas $AMISE$ y, por tanto, a las ventanas $MISE$. Esto implica la correcta selección de ocho ventanas piloto. Es por ello que, con el fin de evitar parte de estas estimaciones piloto que pueden dar lugar a errores en la estimación final de las ventanas, proponemos en la siguiente sección otros procedimientos para seleccionar dichas ventanas basados en el método bootstrap.

4.4. Selectores de las ventanas de tipo bootstrap

Como se definió con anterioridad, las ventanas *MISE* son aquellas que minimizan el error cuadrático medio integrado de $f_n^P(\cdot)$. Usando la notación clásica del método bootstrap es, por tanto, de gran interés conocer la distribución de la variable aleatoria

$$\mathbf{R}_{(s,b)}(\mathbf{X}, F) = \int (f_n^P(v) - f(v))^2 \omega(v) dv,$$

y, en concreto, su media. El método bootstrap aproxima la distribución de $\mathbf{R}_{(s,b)}(\mathbf{X}, F)$ por la distribución en el remuestreo de $\mathbf{R}^* = \mathbf{R}_{(s,b)}(\mathbf{X}^*, F_n)$, es decir, de la variable aleatoria resultante de sustituir en $\mathbf{R}_{(s,b)}(\mathbf{X}, F)$ la distribución teórica $F(\cdot)$ por una estimación $F_n(\cdot)$, y la muestra observada \mathbf{X} por la remuestra bootstrap \mathbf{X}^* obtenida a partir de $F_n(\cdot)$. La técnica más usada para aproximar la distribución de \mathbf{R}^* es mediante Montecarlo, generando B remuestras $\mathbf{X}_1^*, \dots, \mathbf{X}_B^*$ de tamaño n a partir de $F_n(\cdot)$ y aproximando la distribución en el remuestreo de \mathbf{R}^* por la distribución empírica de $\mathbf{R}_1^* = \mathbf{R}_{(s,b)}(\mathbf{X}_1^*, F_n), \dots, \mathbf{R}_B^* = \mathbf{R}_{(s,b)}(\mathbf{X}_B^*, F_n)$.

En resumen, la idea general de los selectores de la ventana bootstrap consiste en estimar la función *MISE* mediante remuestreo, y obtener las ventanas que minimizan esta aproximación. Aunque este método está basado en el método de Montecarlo, y por tanto requiere cálculos intensivos, presenta la ventaja de que no necesita hipótesis sobre el mecanismo que genera los datos, es decir, sobre la función de densidad desconocida $f(\cdot)$ (ver Efron (1979), Hall (1992) y Efron y Tibshirani (1993) entre otros).

El mecanismo de remuestreo ha de adaptarse a cada contexto particular. Así, en una situación de datos censurados, Efron (1981) introduce dos métodos de remuestreo conocidos como bootstrap simple y bootstrap obvio, ambos equivalentes (para más detalle, ver Reid (1981), Akritas (1986) y Cao (1988)). Si además se asume que la muestra \mathbf{X} proviene de una variable aleatoria con función de densidad $f(\cdot)$, es necesario utilizar un remuestreo bootstrap suavizado para aproximar adecuadamente la parte del sesgo de la función *MISE* (ver Marron (1992)). Esto implica obtener las remuestras \mathbf{X}^* de algún estimador adecuado de $f(\cdot)$ en lugar de remuestrear de la distribución empírica $F_n(\cdot)$.

En el contexto con datos censurados que provienen de una distribución con función de densidad, González Manteiga, Cao y Marron (1996) proponen tres planes de remuestreo diferentes, según el orden en que se llevan a cabo la suavización y la censura: el remuestreo *suavizado-censurado SC*, versión suavizada del remuestreo de Efron (1981), en el que primero se remuestra de una estimación suavizada de la distribución de los datos \mathbf{X} , y posteriormente se introduce el mecanismo de censura, y los

remuestreos *censurado-suavizado* *CS1* y *CS2*, ambos equivalentes, en los que primero se generan remuestras censuradas, y después se realiza la suavización.

En los siguientes apartados se adaptan los remuestreos *SC* y *CS2* al contexto presuavizado, en los que la suavización consiste en remuestrear del estimador presuavizado de la densidad $f_n^P(\cdot)$ en lugar del estimador Kaplan-Meier suavizado $f_n^{KM}(\cdot)$. Se proponen dos remuestreos *suavizado-censurado* presuavizados: *SC1* y *SC2*, y un remuestreo *censurado-suavizado* *CS*. Además, en el estudio de simulación de la siguiente sección, compararemos estos tres remuestreos con dos remuestreos no suavizados, que llamaremos *SP* (*simple presuavizado*) y *OP* (*obvio presuavizado*), y que en la práctica dan también buenos resultados.

4.4.1. Remuestreos bootstrap para datos con censura

Los métodos de remuestreo *suavizado-censurado* *SC1* y *SC2* que proponemos son una versión presuavizada del método *SC* introducido por González Manteiga, Cao y Marron (1996). En este remuestreo se intenta introducir la suavización tan pronto como sea posible, en un caso estimando la función de densidad de las variables Y y C mediante los estimadores presuavizados $f_n^P(\cdot)$ y $g_n^P(\cdot)$ respectivamente, en lugar de los estimadores Kaplan-Meier suavizados, y generando posteriormente remuestras de dichos estimadores (*Remuestreo SC1*), o bien de la estimación no paramétrica de la función de densidad de los datos observados (*Remuestreo SC2*). Las remuestras bootstrap se obtiene de acuerdo a los siguientes procedimientos:

Remuestreo SC1

1. Construir los estimadores presuavizados $f_n^P(\cdot)$ y $g_n^P(\cdot)$ de las funciones de densidad de las variables Y y C respectivamente.
2. Generar remuestras, de forma independiente: $Y_i^* \sim f_n^P(\cdot)$ y $C_i^* \sim g_n^P(\cdot)$, para $i = 1, \dots, n$.
3. Contruir $Z_i^* = \min(Y_i^*, C_i^*)$ y $\delta_i^* = \mathbf{1}_{\{Y_i^* \leq C_i^*\}}$ para $i = 1, \dots, n$.

Remuestreo SC2

1. Construir el estimador suavizado de Parzen-Rosenblatt, $h_n(\cdot)$, de la función de densidad de la variable observada Z .
2. Generar la remuestra bootstrap $\{Z_i^*\}_{i=1}^n$ a partir del estimador suavizado $h_n(\cdot)$.
3. Generar la remuestra bootstrap $\{\delta_i^*\}_{i=1}^n$ de modo que la distribución de δ_i^* condicionada a $Z_i^* = t$ sea una Bernoulli de parámetro $p_n(t)$, siendo $p_n(\cdot)$ el estimador de Nadaraya-Watson.

La notación *CS* del siguiente remuestreo hace referencia al hecho de que la parte de la censura se realiza antes que la de suavización. La idea consiste en modificar el remuestreo propuesto por Efron (1981) para incorporar la suavización.

Remuestreo CS

1. Construir los estimadores presuavizados de la función de distribución de Y y C , $F_n^P(\cdot)$ y $G_n^P(\cdot)$.
2. Generar remuestras, de forma independiente: $Y_i^* \sim F_n^P(\cdot)$ y $C_i^* \sim G_n^P(\cdot)$ y construir $\tilde{Z}_i^* = \min(Y_i^*, C_i^*)$, para $i = 1, \dots, n$.
3. Hacer $Z_i^* = \tilde{Z}_i^* + sV_i$, siendo V_i una variable aleatoria con densidad $K(\cdot)$.
4. Construir $\delta_i^* = \mathbf{1}_{\{Y_i^* \leq C_i^*\}}$ para $i = 1, \dots, n$.

Asímismo hemos considerado dos métodos de remuestreo sin suavizar, llamados *SP* y *OP*. De ellos, el llamado *OP* es la versión presuavizada del método obvio propuesto por Efron (1981), mientras que el método *SP* es la versión presuavizada del método simple.

Remuestreo SP

1. Construir el estimador empírico de la función de distribución $H_n(\cdot)$ de los datos observados $\{Z_i\}_{i=1}^n$.
2. Generar muestras $\{Z_i^*\}_{i=1}^n$ a partir de $H_n(\cdot)$, es decir, utilizando muestreo con reemplazamiento y masa de probabilidad $1/n$ en cada valor observado Z_i .
3. Generar la remuestra bootstrap $\{\delta_i^*\}_{i=1}^n$ de modo que la distribución de δ_i^* condicionada $Z_i^* = t$ es una Bernoulli de parámetro $p_n(t)$, siendo $p_n(\cdot)$ el estimador de Nadaraya-Watson de la función $p(\cdot)$.

Remuestreo OP

1. Construir los estimadores presuavizados de la función de distribución de Y y C , $F_n^P(\cdot)$ y $G_n^P(\cdot)$.
2. Generar muestras, de forma independiente: $Y_i^* \sim F_n^P(\cdot)$ y $C_i^* \sim G_n^P(\cdot)$ para $i = 1, \dots, n$.
3. Contruir $Z_i^* = \min(Y_i^*, C_i^*)$ y $\delta_i^* = \mathbf{1}_{\{Y_i^* \leq C_i^*\}}$ para $i = 1, \dots, n$.

Observación 4.4.1

Los dos métodos no suavizados, SP y OP , son equivalentes si, a la hora de construir los estimadores presuavizados de la función de distribución de Y (análogamente para C), se considera el estimador

$$F_n^{P(2)}(t) = 1 - \prod_{Z_{(i)} \leq t} \left(1 - \frac{1}{n-i+1} \right)^{p_n(Z_{(i)})}$$

en lugar de

$$F_n^P(t) = 1 - \prod_{Z_{(i)} \leq t} \left(1 - \frac{p_n(Z_{(i)})}{n-i+1} \right).$$

En ese caso, el producto de los estimadores presuavizados de las funciones de supervivencia $1 - F(\cdot)$ y $1 - G(\cdot)$ coincide con la supervivencia empírica de la variable observada $1 - H(\cdot)$, es decir,

$$\left(1 - F_n^{P(2)}(t) \right) \left(1 - G_n^{P(2)}(t) \right) = 1 - H_n(t).$$

Obsérvese que ambos estimadores son asintóticamente equivalentes, ya que

$$F_n^P(t) - F_n^{P(2)}(t) = O_P(n^{-1}).$$

Sea cual sea el procedimiento para obtener las remuestras bootstrap, las ventanas bootstrap se calculan siguiendo los pasos:

1. Definir el análogo bootstrap del estimador presuavizado de la densidad:

$$f_n^{P*}(t) = \int K_s(t-v) dF_n^{P*}(v)$$

con ventana de suavizado s , donde

$$F_n^{P*}(t) = 1 - \prod_{Z_{(i)}^* \leq t} \left(1 - \frac{p_n(Z_{(i)}^*)}{n-i+1} \right) \text{ y } p_n(Z_{(i)}^*) = \frac{\frac{1}{nb} \sum_{j=1}^n K\left(\frac{Z_{(i)}^* - Z_j^*}{b}\right) \delta_{[j]}^*}{\frac{1}{nb} \sum_{j=1}^n K\left(\frac{Z_{(i)}^* - Z_j^*}{b}\right)}.$$

con ventana de suavizado b .

2. Obtener la versión bootstrap del $MISE(f_n^P(\cdot))$:

$$MISE^*(f_n^{P*}(\cdot)) = \mathbb{E}^* \left[\int (f_n^{P*}(v) - f_n^P(v))^2 \omega(v) dv \right].$$

En la práctica, se usa una aproximación de Montecarlo, calculando el promedio de los valores que toma la función

$$ISE^* (f_n^{P^*}(\cdot)) = \int (f_n^{P^*}(v) - f_n^P(v))^2 \omega(v) dv$$

en un número B de remuestras:

$$MISE^* (f_n^{P^*}(\cdot)) \simeq \frac{1}{B} \sum_{i=1}^B \int (f_n^{P^*i}(v) - f_n^P(v))^2 \omega(v) dv.$$

3. Calcular el par de ventanas (s, b) que minimizan la aproximación de $MISE^* (f_n^{P^*}(\cdot))$.

En la siguiente sección se lleva a cabo un estudio de simulación en el que se comprobará el comportamiento en la práctica de los diferentes planes de remuestro en la selección de los parámetros ventana s y b .

4.4.2. Estudio de simulación

Para implementar los selectores bootstrap de las ventanas, hemos aproximado la función $MISE^* (f_n^{P^*}(\cdot))$ mediante la media muestral de $ISE^* (f_n^{P^*}(\cdot))$:

$$\begin{aligned} MISE^* (f_n^{P^*}(\cdot)) &\simeq \frac{1}{B} \sum_{i=1}^B ISE^* (f_n^{P^*}(\cdot)) \\ &= \frac{1}{B} \sum_{i=1}^B \int (f_n^{P^*i}(v) - f_n^P(v))^2 \omega(v) dv \end{aligned}$$

calculado con $B = 500$ remuestras de tamaño $n = 100$ de los cuatro modelos considerados en estudios de simulación previos, y hemos obtenido las ventanas s^* y b^* que minimizan dicha aproximación.

Nótese que se necesitan dos ventanas piloto, s_0 y b_0 , para calcular $f_n^P(\cdot)$. La ventana piloto de presuavizado, b_0 , se usará para calcular el estimador $p_n(\cdot)$ de Nadaraya-Watson y, a partir de $p_n(\cdot)$ obtener, si el método de remuestreo lo requiere, las estimaciones presuavizadas de la función de distribución $F_n^P(\cdot)$ y $G_n^P(\cdot)$. Esta ventana se obtuvo aplicando validación cruzada (ver Stone (1974)). La expresión de otra ventana piloto, la de suavizado s_0 , es exactamente igual a la usada en el estudio de simulación para las ventanas plug-in, es decir

$$s_0 = \left(\frac{i_K}{d_K \int \hat{f}'''(v)^2 \omega(v) dv} \right)^{1/7} n^{-1/7},$$

Tabla 4.3. Porcentaje de veces que $MISE^*(f_n^{P^*}(\cdot))$ no presenta un mínimo local, b^* , para cada remuestreo, en los modelos 1-4.

	Modelo 1	Modelo 2	Modelo 3	Modelo 4
<i>SP</i>	39.6	10.0	27.95	2.8
<i>OP</i>	42.4	17.0	13.62	31.6
<i>SC1</i>	57.0	22.4	14.89	23.4
<i>SC2</i>	46.2	11.4	19.80	15.4
<i>CS</i>	63.2	19.8	23.87	35.0

donde la derivada tercera de la función de densidad $f(\cdot)$, que aparece en el denominador, ha sido estimada suponiendo una mezcla de tres lognormales, cuyos parámetros han sido estimados por el método de máxima verosimilitud ponderada por los pesos presuavizados calculados con ventana b_0 .

En el plan de remuestreo *SC2* los datos Z_i^* se obtienen del estimador de Parzen-Rosenblatt, $h_n(\cdot)$, de la función de densidad de la variable observada, Z . Para su construcción hemos considerado una tercera ventana piloto b_1 , obtenida al aplicar a los valores observados $\{Z_i\}_{i=1}^n$ la fórmula (4.76), análoga a la expresión de la ventana s_0 , pero sustituyendo la estimación pertinente de $f'''(\cdot)$ por la de $h'''(\cdot)$. De nuevo hemos supuesto un modelo paramétrico para la densidad, $h(\cdot)$, en este caso una distribución de Weibull cuyos parámetros han sido estimados también mediante máxima verosimilitud.

A diferencia del método de estimación de las ventanas tipo plug-in, el método bootstrap requiere, dependiendo del remuestreo, la estimación de dos o, a lo sumo, de tres funciones poblacionales que involucren suavizado: la función $p(\cdot)$, con cuya estimación $p_n(\cdot)$ se calcularán las estimaciones presuavizadas de $F(\cdot)$ y $G(\cdot)$, y finalmente las funciones de densidad $f(\cdot)$ y $h(\cdot)$.

En el caso de la ventana de presuavizado b^* , para algunas muestras la función $MISE^*(f_n^{P^*}(\cdot))$ no presenta un mínimo local dentro del intervalo de búsqueda. Estos intervalos han sido $[0.1, 2.5]$, $[0.05, 0.7]$, $[0.01, 0.4]$ y $[0.005, 0.3]$ para los modelos 1-4 respectivamente. La frecuencia con que ocurre esto depende del remuestreo y del modelo considerado (ver tabla 4.3). Este porcentaje se puede considerar alto en el modelo 1, pero esto es lógico, puesto que en dicho modelo la función $p(\cdot)$ es prácticamente plana, de modo que cualquier ventana b suficientemente grande proporciona buenas estimaciones $p_n(\cdot)$.

Las figuras 4.9-4.12 representan la estimación no paramétrica de Parzen-Rosenblatt de la función de densidad de las ventanas bootstrap s^* y b^* . La ventana de suavizado usada ha sido la propuesta por Sheather-Jones. La línea vertical indica el valor de la correspondiente ventana $MISE$.

En el modelo 1 se observa un ligero sesgo de la ventana de suavizado bootstrap s^* . Esto es debido a que en ese modelo la función de densidad teórica $f(\cdot)$ presenta una asíntota en $t = 1$ (ver figura 4.1), y por tanto es muy difícil de aproximar mediante un modelo paramétrico de una mixtura de lognormales. Esto produce infraestimaciones en la ventana piloto s_0 y, en consecuencia, ventanas bootstrap s^* ligeramente sesgadas.

Por otra parte, las ventanas de presuavizado bootstrap b^* presentan cierto sesgo en los modelos 1, 2 y 4. Esto, sin embargo, no influye demasiado en la estimación final de la densidad, puesto que la eficiencia de $f_n^P(\cdot)$ depende principalmente de la elección de la ventana s .

Los métodos de remuestreo bootstrap que parecen proporcionar ventanas más próximas a las ventanas $MISE$ son los remuestreos SP y $SC2$. Ambos tienen en común, a diferencia de los cuatro métodos restantes, la forma de remuestrear los valores δ_i^* . Estos valores se generan de modo que la distribución de δ_i^* condicionada $Z_i^* = t$ sea una Bernoulli de parámetro $p_n(t)$, siendo $p_n(\cdot)$ el estimador de Nadaraya-Watson. En los demás métodos de remuestreo los valores δ^* se generan como $\delta^* = \mathbf{1}_{\{Y^* \leq C^*\}}$. Es llamativo el buen comportamiento del remuestreo SP al no ser un método suavizado. Aunque, en principio, este método no debería resultar adecuado para imitar el sesgo, posiblemente la existencia de presuavización en el remuestreo mitigue este efecto.

Se ha aproximado $MISE(s, b)$, el error cuadrático medio cometido al estimar $f(\cdot)$ mediante $f_n^P(\cdot)$ con la ventana de suavizado s y de presuavizado b , mediante la media muestral de $ISE(f_n^P(\cdot))$ calculado con $m = 500$ muestras:

$$MISE(s, b) \simeq \frac{1}{m} \sum_{i=1}^m ISE(s, b)$$

Para cada par de ventanas bootstrap $(s_i^*, b_i^*)_{i=1}^m$ se ha calculado el cociente

$$\psi_i = \frac{MISE(s_i^*, b_i^*)}{MISE(s_{MISE}, b_{MISE})} \quad i = 1, \dots, m$$

donde (s_{MISE}, b_{MISE}) son las ventanas que minimizan $MISE(f_n^P(\cdot))$. Estos valores se han usado para aproximar la función de densidad de la variable aleatoria

$$\psi = \frac{MISE(s^*, b^*)}{MISE(s_{MISE}, b_{MISE})} = \phi(s^*, b^*) + 1.$$

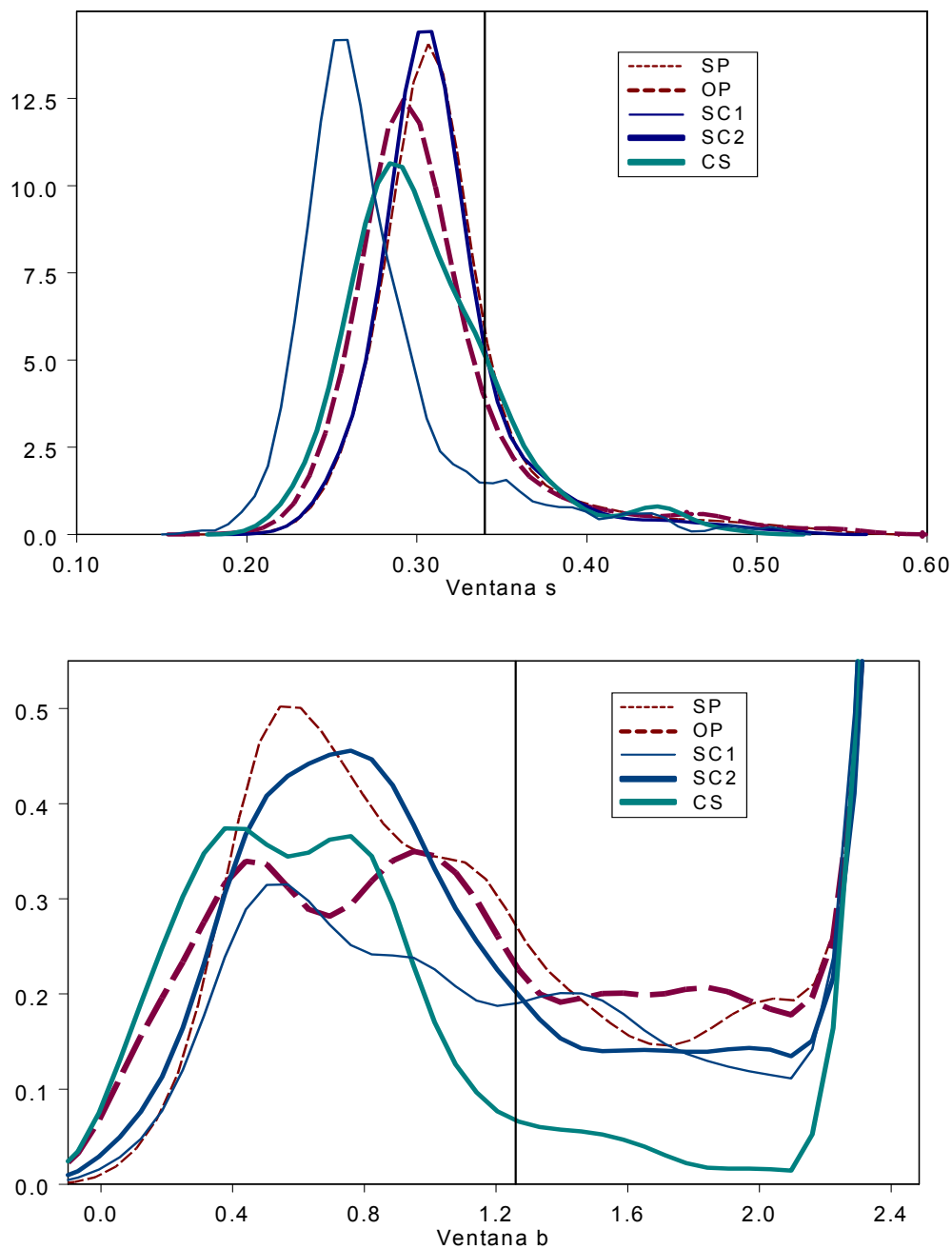


Figura 4.9. Estimación de la función de densidad de las ventanas bootstrap s^* y b^* para el modelo 1. La línea vertical indica el valor de la correspondiente ventana $MISE$.

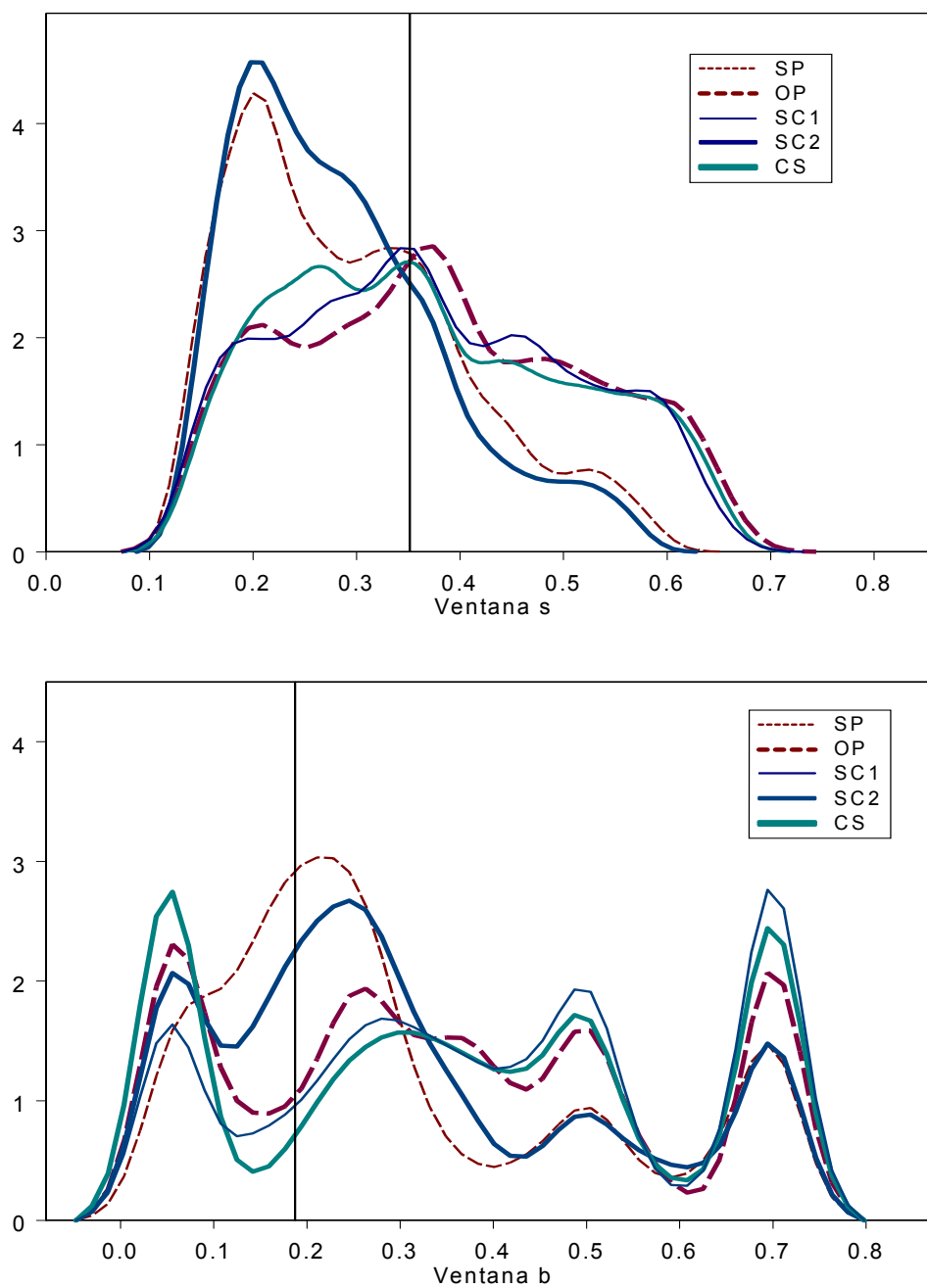


Figura 4.10. Estimación de la función de densidad de las ventanas bootstrap s^* y b^* para el modelo 2. La línea vertical indica el valor de la correspondiente ventana $MISE$.

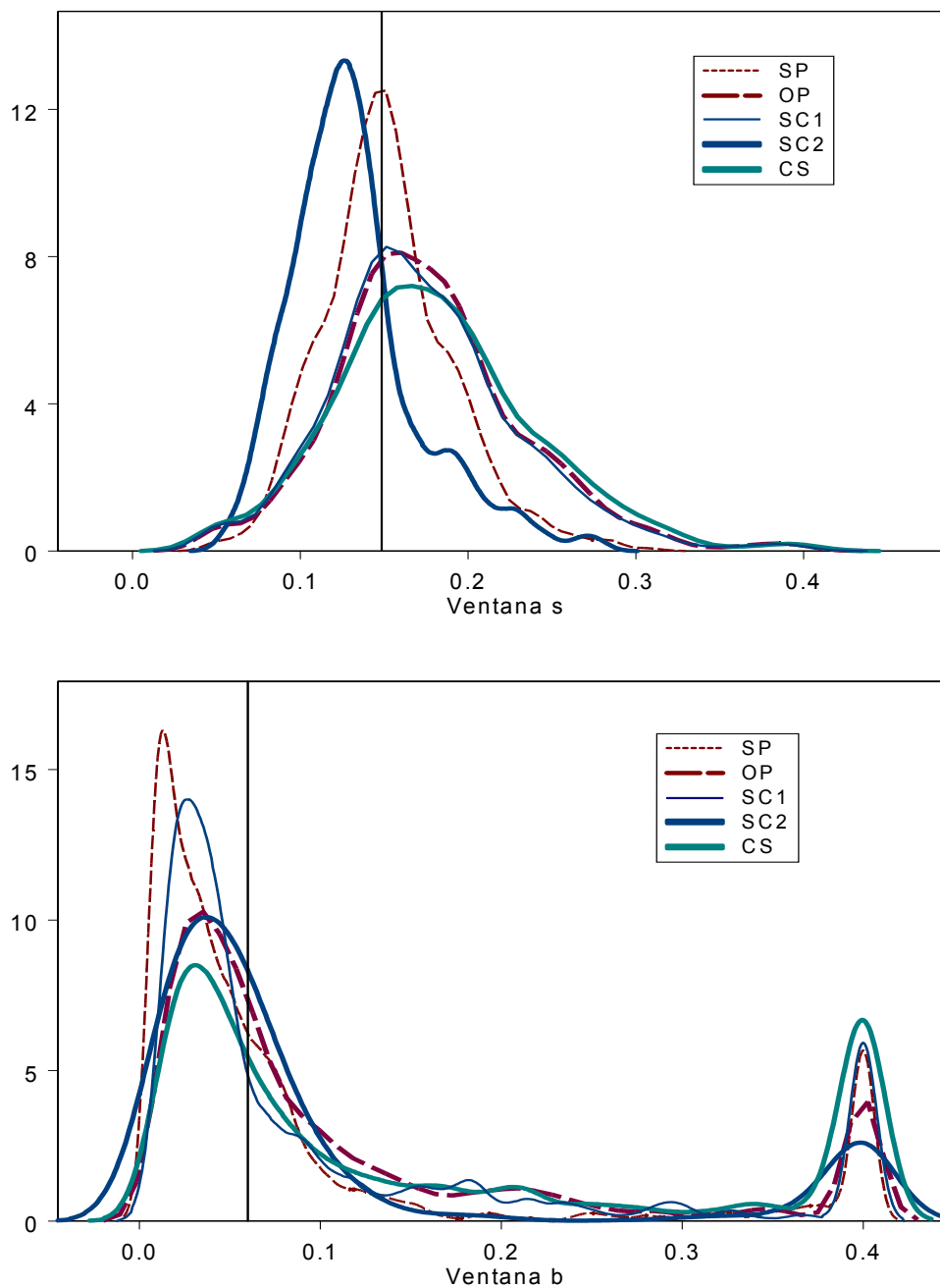


Figura 4.11. Estimación de la función de densidad de las ventanas bootstrap s^* y b^* para el modelo 3. La línea vertical indica el valor de la correspondiente ventana *MISE*.

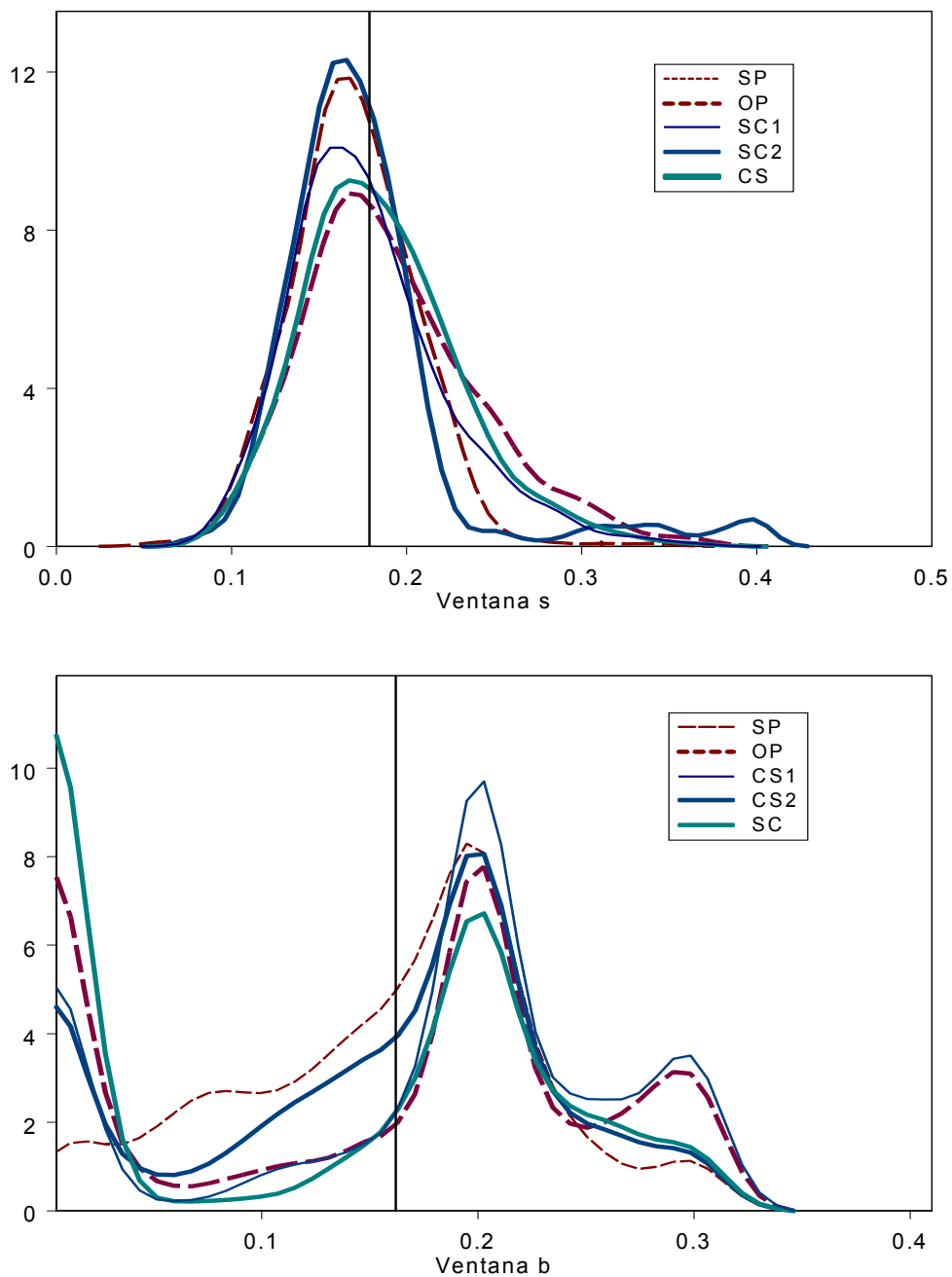


Figura 4.12. Estimación de la función de densidad de las ventanas bootstrap s^* y b^* para el modelo 4. La línea vertical indica el valor de la correspondiente ventana $MISE$.

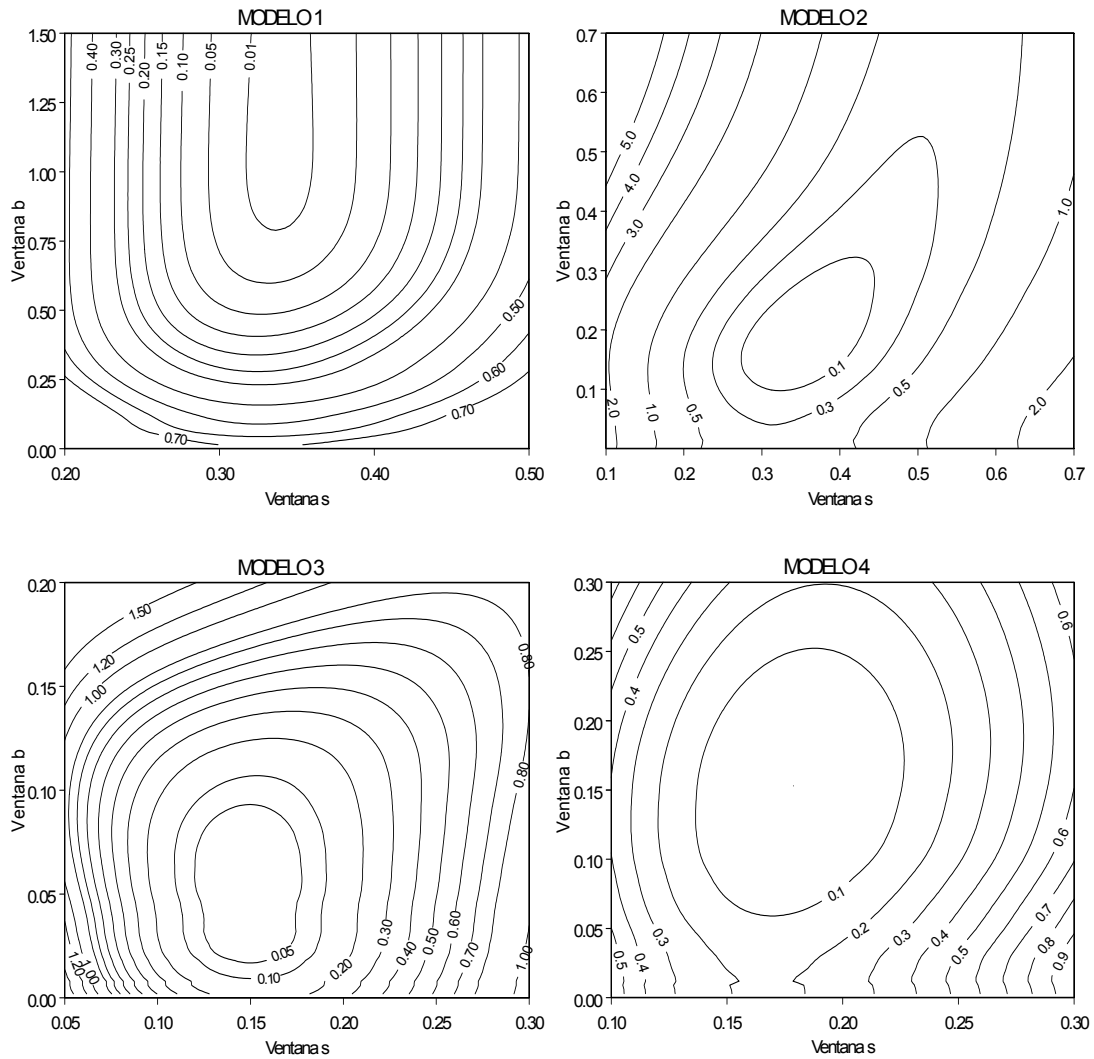


Figura 4.13. Función $\phi(\cdot, \cdot)$ para los modelos 1-4.

Asimismo, también se calcula la función

$$\phi(s, b) = \frac{MISE(s, b) - MISE(s_{MISE}, b_{MISE})}{MISE(s_{MISE}, b_{MISE})},$$

que representa la pérdida relativa de eficiencia en la estimación presuavizada de la densidad, en términos de $MISE(f_n^P(\cdot))$, cuando se usa el par de ventanas (s, b) en lugar de las ventanas óptimas (s_{MISE}, b_{MISE}) .

Para comprobar cómo afecta la distribución de las ventanas bootstrap (ver figuras 4.9-4.12) al valor de $MISE(f_n^P(\cdot))$ se ha realizado un gráfico de contorno de la función $\phi(\cdot, \cdot)$ con su valor en una rejilla de ventanas (s, b) . En la figura 4.13 se puede

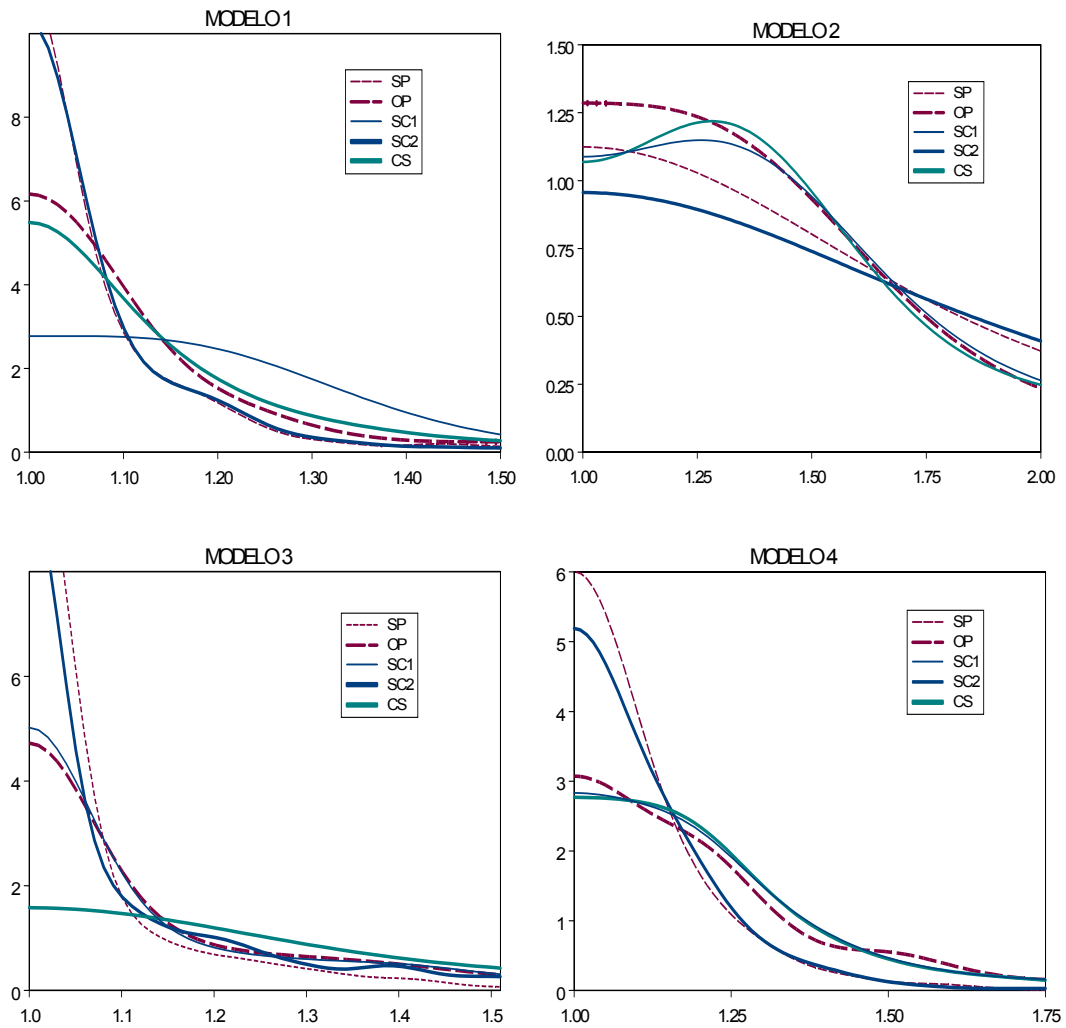


Figura 4.14. Estimación no paramétrica de la función de densidad de ψ con los cinco remuestreos distintos para los modelos 1-4.

comprobar que la región que representa el soporte de las funciones de densidad de s^* y b^* para cada modelo se corresponden con valores de $\phi(\cdot, \cdot)$ no demasiado altos.

La figura 4.14 muestra las estimaciones de Parzen-Rosenblatt de la función de densidad de la variable ψ calculadas con ventana de Sheather-Jones con los $m = 500$ valores obtenidos al evaluar $\phi(\cdot, \cdot)$ en los m pares de ventanas bootstrap $(s_i^*, b_i^*)_{i=1}^{500}$, obtenidos a su vez con los cinco remuestreos distintos propuestos. Puesto que la variable ψ está acotada inferiormente por 1, el estimador clásico de Parzen-Rosenblatt no es consistente en un entorno del punto $x = 1$. Para asegurar que la estimación no paramétrica de la función de densidad de ψ tome el valor cero en $x < 1$ hemos usado

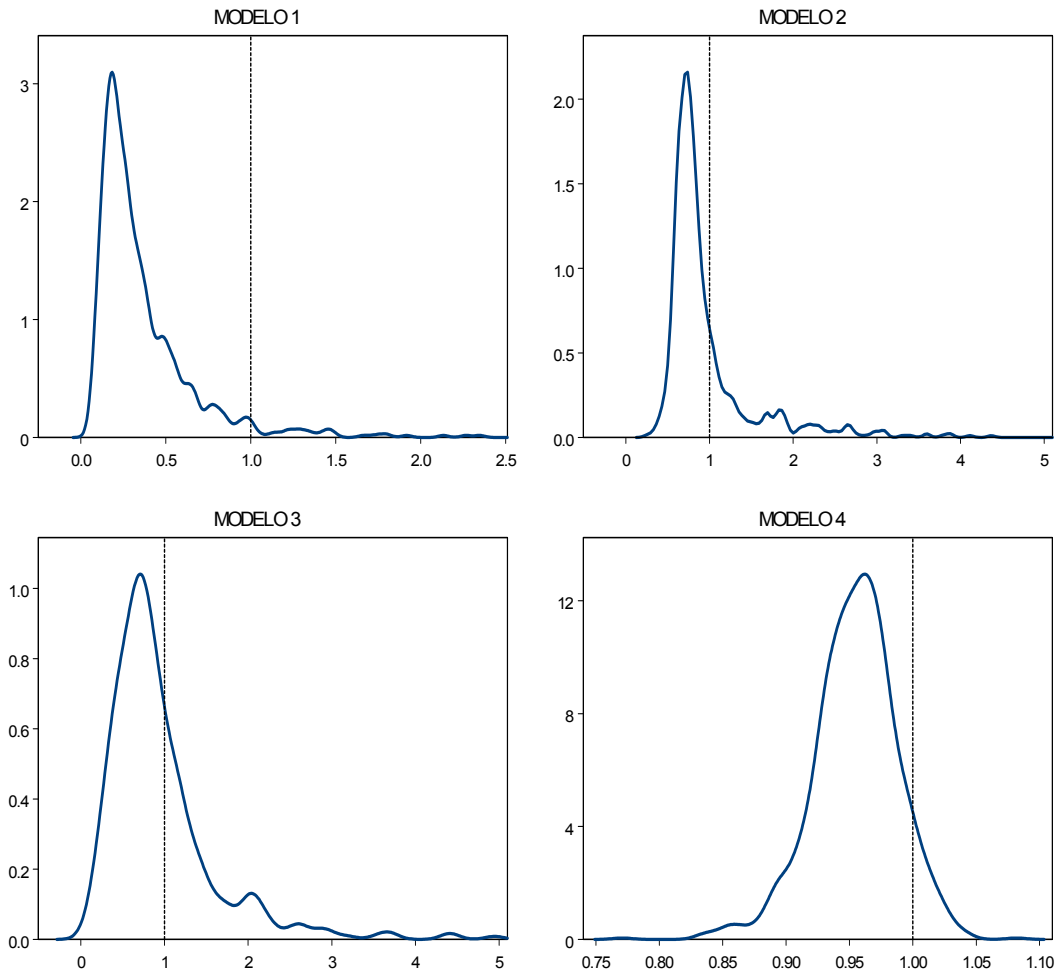


Figura 4.15. Estimaciones de Parzen-Rosenblatt de las funciones de densidad de la eficiencia relativa $RE_{ISE}^{KM,P}(\cdot, \cdot, \cdot)$ para $m = 500$ muestras de tamaño $n = 100$ de los modelos 1-4. El remuestreo considerado ha sido el *SC2*.

para estimarla una modificación del método tipo núcleo clásico llamado método de reflexión (ver, entre otros, Schuster (1985) y Silverman (1986)).

Puede verse que, para el modelo 2, la eficiencia de los cinco remuestreos bootstrap propuestos son similares, debido a la alta variabilidad de las ventanas de suavizado bootstrap obtenidas para ese modelo. Quizás puede sorprender que el mejor método de remuestreo en este modelo sea un método no suavizado, el *OP*. A pesar de que la eficiencia de este método es parecida a la de los demás, el hecho de presuavizar puede mitigar, en principio, la mala imitación del sesgo. Para los demás modelos, los mejores métodos de remuestreo bootstrap parecen ser el suavizado *SC2* y el no suavizado *SP*.

Tabla 4.4 Eficiencia relativa del estimador con pesos de Kaplan-Meier $f_n^{KM}(\cdot)$ (con selector de ventana de tipo plug-in) con respecto al estimador presuavizado $f_n^P(\cdot)$ (con selector de ventana de tipo bootstrap, mediante el remuestreo SC2).

MODELO	$n = 50$	$n = 100$	$n = 200$
1	0.367	0.304	0.291
2	0.861	0.859	0.815
3	1.126	0.920	0.622
4	0.941	0.955	0.990

Al igual que se había hecho para el selector de tipo plug-in, se ha comparado el estimador presuavizado con selector de ventana bootstrap con el estimador de Kaplan-Meier. De los cinco remuestreos bootstrap estudiados, hemos seleccionado para este fin el remuestreo *SC2*. Respecto al estimador de Kaplan-Meier, hemos considerado el mismo selector que el usado en el estudio de simulación de la sección anterior, es decir, el selector de tipo plug-in propuesto por Sánchez Sellero *et al.* (1999). Con cada una de las $m = 500$ muestras se ha calculado la eficiencia relativa del estimador de Kaplan-Meier, $f_n^{KM}(\cdot)$, con respecto al estimador presuavizado, $f_n^P(\cdot)$, en términos del *ISE* cometido, es decir,

$$RE_{ISE}^{KM,P}(\hat{s}^P, \hat{b}^P; \hat{s}^{KM}) = \frac{ISE(f_{n, \hat{s}^P, \hat{b}^P}^P(\cdot))}{ISE(f_{n, \hat{s}^{KM}}^{KM}(\cdot))}.$$

Valores de $RE_{ISE}^{KM,P}(\cdot, \cdot, \cdot)$ menores que 1 indican que el estimador presuavizado con selector de ventana de tipo bootstrap (y remuestreo *SC2*) comete un menor error *ISE* que el estimador con pesos de Kaplan-Meier con selector de la ventana de tipo plug-in. Las estimaciones de Parzen-Rosenblatt de la función de densidad de $RE_{ISE}^{KM,P}(\cdot, \cdot, \cdot)$, con núcleo gaussiano y selector de ventana de Sheather-Jones, puede verse en la figura 4.15. De forma análoga a lo que ocurría con el selector de tipo plug-in, el estimador presuavizado es, en general, más eficiente en términos de *ISE* que el estimador $f_n^{KM}(\cdot)$.

En la tabla 4.4 figura, para cada modelo y para cada tamaño muestral, la eficiencia relativa del estimador de Kaplan-Meier, $f_n^{KM}(\cdot)$, con selector de ventana de tipo plug-in, con respecto al estimador presuavizado, $f_n^P(\cdot)$, calculado con el selector de ventana de tipo bootstrap:

$$RE^{KM,P}(\hat{s}^P, \hat{b}^P; \hat{s}^{KM}) = \frac{MISE(f_{n, \hat{s}^P, \hat{b}^P}^P(\cdot))}{MISE(f_{n, \hat{s}^{KM}}^{KM}(\cdot))}. \quad (4.77)$$

Para todos los modelos y todos los tamaños muestrales considerados, la eficiencia relativa del estimador de Kaplan-Meier, $f_n^{KM}(\cdot)$, con selector de ventana de tipo plug-in, respecto al estimador presuavizado, $f_n^P(\cdot)$, con selector de tipo bootstrap, es menor que respecto al mismo estimador con selector de tipo plug-in (ver tablas 4.2 y 4.4). Es decir, siempre es mayor la eficiencia relativa en términos de *MISE* del estimador presuavizado con selector de la ventana de tipo bootstrap que con uno de tipo plug-in.

Esta mejora se manifiesta claramente en los modelos 1 y 3. Para el primero de ellos se pasa de tener una eficiencia relativa, para el estimador presuavizado, de un 150% para $n = 50$ con selector de tipo plug-in a, aproximadamente, un 270% para $n = 50$ con selector de tipo bootstrap. En relación al modelo 3, la eficiencia relativa (4.77) pasa de ser mayor que 1 (mejor comportamiento, en términos de *MISE*, del estimador de Kaplan-Meier) cuando el selector es de tipo plug-in, a ser menor que 1 para los tamaños muestrales $n = 100$ y $n = 200$ cuando se considera el selector de tipo bootstrap.

El estudio en profundidad de estos selectores de las ventanas tipo bootstrap será objeto de nuestra investigación futura.

Capítulo 5

Comparación de los estimadores de Nadaraya-Watson y local lineal en la estimación presuavizada de la densidad

5.1. Los estimadores de Nadaraya-Watson y local lineal de la regresión

Dado que la función $p(t) = P(\delta = 1|Z = t) = \mathbb{E}(\delta|Z = t)$ es la función de regresión de la variable δ condicionada a $Z = t$, el método clásico de estimación no paramétrica de $p(\cdot)$ es el estimador tipo núcleo de Nadaraya-Watson (ver Nadaraya (1964), Watson (1964) y Härdle (1990) para más detalles):

$$p_n^{NW}(t) = \frac{\frac{1}{n} \sum_{i=1}^n K_b(t - Z_i) \delta_i}{\frac{1}{n} \sum_{i=1}^n K_b(t - Z_i)}.$$

Este estimador aproxima localmente $p(\cdot)$ por una constante a , es decir, es el estimador mínimo-cuadrático local ponderado, puesto que en cada punto t minimiza la suma de los residuos al cuadrado ponderados localmente según los pesos $K_b(t - Z_i)$:

$$\arg \min_a \sum_{i=1}^n K_b(t - Z_i) (\delta_i - a)^2 = \hat{a} = \frac{\sum_{i=1}^n K_b(t - Z_i) \delta_i}{\sum_{i=1}^n K_b(t - Z_i)} = p_n^{NW}(t).$$

De lo anterior se deduce que $p_n^{NW}(\cdot)$ es el polinomio de grado 0 que mejor ajusta los datos en el sentido de minimizar la suma de los residuos al cuadrado en un entorno de centro t determinado en forma y diámetro por $K_b(\cdot) = \frac{1}{b}K(\frac{\cdot}{b})$.

Sin embargo, en los últimos años se ha dedicado mucha más atención a los estimadores polinómico locales (ver Fan (1992, 1993), Fan y Gijbels (1992) y Ruppert y Wand (1994) entre otros). Dentro de ellos se incluyen, como caso particular, los de tipo núcleo y los local lineales, que consisten en estimar la función de regresión ajustando localmente una recta por mínimos cuadrados ponderados. La idea es la siguiente: supongamos que existe la primera derivada de $p(\cdot)$. En un entorno de un punto $t > 0$, y mediante un desarrollo de Taylor, se tiene que

$$p(x) \simeq p(t) + p'(t)(x - t) = a + b(x - t).$$

Esto sugiere la estimación de $p(\cdot)$ encontrando los parámetros a y b que minimizan la suma de cuadrados de residuos ponderados de una regresión lineal local ponderada (un polinomio local de grado $p = 1$):

$$\sum_{i=1}^n K_b(t - Z_i) [\delta_i - (a + b(Z_i - t))]^2.$$

Así se obtiene el estimador local lineal de $p(\cdot)$:

$$p_n^{LL}(t) = \hat{a} = \left(\sum_{i=1}^n \omega_i \right)^{-1} \sum_{i=1}^n \delta_i \omega_i$$

siendo los pesos iguales a

$$\omega_i = K_b(Z_i - t) (S_{n,2} - (Z_i - t) S_{n,1}) \quad \text{con} \quad S_{n,j} = \sum_{i=1}^n K_b(Z_i - t) (Z_i - t)^j.$$

En consecuencia, ambos métodos pueden pensarse como dos casos particulares de un suavizador local polinómico de grados $p = 0$ y $p = 1$ respectivamente.

Se ha probado ampliamente en la literatura que el estimador polinómico local, y en particular el estimador local lineal, presenta muchas ventajas con respecto al estimador de Nadaraya-Watson (ver Fan (1992, 1993), Ruppert y Wang (1994) y Fan y Gijbels (1996), entre otros). Este estimador es independiente del diseño del modelo (fijo o aleatorio, altamente agrupado o casi uniforme), es eficiente en el sentido minimax, y tiene muy buen comportamiento en los puntos frontera. Además, presenta menor sesgo, en valor absoluto, que el estimador de Nadaraya-Watson en los puntos en los que $p'(\cdot)h'(\cdot)$ tiene el mismo signo que $p''(\cdot)$, como puede verse en la tabla 5.1. De hecho, cuanto mayor sea el orden del polinomio local, mayor es la reducción del sesgo que se

obtiene. Sin embargo, la pregunta a contestar es: ¿hereda la estimación presuavizada de la densidad, con estimación local lineal de la función $p(\cdot)$, todas estas ventajas con respecto al ajuste de $p(\cdot)$ con el estimador de Nadaraya-Watson? Esta cuestión se abordará en el resto del capítulo.

5.2. Estimación presuavizada de la densidad con los diferentes ajustes de p

Para explorar los beneficios de la estimación local lineal de $p(\cdot)$ en la presuavización, nos centraremos en la representación i.i.d. de $f_n^P(\cdot)$ dada en el capítulo 3. Puesto que la varianza asintótica de ambos estimadores de la regresión es la misma (ver tabla 5.1), la ventaja de estimar $p(\cdot)$ mediante un ajuste local lineal viene dada por el sesgo de $f_n^P(\cdot)$, y en particular, por el de su parte dominante $\overline{f_n^P}(\cdot)$.

Propiedad 5.2.1

Sean $B_n^{NW}(\cdot)$ y $B_n^{LL}(\cdot)$ las expresiones asintóticas del sesgo del estimador de Nadaraya-Watson y local lineal de la función $p(\cdot)$ respectivamente, dadas en la tabla 5.1, y sean $f_n^{P,NW}(\cdot)$ y $f_n^{P,LL}(\cdot)$ las estimaciones presuavizadas de la función de densidad con los diferentes ajustes de $p(\cdot)$. Entonces,

$$\mathbb{E} \left[\overline{f_n^{P,NW}}(t) - f(t) \right] = \frac{1}{2} s^2 f''(t) d_K + \left[(1 - F(t)) \int_0^t \frac{h(v) B_n^{NW}(v)}{1 - H(v)} dv \right]' \quad (5.1)$$

$$\mathbb{E} \left[\overline{f_n^{P,LL}}(t) - f(t) \right] = \frac{1}{2} s^2 f''(t) d_K + \left[(1 - F(t)) \int_0^t \frac{h(v) B_n^{LL}(v)}{1 - H(v)} dv \right]' \quad (5.2)$$

Demostración.

En el teorema 3.2.1 se probó la representación casi segura y en media del estimador presuavizado $f_n^P(\cdot)$:

$$f_n^P(t) = f(t) + \beta_n(t) + \sigma_n(t) + e_n(t),$$

con

$$\beta_n(t) = \int f(t - sv) K(v) dv - f(t), \quad (5.3)$$

$$\sigma_n(t) = \frac{1}{ns} \sum_{i=1}^n \int \varepsilon_i(t - sv) K'(v) dv,$$

$$\varepsilon_1(t) = (1 - F(t)) (g_1(t, Z_1) - g_2(t, Z_1) + g_3(t, Z_1, \delta_1))$$

y donde las funciones $g_1(\cdot, \cdot)$, $g_2(\cdot, \cdot)$ y $g_3(\cdot, \cdot, \cdot)$ han sido definidas en (2.12)-(2.14).

La expresión del estimador usado para la función $p(\cdot)$, el estimador de Nadaraya-Watson, está, en cierto modo, integrada dentro de la función $g_3(\cdot, Z, \delta)$. Para verlo más claramente, reescribimos la expresión de $g_3(\cdot, Z, \delta)$ de la siguiente manera:

$$\begin{aligned}
& \frac{1}{n} \sum_{i=1}^n g_3(t, Z_i, \delta_i) \\
&= \frac{1}{n} \sum_{i=1}^n \int_0^t K_b(v - Z_i) \frac{\delta_i - p(v)}{1 - H(v)} dv \\
&= \int_0^t \frac{h_n(v)}{1 - H(v)} \left(\frac{\frac{1}{n} \sum_{i=1}^n K_b(v - Z_i) \delta_i}{h_n(v)} - p(v) \right) dv \\
&= \int_0^t \frac{h_n(v)}{1 - H(v)} (p_n^{NW}(v) - p(v)) dv \\
&= \int_0^t \frac{h(v)}{1 - H(v)} (p_n^{NW}(v) - p(v)) dv - \int_0^t \frac{h_n(v) - h(v)}{1 - H(v)} (p_n^{NW}(v) - p(v)) dv.
\end{aligned} \tag{5.4}$$

El segundo sumando es despreciable. De esta forma, la representación del estimador presuavizado de $f(\cdot)$ con un ajuste $\hat{p}_n(\cdot)$ de la función $p(\cdot)$ es igual a

$$\begin{aligned}
f_n^P(t) &= f(t) + \beta_n(t) + \frac{1}{ns} \sum_{i=1}^n \int \varepsilon_i^{\hat{p}_n}(t - sv) K'(v) dv \\
&\quad + \frac{1}{s} \int \gamma_3^{\hat{p}_n}(t - sv) K'(v) dv + e_n^{\hat{p}_n}(t),
\end{aligned}$$

donde $\beta_n(\cdot)$ es la función dada en (5.3),

$$\begin{aligned}
\varepsilon_1^{\hat{p}_n}(t) &= (1 - F(t)) (g_1(t, Z_1) - g_2(t, Z_1)), \\
\gamma_3^{\hat{p}_n}(t) &= (1 - F(t)) \int_0^t \frac{h(u)}{1 - H(u)} (\hat{p}_n(u) - p(u)) du.
\end{aligned}$$

y $e_n^{\hat{p}_n}(\cdot)$ es un nuevo término de error que incorpora el último sumando en (5.4).

Puesto que $\mathbb{E}[g_1(t, Z)] = \mathbb{E}[g_2(t, Z)] = 0$, con cálculos análogos a los usados en la demostración del teorema 3.2.2 se obtiene (5.1) y (5.2). ■

Las varianzas asintóticas de $p_n^{NW}(\cdot)$ y $p_n^{LL}(\cdot)$ coinciden puesto que

$$\begin{aligned}
& Var(p_n^{LL}(\cdot)) = Var(p_n^{NW}(\cdot)) \\
& + \frac{b}{n} \left(\frac{h'(t)^2}{h^3(t)} \sigma^2(t) (2c_K d_K + m_K) - 2 \frac{h'(t) (\sigma^2 h)'(t)}{h^3(t)} m_K \right) + o\left(\frac{b}{n}\right),
\end{aligned}$$

Tabla 5.1. Expresiones puntuales del sesgo y varianza asintóticos de ambos estimadores de la función de regresión.

Método	Sesgo	Varianza
Nadaraya-Watson	$\frac{1}{2} \left(p''(t) + \frac{2p'(t)h'(t)}{h(t)} \right) d_K b^2$	$\frac{1}{nb} \frac{\sigma^2(t)}{h(t)} c_K$
Local Lineal	$\frac{1}{2} p''(t) d_K b^2$	$\frac{1}{nb} \frac{\sigma^2(t)}{h(t)} c_K$

con $\sigma^2(t) = Var(\delta|Z=t) = \frac{p(t)}{f_n^{P,NW}(\cdot)} (1-p(t))$. En consecuencia, no se esperan diferencias entre las varianzas de $f_n^{P,NW}(\cdot)$ y $f_n^{P,LL}(\cdot)$, excepto en términos despreciables.

Por tanto, el beneficio de la presuavización mediante el ajuste de $p(\cdot)$ por un estimador o por otro viene dado por el término dominante del sesgo $B_n(\cdot)$. Cuando $p'(\cdot)h'(\cdot)$ tiene el mismo signo que $p''(\cdot)$, el sesgo del estimador local lineal es menor, en valor absoluto, que el del estimador de Nadaraya-Watson, y por tanto el estimador presuavizado de la densidad con ajuste local lineal es más eficiente que aquel con ajuste de Nadaraya-Watson, como comprobaremos en las simulaciones de la última sección de este capítulo. Sin embargo, ¿qué ocurre en aquellas situaciones con un diseño del modelo casi uniforme, en el que $|h'(\cdot)|$ es muy pequeño, y por tanto los sesgos $B_n^{NW}(\cdot)$ y $B_n^{LL}(\cdot)$ son prácticamente iguales, así como los términos de tercer orden de las varianzas? Se esperaría un comportamiento muy similar en el estimador presuavizado de la función de densidad con ambos ajustes, sin embargo algunas simulaciones lo contradicen.

En resumen, los estimadores polinómico locales, y en particular el estimador local lineal, presentan muchas ventajas con respecto al estimador de Nadaraya-Watson. No obstante, esta eficiencia en el ajuste de $p(\cdot)$ no implica una mayor eficiencia en la estimación presuavizada de $f(\cdot)$. De hecho, puede haber situaciones en las que es preferible presuavizar estimando $p(\cdot)$ con el estimador de Nadaraya-Watson en vez de usar un ajuste local lineal.

5.2.1. Estimación local logística de p

Una característica importante de la función $p(\cdot)$ es que se trata de una función de probabilidad condicional, y por tanto únicamente toma valores entre 0 y 1. Esta propiedad la verifica también el estimador de Nadaraya-Watson $p_n^{NW}(\cdot)$, pero no el estimador local lineal $p_n^{LL}(\cdot)$ que puede fácilmente tomar valores fuera de ese intervalo.

Un ajuste de la función $p(\cdot)$ que mantenga la flexibilidad del estimador local lineal, es decir, que no estime $p(\cdot)$ localmente como una constante tal como hace el estimador de Nadaraya-Watson, pero que siempre tome valores entre 0 y 1 puede ser el estimador local logístico. Para estimar la función $p(\cdot)$ en un punto $t \geq 0$, se ajusta localmente una función logística en un entorno de t , cuya amplitud dependerá de la ventana b .

La función de regresión logística es el ajuste más común para datos binarios. Consiste en ajustar la función *logit* de la probabilidad de la variable respuesta como una función lineal de la forma

$$\text{logit}[p(t)] = \log \left[\frac{p(t)}{1-p(t)} \right] = a + bt,$$

es decir,

$$p(t) = \frac{\exp(a + bt)}{1 + \exp(a + bt)}.$$

El ajuste logístico local consiste entonces en minimizar la siguiente función objetivo

$$\sum_{i=1}^n l(Z_i, \delta_i, g) K \left(\frac{Z_i - t}{b} \right), \quad (5.5)$$

donde $g(t) = \exp(a + bt) [1 + \exp(a + bt)]^{-1}$, y el funcional $l(\cdot, \cdot, \cdot)$ puede ser

$$l(z, d, g) = -\log \left[g(z)^d (1 - g(z))^{1-d} \right],$$

que da lugar al estimador local logístico máximo-verosímil $p_n^{LLog-MV}(\cdot)$, o

$$l(z, d, g) = (d - g(z))^2,$$

que corresponde con el ajuste mínimo cuadrático $p_n^{LLog-MC}(\cdot)$.

El método estándar para resolver de forma global estas ecuaciones es el procedimiento *local scoring*, es decir, el algoritmo de Newton-Raphson usando la matriz de información esperada en vez de la observada. Sin embargo, este algoritmo no se puede usar en el ajuste local. En este caso, es necesario calcular para cada punto $t > 0$ de estimación la función objetivo dada en (5.5) y calcular los parámetros a y b que la minimizan. Puesto que las ecuaciones no son lineales en los parámetros, las soluciones locales tienen que ser aproximadas de forma iterativa.

En el estudio de simulación de la siguiente sección se han comparado los ajustes locales de $p(\cdot)$ por una constante (estimador de Nadaraya-Watson), por una recta (estimador local lineal) y por la función logística (estimador local logístico) en cuatro modelos diferentes. Veremos cómo los resultados no son siempre los esperados.

5.3. Estudio de simulación

En esta sección hemos realizado un estudio de simulación con cuatro modelos diferentes, de acuerdo a diferentes formas de la función $p(\cdot)$, para comparar el comportamiento en la práctica de los estimadores de la función de densidad con pesos presuavizados estimando la función $p(\cdot)$ mediante el estimador de Nadaraya-Watson, $f_n^{P,NW}(\cdot)$, mediante el estimador local lineal, $f_n^{P,LL}(\cdot)$, o bien a través del ajuste local logístico, $f_n^{P,LLog}(\cdot)$. A su vez, hemos comparado estos estimadores con el estimador clásico sin presuavización, es decir, el estimador con pesos de Kaplan-Meier, $f_n^{KM}(\cdot)$.

Los modelos 1-3 son los usados en las simulaciones de los capítulos anteriores, pero aquí se ha considerado también un cuarto modelo. Mientras que en los modelos 1 y 2, la variable observable sigue la distribución $Z \stackrel{d}{=} U[0, 1]$, en los modelos 3 y 4 se tiene $Z \stackrel{d}{=} W(1.75, 5)$ y $Z \stackrel{d}{=} W(2, 4)$ respectivamente, donde $W(a, b)$ denota la distribución de Weibull con parámetro de escala a y parámetro de forma b , cuya densidad es $f_{W(a,b)}(t) = a^b b t^{b-1} \exp[-(at)^b]$.

En la tabla 5.2 se muestran las funciones de densidad teóricas a estimar $f(\cdot)$ y las probabilidades condicionales de no censura $p(\cdot)$ para cada uno de los modelos. De nuevo, las funciones $p(\cdot)$ se han elegido de manera que se ajusten a cuatro situaciones diferentes entre sí. En el modelo 1 la función $p(\cdot)$ es prácticamente constante, en el modelo 2 es unimodal, en el modelo 3 es monótona decreciente, y en el modelo 4 presenta un mínimo y un máximo local (ver figura 5.1).

En la figura 5.1 se representan la función de densidad $f(\cdot)$ de la variable de interés (línea con puntos), la función de densidad de la variable observable $h(\cdot)$ (línea continua gruesa) y la función de probabilidad condicional de no censura $p(\cdot)$ (línea continua fina) para cada uno de los modelos.

Dado que en los modelos 1 y 2 se tiene que $h(\cdot) = \mathbf{1}_{[0,1]}$, y por tanto $h'(t) = 0$ para todo $t \in [0, 1]$, en ellos las expresiones asintóticas del sesgo de los estimadores de Nadaraya-Watson y local lineal son iguales, es decir, $B_n^{NW}(t) = B_n^{LL}(t)$ para todo $t \in [0, 1]$. Puesto que las expresiones asintóticas de las varianzas son siempre iguales $V_n^{NW}(\cdot) = V_n^{LL}(\cdot)$, se esperaría muy poca diferencia entre ambos ajustes en la estimación presuavizada de la función de densidad. Sin embargo, aunque el estimador local lineal de $p(\cdot)$ presenta asintóticamente buenas propiedades con respecto al estimador de Nadaraya-Watson, su comportamiento con tamaños muestrales pequeños es peor.

Para comparar las estimaciones presuavizadas con los distintos ajustes de $p(\cdot)$ se ha aproximado, por simulación, el error cuadrático medio integrado ponderado:

$$MISE(f_n^P(\cdot)) = \mathbb{E} \left[\int (f_n^P(v) - f(v))^2 \omega(v) dv \right].$$

Tabla 5.2. Funciones de densidad, $f(\cdot)$, y probabilidad condicional de no censura, $p(\cdot)$, en los modelos 1-4.

Función de densidad	Función $p(t)$
$f(t) = (1-t)^{3/8} e^{-1/4t^2} \left[\frac{3}{8} (1-t)^{-1} + \frac{1}{2}t \right]$	$p(t) = \frac{1}{2} \left[1 - \left(t - \frac{1}{2}\right)^2 \right]$
$f(t) = 3te^{-3/2t^2}$	$p(t) = \left[\frac{3}{4} - 3 \left(t - \frac{1}{2}\right)^2 \right] \mathbf{1}_{\{0 \leq t \leq 1\}}$
$f(t) = \frac{5}{2}t^4 e^{-t^5} + 80t^4 e^{-32t^5}$	$p(t) = \frac{\frac{5}{2}e^{-t^5} + 80e^{-32t^5}}{41.0327(e^{-t^5} + e^{-32t^5})}$
$f(t) = 120t^4 e^{-32t^5} + 2.25t^8 e^{-t^9}$	$p(t) = \frac{7.5te^{-32t^5} + 0.14063t^5 e^{-t^9}}{3e^{-32t^5} + e^{-t^9}}$

La función de peso $\omega(\cdot)$ es la función indicadora de los intervalos $[0, 0.9]$ para el modelo 1, $[0, 0.95]$ para el modelo 2, $[0.2, 0.75]$ para el modelo 3 y $[0, 0.75]$ para el modelo 4, con el propósito de evitar problemas de definición en la cola derecha de los intervalos de integración. Los extremos de dichos intervalos vienen dados por las líneas verticales en la figura 5.1. Estos intervalos han sido elegidos de manera que la función de peso $\omega(\cdot)$ desecha un máximo de 5% de los datos observados en la cola superior del intervalo (modelos 2-4), o bien el 10% de los datos para el modelo 1, debido a la asíntota que presenta la función $f(\cdot)$ en $t = 1$ (ver figura 5.1).

Las ventanas que se han usado en la estimación de $f(\cdot)$, con cada uno de los estimadores considerados en este estudio de simulación, han sido las correspondientes ventanas *MISE* óptimas.

Además, puesto que el estimador local lineal $p_n^{LL}(\cdot)$ puede tomar valores fuera del intervalo $[0, 1]$, hemos truncado sus valores de la siguiente manera:

$$\min(\max(0, p_n^{LL}(\cdot)), 1)$$

La tabla 5.3 muestra los valores mínimos de *MISE* para los cuatro estimadores de $f(\cdot)$, obtenidos por Montecarlo con $m = 500$ muestras de tamaños $n = 50, 100$ y 200 , y usando el núcleo de Epanechnikov. Además, se ha calculado la eficiencia relativa de cada estimador de la densidad en comparación con el estimador presuavizado con ajuste de Nadaraya-Watson $f_n^{P,NW}(\cdot)$. Esta eficiencia se define como:

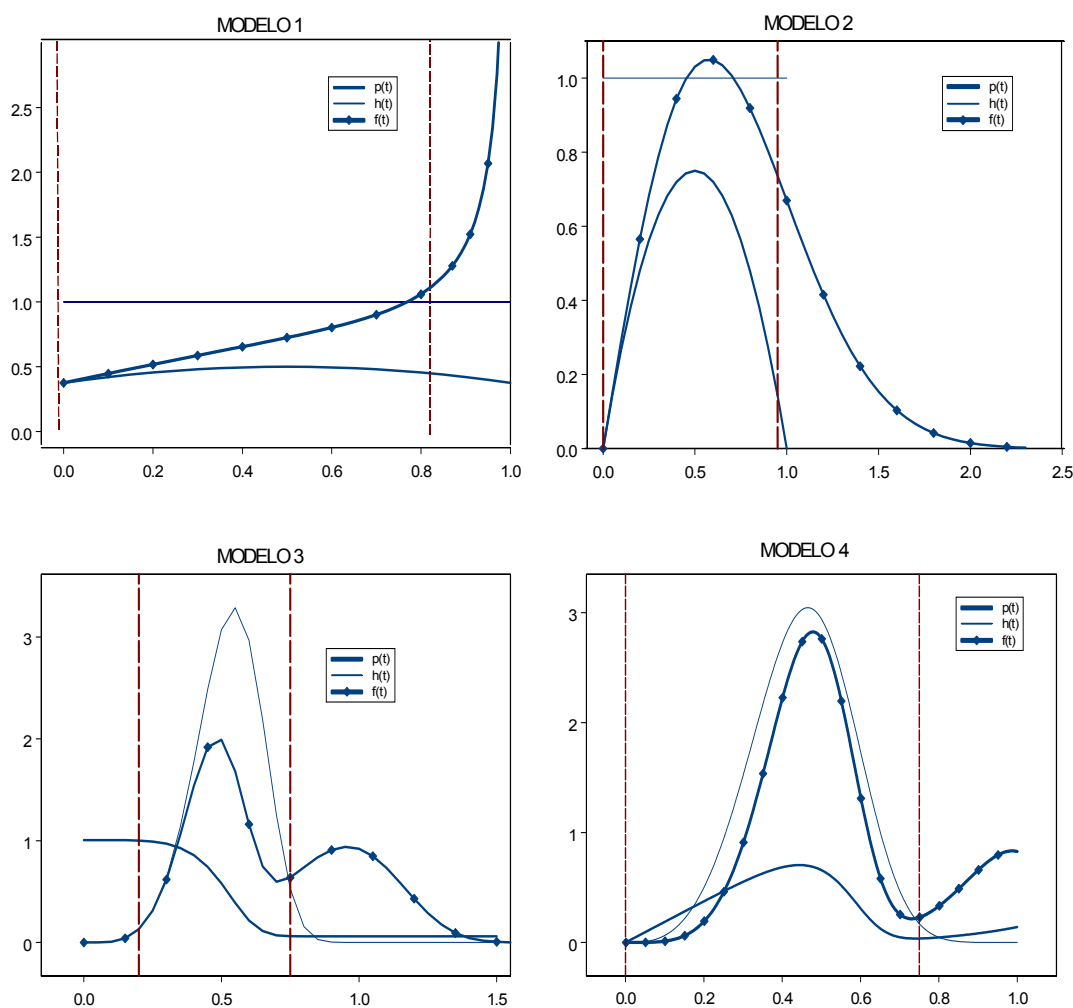


Figura 5.1. Funciones $p(\cdot)$, $h(\cdot)$ y $f(\cdot)$ para los modelos 1-4. La función de peso $\omega(\cdot)$ es la función indicadora del intervalo marcado por las líneas verticales.

$$ER \left[\hat{f}(\cdot) \right] = \frac{MISE_{\min} \left(f_n^{P,NW}(\cdot) \right)}{MISE_{\min} \left(\hat{f}(\cdot) \right)},$$

siendo $\hat{f}(\cdot)$ un estimador genérico de $f(\cdot)$, es decir, el cociente entre el menor valor de $MISE$ cometido con el estimador $f_n^{P,NW}(\cdot)$ y el menor valor de $MISE$ cometido con el estimador $\hat{f}(\cdot)$. Valores de $ER \left[\hat{f}(\cdot) \right]$ por debajo de 1 significan que el estimador de $f(\cdot)$ con ajuste de Nadaraya-Watson es más eficiente que el estimador $\hat{f}(\cdot)$, ambos con el valor óptimo para las respectivas ventanas.

Tabla 5.3. MISE de los 5 estimadores de la función de densidad, y Eficiencia Relativa (%) para los modelos 1-4 con tamaños muestrales $n = 50, n = 100$ y 200 .

MODELO	n	NW	LLineal	LLog_MV	LLog_MC	KM
1	50	0.01846	0.02411 (76.59)	0.02094 (88.16)	0.02087 (88.45)	0.03298 (55.97)
	100	0.01197	0.01484 (80.66)	0.01400 (85.5)	0.01400 (85.5)	0.0198 (60.45)
	200	0.09911	0.01129 (87.79)	0.01091 (90.88)	0.01088 (91.08)	0.01354 (73.2)
2	50	0.02667	0.03297 (80.87)	0.02689 (99.15)	0.02659 (100.29)	0.03373 (79.07)
	100	0.01591	0.01973 (80.66)	0.01557 (102.21)	0.01553 (102.44)	0.02092 (76.06)
	200	0.01116	0.01277 (87.39)	0.01009 (110.55)	0.01011 (110.46)	0.01317 (84.76)
3	50	0.07127	0.05342 (133.41)	0.0492 (144.85)	0.05067 (140.65)	0.0808 (88.21)
	100	0.04649	0.03604 (128.99)	0.03214 (144.65)	0.03258 (142.69)	0.0569 (81.75)
	200	0.03099	0.02928 (105.85)	0.02253 (137.58)	0.02181 (137.58)	0.03216 (96.38)
4	50	0.08951	0.08918 (100.37)	0.08881 (100.79)	0.08884 (100.76)	0.09419 (95.03)
	100	0.05758	0.05544 (103.87)	0.05793 (99.41)	0.05796 (99.35)	0.05984 (96.23)
	200	0.03832	0.03506 (109.31)	0.03943 (97.19)	0.03946 (97.13)	0.03841 (99.79)

Como puede verse en la tabla 5.3, la eficiencia relativa del estimador tipo núcleo con pesos Kaplan-Meier, $f_n^{KM}(\cdot)$, con respecto al estimador presuavizado con ajuste de Nadaraya-Watson, $f_n^{P,NW}(\cdot)$, es mucho menor que uno para los cuatro modelos. Esto coincide con las simulaciones realizadas en capítulos anteriores, y por las cuales se concluyó que siempre es preferible estimar la densidad presuavizando antes que con el estimador de Kaplan-Meier, sobre todo para tamaños muestrales pequeños.

El comportamiento del estimador presuavizado con ajuste local lineal, $f_n^{P,LL}(\cdot)$, es mejor que el del estimador con presuavización de Nadaraya-Watson, $f_n^{P,NW}(\cdot)$, en los modelos 3 y 4. La razón es que el valor absoluto del sesgo asintótico de $p_n^{LL}(\cdot)$ es menor que el valor absoluto del sesgo de $p_n^{NW}(\cdot)$, en el modelo 3 en todo punto del intervalo de integración $t \in [0.2, 0.75]$, y en el modelo 4 para una amplia región del intervalo $[0, 0.75]$, como puede verse en la figura 5.2.

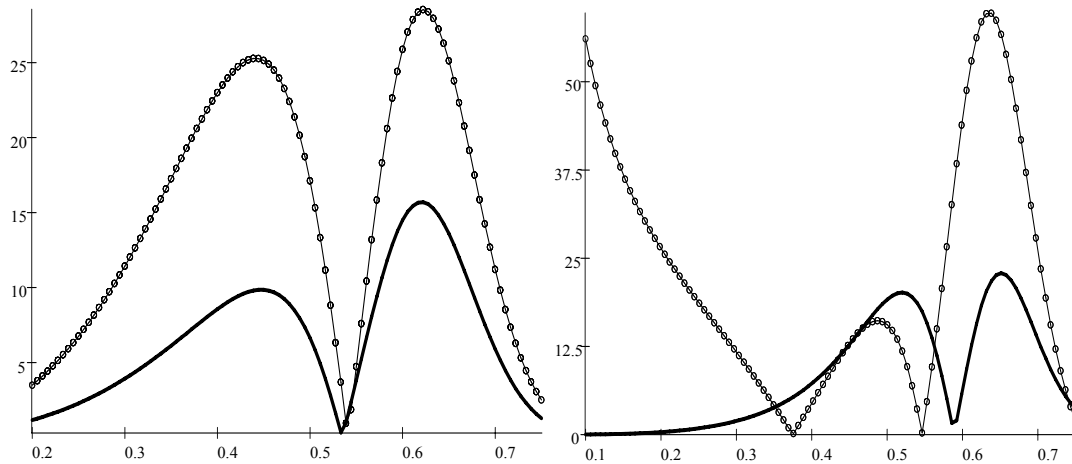


Figura 5.2. Valor absoluto del sesgo asintótico del estimador de $p(\cdot)$ de Nadaraya-Watson (línea con puntos) y del estimador local lineal (línea continua) para los modelos 3 (izquierda) y 4 (derecha).

No obstante, para el modelo 4, si en el cálculo de $MISE(f_n^P(\cdot))$ se considera una función peso constante en, digamos, $[0.35, 0.575]$ en vez de en el intervalo $[0, 0.75]$, entonces la eficiencia de $f_n^{P,NW}(\cdot)$ sería mayor que la del estimador $f_n^{P,LL}(\cdot)$. Esto se debe a que en dicho intervalo el valor absoluto de $B_n^{NW}(\cdot)$ es menor que el de $B_n^{LL}(\cdot)$. Por tanto, en los puntos de dicho intervalo, se espera una mejor estimación de $f(\cdot)$ presuavizando con un ajuste de Nadaraya-Watson antes que con un ajuste local lineal.

Con respecto a los modelos 1 y 2, el sesgo y la varianza asintóticos de $p_n^{NW}(\cdot)$ y $p_n^{LL}(\cdot)$ son iguales puesto que $h'(t) = 0$ para $t \in [0, 1]$. Aunque entonces se esperaría un comportamiento similar del estimador presuavizado con ambos ajustes de $p(\cdot)$, los resultados son mejores con el estimador de Nadaraya-Watson. En cualquier caso, la eficiencia relativa de $f_n^{P,LL}(\cdot)$ con respecto a $f_n^{P,NW}(\cdot)$ aumenta cuando $n \rightarrow \infty$, tal como se puede ver en la tabla 5.3.

Esta inesperada mayor eficiencia de $f_n^{P,NW}(\cdot)$ con respecto a $f_n^{P,LL}(\cdot)$ en estos modelos puede ser debida al hecho de que la función $p(\cdot)$ es una función de probabilidad y, como tal, sólo toma valores entre 0 y 1, de modo que cualquier estimador de $p(\cdot)$ debería tomar valores dentro de dicho intervalo. Esta condición, que cumple el estimador de Nadaraya-Watson pero no siempre el estimador local lineal, la verifica también el estimador local logístico. A pesar de que se han considerado valores truncados de $p_n^{LL}(\cdot)$ para forzar al estimador local lineal a tomar valores en $[0, 1]$, el ajuste local logístico de $p(\cdot)$ ofrece, en general, mejores resultados.

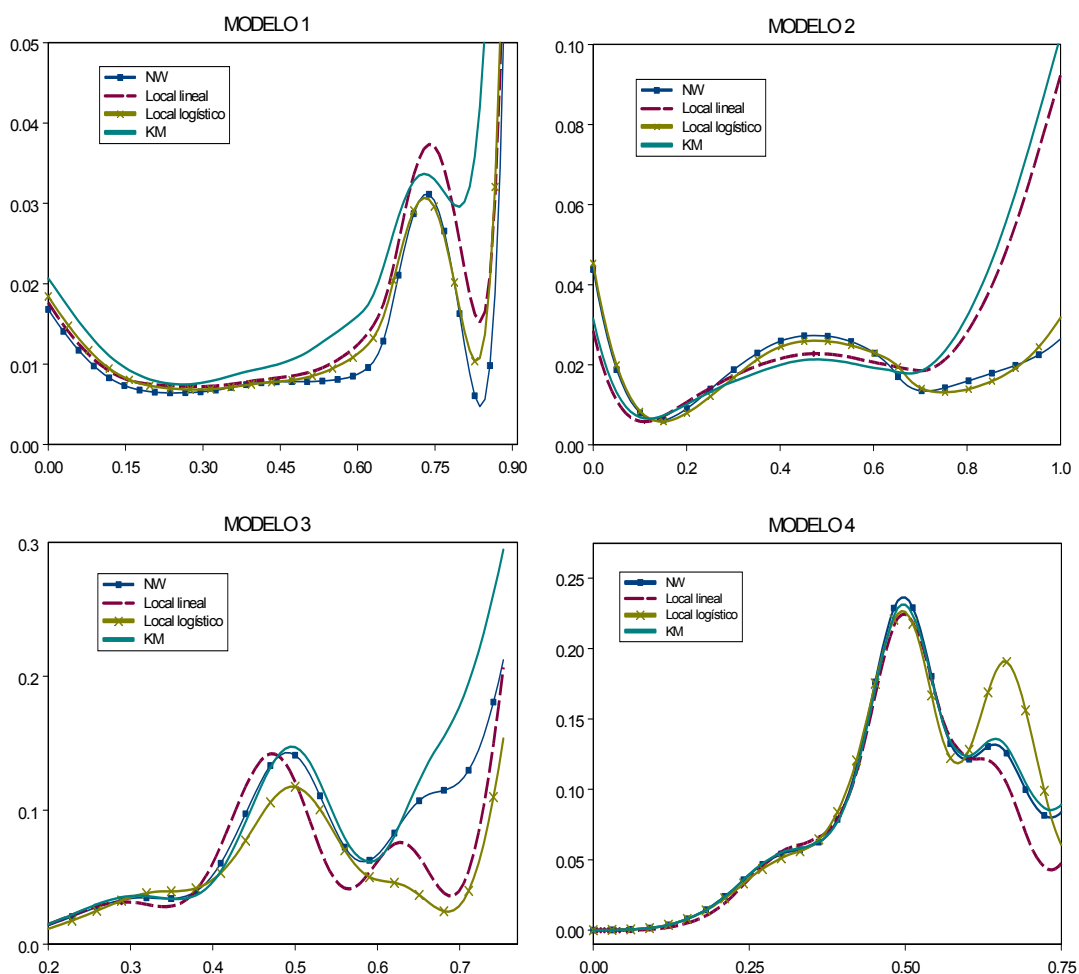


Figura 5.3. Error cuadrático medio (MSE) cometido por los tres estimadores presuavizados considerados, y por el clásico con pesos Kaplan-Meier para los modelos 1-4 con tamaño muestral $n = 100$.

La eficiencia del estimador presuavizado de $f(\cdot)$ aumenta con la flexibilidad de la función con la que se ajuste localmente la función $p(\cdot)$.

El comportamiento tan bueno de $f_n^{P,NW}(\cdot)$ en el modelo 1 se debe a la forma casi plana de $p(\cdot)$ para todo $t \in [0, 1]$. Puesto que la función $p(\cdot)$ es casi constante, la mejor forma de estimarla localmente es mediante un ajuste local por una constante, tal como hace el estimador de Nadaraya-Watson.

Además, para $n = 100$, se ha estimado el error cuadrático medio (MSE) puntualmente para cada uno de los modelos anteriores, cada estimador obtenido con las ventanas óptimas que minimizan su correspondiente $MISE$. El ajuste local logístico considerado ha sido el basado en la minimización por máxima verosimilitud.

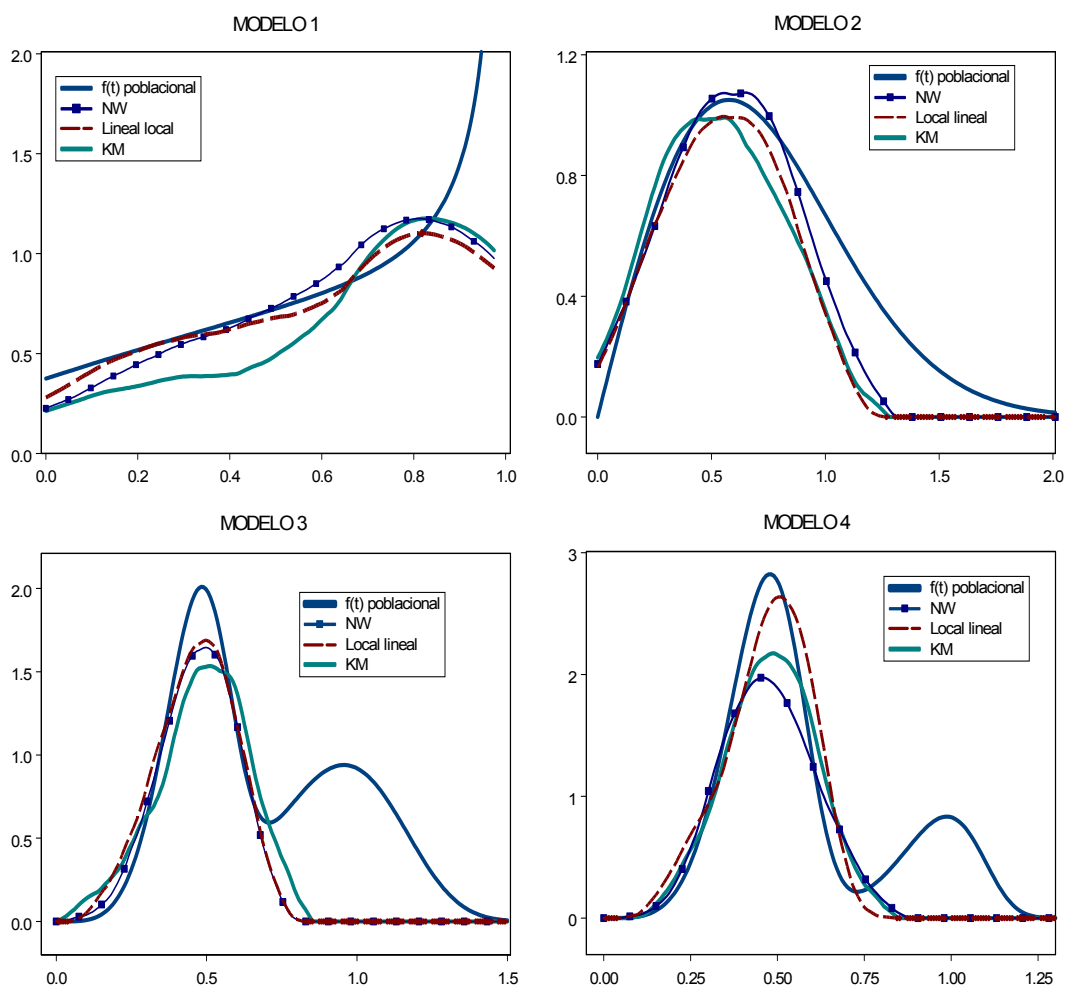


Figura 5.4. Función de densidad poblacional (línea gruesa), estimación presuavizada con ajuste de Nadaraya-Watson (línea con puntos), ajuste local lineal (línea discontinua), y estimador clásico con pesos de Kaplan-Meier (línea fina), para los modelos 1-4 con muestras de tamaño $n = 100$, correspondientes al comportamiento mediano en términos de ISE .

En la figura 5.3 se representa el error cometido con el estimador con pesos Kaplan-Meier, $f_n^{KM}(\cdot)$, el cometido por el estimador presuavizado que usa para estimar $p(\cdot)$ el estimador de Nadaraya-Watson, $f_n^{P,NW}(\cdot)$, y finalmente el error cometido por el estimador $f_n^{P,LL}(\cdot)$. Se puede comprobar cómo los estimadores presuavizados, $f_n^P(\cdot)$, tanto el que estima la función $p(\cdot)$ mediante el estimador de Nadaraya-Watson como el que usa el estimador local lineal, cometen un error menor que el clásico con pesos Kaplan-Meier, $f_n^{KM}(\cdot)$, en la mayor parte de los puntos observados. Esta mejora se produce incluso en los modelos 2 y 3, para los que se obtuvo $L_0 = 0$, es decir, que el cociente entre las ventanas óptimas s y b tiene a cero, y por tanto el comportamiento asintótico del estimador presuavizado y el clásico con pesos Kaplan-Meier serían

equivalentes, ya que la eficiencia del primero respecto al segundo es, en esos modelos, tan sólo de tercer orden.

Finalmente, la figura 5.4 muestra, en cada gráfica, cuatro curvas: la función de densidad poblacional, $f(\cdot)$, la estimación con pesos de Kaplan-Meier, $f_n^{KM}(\cdot)$, el estimador con estimación previa de $p(\cdot)$ mediante Nadaraya-Watson, $f_n^{NW}(\cdot)$, y con ajuste local lineal, $f_n^{LL}(\cdot)$, para una muestra de tamaño $n = 100$. Todos los estimadores han sido calculados con el valor de las correspondientes ventanas $MISE$. Las curvas muestran el comportamiento mediano de los estimadores, es decir, para cada modelo se han generado $m = 500$ muestras diferentes, y para cada una se ha calculado su error cuadrático integrado (ISE). Las muestras han sido ordenadas de acuerdo a sus valores ISE , y se han seleccionado las muestras que corresponden a la mediana de estas series ordenadas.

Se observa un efecto frontera en $t = 0$ en los modelos 1 y 2 (ver figura 5.4), pero ninguno de los estimadores propuestos lo recogen adecuadamente. Aunque todos se comportan de manera similar en un entorno de $t = 0$, el estimador con pesos de Nadaraya-Watson es mejor en el modelo 1. Esto es así aunque la función $p(\cdot)$ no es continua en $t = 0$, y por lo tanto se espera también un efecto frontera en su estimación (ver figura 5.1), y a pesar de que el estimador $p_n^{LL}(\cdot)$ presenta un mejor comportamiento en los puntos de la frontera que $p_n^{NW}(\cdot)$. Esto puede ser debido a la selección de ventanas globales en la estimación de $p(\cdot)$ y $f(\cdot)$. La estimación presuavizada con ajuste local lineal de $p(\cdot)$ probablemente mejorará en las fronteras de los intervalos de observación seleccionando adecuadamente ventanas locales.

En resumen, si $B_n^{NW}(\cdot) = B_n^{LL}(\cdot)$, entonces es mejor presuavizar mediante la estimación de $p(\cdot)$ con el estimador de Nadaraya-Watson antes que con el suavizador local lineal. En caso contrario, cuando $|B_n^{NW}(\cdot)| > |B_n^{LL}(\cdot)|$, la estimación local lineal de $p(\cdot)$ proporciona estimaciones presuavizadas más eficientes que el ajuste de Nadaraya-Watson.

En el caso concreto en el que se tenga una función $p(\cdot)$ casi constante, de modo que se verifique aproximadamente el modelo de Koziol-Green, entonces el estimador de $p(\cdot)$ que proporciona estimaciones presuavizadas de $f(\cdot)$ más eficientes es el estimador de Nadaraya-Watson.

En cualquier caso, independientemente de los sesgos $B_n^{NW}(\cdot)$ y $B_n^{LL}(\cdot)$, y de la forma de la función $p(\cdot)$, el ajuste local logístico es siempre una buena elección en la estimación presuavizada de la función de densidad.

Capítulo 6

Aplicación a datos reales

Este capítulo pretende ilustrar el comportamiento práctico de los métodos de estimación de las funciones de distribución y densidad, expuestos en capítulos anteriores, mediante su aplicación a un conjunto de datos reales procedentes del campo de la Medicina.

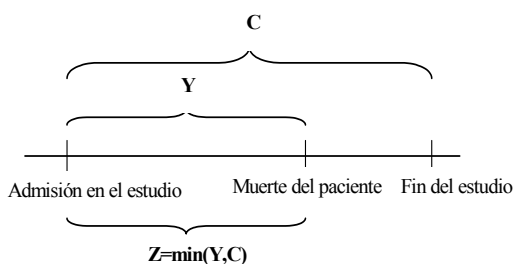
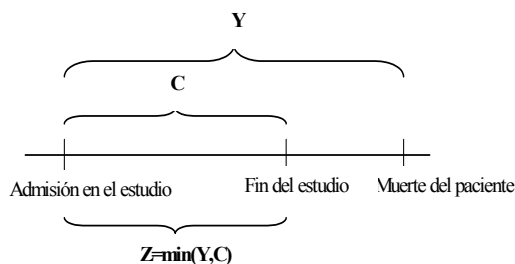
Los datos analizados son los del programa de trasplante de corazón de Stanford (Crowley y Hu (1977)). Estos datos han sido objeto de numerosos estudios en el ámbito del análisis de supervivencia, sobre todo para examinar el efecto de covariables (dependientes del tiempo o no) en la supervivencia de los enfermos de corazón. La variable de interés es el tiempo de supervivencia, para la cual se dispone de los 103 pacientes admitidos en el programa de trasplante de corazón de Stanford en el período comprendido entre el 1 de octubre de 1967 hasta el 1 de abril de 1974. Asociado a cada paciente se tiene la fecha de aceptación en el programa, T_1 , y la fecha de su muerte o, si ésta no llega a observarse, la fecha en la que se finalizó su seguimiento, T_2 . De esta manera, la supervivencia observada para cada paciente, en días, viene dada por $Z = T_2 - T_1$. La censura aleatoria por la derecha surge debido a que, al final del estudio, hay pacientes que todavía viven o a los que se interrumpió su seguimiento, de modo que su tiempo de supervivencia real, Y , desconocido, es en todo caso superior al tiempo de supervivencia observado Z . Usando la notación introducida en el primer capítulo tenemos:

Z = "Tiempo, en días, desde la entrada en el estudio hasta el fin del seguimiento del paciente".

Y = "Tiempo, en días, desde la entrada en el estudio hasta la muerte del paciente".

C = "Tiempo, en días, desde la entrada en el estudio hasta el fin del mismo o, en su caso, hasta que se pierde el seguimiento al paciente".

δ = "Indicador de censura (0=no censurado, 1=censurado)".

**Figura 6.1.** Ejemplo de caso no censurado.**Figura 6.2.** Ejemplo de caso censurado.

Para entender mejor la notación, se han representado, en las figuras 6.1 y 6.2, las variables Y , C y Z en un ejemplo de un dato censurado y otro no censurado.

La censura está presente en el 27.18 % de los casos, la mayor parte de ella (25.24 %) debido a que el paciente llega vivo al fin del estudio. El restante 1.94 % son pacientes a los que se perdió el seguimiento antes de que fallecieran.

Para cada paciente se han registrado otras variables de interés, que se describen brevemente a continuación, analizadas en trabajos previos con el fin de estudiar su efecto sobre el tiempo de supervivencia. En la tabla 6.1 figura la frecuencia de las principales variables consideradas:

1. Fecha de nacimiento
2. Fecha de admisión en el programa.
3. Fecha del trasplante de corazón (de ser el caso).
4. Fecha del fin de seguimiento del paciente.
5. Estado al finalizar el seguimiento del paciente (vivo o muerto).
6. Operaciones a corazón abierto previas (sí o no).
7. Presencia del antígeno HLA-A2 en el donante pero no en el receptor (sí o no).
8. N° de alelos del donante que no coinciden con los del receptor (desde 1 hasta 4).

A partir de las variables anteriores se han registrado también las siguientes: tiempo de supervivencia, indicador de censura, tiempo de espera hasta haber recibido un trasplante (de ser el caso), edad en el momento de ser admitido en el programa, edad en el momento del trasplante, edad al finalizar el seguimiento e indicador de si la edad al finalizar el estudio es mayor o menor que 48 años.

Las diferentes variables que registran la edad han sido consideradas puesto que los riesgos de la operación, la respuesta al trasplante y la terapia posterior pueden variar con la edad del paciente. Asimismo, el rechazo a los órganos transplantados depende, entre otras cosas, de la presencia de antígenos HLA y de la compatibilidad entre el donante y el receptor. Finalmente, el número de operaciones previas es importante

Tabla 6.1 Principales características de los pacientes del programa de transplante de corazón de Stanford.

Característica	Frecuencia (%)
EDAD	
Menor de 48 años	46.6
48 años o más	53.4
OPERADO PREVIAMENTE	
Sí	15.5
No	84.5
TRANSPLANTADO	
Sí	67
No	33
HLA-A2	
Sí	26.2
No	73.8

puesto que van acompañadas de transfusiones de sangre con diferentes antígenos HLA que pueden provocar en el paciente, a la hora de recibir un transplante de corazón, una mejor respuesta inmunitaria.

En la figura 6.3 se puede ver, en un gráfico de dispersión, el tiempo de supervivencia observado (en días) en función de la edad en el momento de admisión en el programa. En él se puede comprobar que los mayores tiempos de supervivencia coinciden con valores censurados, representados por \circ . Además, para evitar la fuerte asimetría de la distribución de la variable Z , a la hora de estimar la función de densidad trabajaremos con los datos transformados por la función logaritmo neperiano.

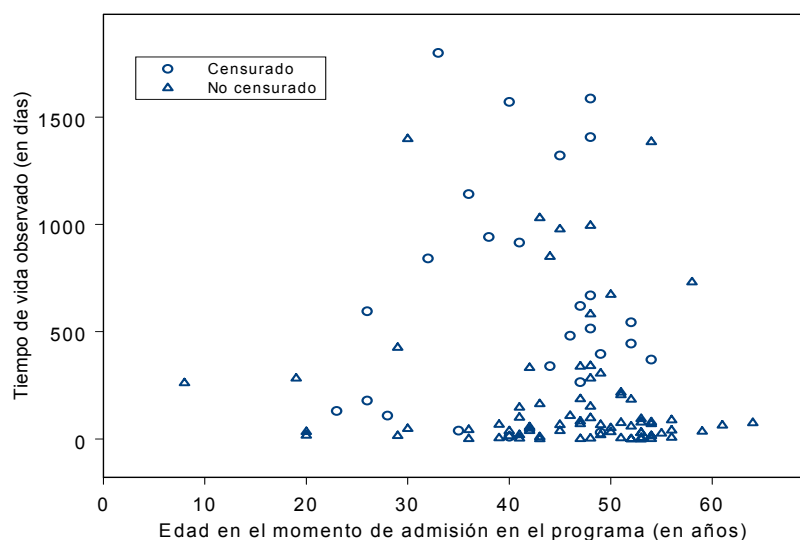


Figura 6.3. Gráfico de dispersión de los tiempos de supervivencia observados en función de la edad. El símbolo \circ denota las observaciones censuradas, mientras que las observaciones no censuradas están representadas por \triangle .

6.1. Estimación de la función de distribución

En el capítulo 2 se estudiaron las propiedades asintóticas del estimador presuavizado de la función de supervivencia $1 - F(\cdot)$, presentado como alternativa al estimador de Kaplan-Meier por Cao *et al.* (2005). Así, entre otros resultados, se daba una expresión para el parámetro de suavización b que minimiza la parte dominante de $MSE(\overline{F}_n^P(t))$:

$$b_{OPT}(t) = \left(\frac{q(t) e_K}{2\alpha^2(t) d_K^2} \right)^{1/3} n^{-1/3}. \quad (6.1)$$

La expresión de esta ventana es la misma que la de la ventana asintóticamente óptima del estimador presuavizado de la función de fallo acumulativa (ver Cao *et al.* (2005)):

$$\Lambda_F(t) = \int_0^t \lambda_F(v) dv, \quad \text{siendo} \quad \lambda_F(t) = \frac{f(t)}{1 - F(t)}$$

la función de riesgo. Ambas funciones son muy prácticas para describir cambios en el tiempo de la probabilidad de que un individuo experimente un fallo. Además, la función de riesgo acumulativa y la función de supervivencia están directamente relacionadas de la siguiente manera:

$$1 - F(t) = \exp[-\Lambda_F(t)].$$

Del mismo modo, el estimador presuavizado de la función de fallo acumulativa, el llamado estimador de Nelson-Aalen presuavizado, $\Lambda_n^P(\cdot)$, y el estimador de la función de distribución presuavizado, $F_n^P(\cdot)$, están directamente relacionados (ver proposición 2.4.1 en el capítulo 2). Es por ello que, como paso previo a la estimación de la función de distribución, hemos estimado la función de fallo acumulativa.

La ventana (6.1) es una ventana local, es decir, requiere calcularse en cada punto de estimación. La ventana global asintóticamente óptima en el sentido de minimizar el $AMISE(\overline{\Lambda}_n^P(t))$ es

$$b_{OPT} = \left(\frac{Q e_K}{2d_K^2 A} \right)^{1/3} n^{-1/3}, \quad (6.2)$$

donde

$$Q = \int_0^\infty q(v) \omega(v) dv = \int_0^\infty \frac{p(v)(1-p(v))h(v)}{(1-H(v))^2} \omega(v) dv$$

y

$$A = \int_0^\infty \alpha^2(v) \omega(v) dv,$$

siendo $\alpha(\cdot)$ la función dada en (2.38) en el capítulo 2.

El selector propuesto en Cao *et al.* (2005) es el tipo plug-in:

$$\hat{b} = \left(\frac{\hat{Q}e_K}{2d_K^2 \hat{A}} \right)^{1/3} n^{-1/3}, \quad (6.3)$$

donde

$$\hat{Q} = \frac{1}{n} \sum_{i=1}^n \frac{p_n(Z_i) (1 - p_n(Z_i)) \omega(Z_i)}{\left(1 - H_n(Z_i) + \frac{1}{n}\right)^2}$$

y la integral \hat{A} es

$$\hat{A} = \int_0^\infty \alpha_n^2(v) \omega(v) dv,$$

siendo

$$\alpha_n(t) = \int_0^t \frac{\frac{1}{2}p_n''(v)h_n(v) + p_n'(v)h_n'(v)}{1 - H_n(v) + \frac{1}{n}} dv$$

con $p_n(\cdot)$, $p_n'(\cdot)$ y $p_n''(\cdot)$ los estimadores de Nadaraya-Watson de $p(\cdot)$ y de sus derivadas primera y segunda respectivamente, y $h_n(\cdot)$ y $h_n'(\cdot)$ los estimadores de Parzen-Rosenblatt de la densidad $h(\cdot)$ y de su derivada primera.

De esta forma, el cálculo de (6.2) requiere la elección de dos ventanas piloto, g_1 y g_2 , para los estimadores \hat{A} y \hat{Q} respectivamente. Sus valores asintóticamente óptimos vienen dados por las siguientes propiedades.

Propiedad 6.1.1 (Teorema 7 de Cao *et al.* (2005))

Bajo las hipótesis (K1), (p2), (p3), (H3), (H4) y si la ventana g_1 verifica $ng_1^3 \left(\log \frac{1}{g_1}\right)^{-3} \rightarrow \infty$ y $ng_1^5 \left(\log \frac{1}{g_1}\right)^{-1} \rightarrow 0$, entonces

$$g_{1,AMSE} = Cn^{-1/5} \quad \text{donde } C = \begin{cases} \left(-\frac{C_2}{C_1}\right)^{1/5} & \text{si } C_1 < 0 \\ \left(\frac{3C_2}{2C_1}\right)^{1/5} & \text{si } C_1 > 0 \end{cases}$$

siendo

$$C_1 = \frac{1}{2}d_K \int_\varepsilon^\infty \int_\varepsilon^t \frac{1}{1 - H(r)} \left(p^{(4)}(r)h(r) + 4p'''(r)h'(r) + 5p''(r)h''(r) \right. \\ \left. + 4p'(r)h'''(r) - 2p'(r)h^{-1}(r)h'(r)h''(r) \right) dr \alpha(t) dt,$$

y

$$C_2 = \frac{1}{4}c_K \int_\varepsilon^\infty \frac{(1 - p(v))p(v)h(v)}{(1 - H(v))^2} \omega(v) dv.$$

Propiedad 6.1.2 (Teorema 8 de Cao et al. (2005))

Bajo las hipótesis (K1), (p2), (p3), (H3), (H4) y si la ventana g_2 verifica $ng_2^{8/3} \rightarrow \infty$ y $ng_2^4 \rightarrow 0$, entonces

$$g_{2,AMSE} = Dn^{-1/3} \text{ donde } D = \begin{cases} \left(\frac{D_2}{2D_1}\right)^{1/3} & \text{si } D_1 < 0 \\ \left(-\frac{D_2}{D_1}\right)^{1/3} & \text{si } D_1 > 0 \end{cases}$$

siendo

$$D_1 = d_K \int_{\varepsilon}^{\infty} \frac{(1 - 2p(v))(p'(v)h'(v) + \frac{1}{2}p''(v)h(v))}{(1 - H(v))^2} dv$$

y

$$D_2 = -c_K \int_{\varepsilon}^{\infty} \frac{p(v)(1 - p(v))}{(1 - H(v))^2} \omega(v) dv.$$

Una de las condiciones para que garantizan la consistencia de este procedimiento de selección de la ventana para $\Lambda_n^P(\cdot)$, así como la consistencia del selector tipo plug-in de las ventanas para $f_n^P(\cdot)$, es la hipótesis (p3), por la cual $p(t) = 1$ para todo $t > 0$ suficientemente próximo a 0. Para evaluar si es admisible esta hipótesis para este conjunto de datos, se ha obtenido la estimación local lineal de $p(\cdot)$ (ver figura 6.4). El selector de la ventana ha sido el plug-in propuesto por Ruppert, Sheather y Wand (1995) el cual, suponiendo que $H(\cdot)$ es una función de distribución de Weibull, y ajustando, en el caso de $p(\cdot)$, un modelo logístico, proporcionó la ventana $\hat{b}^{LL} = 748.9$.

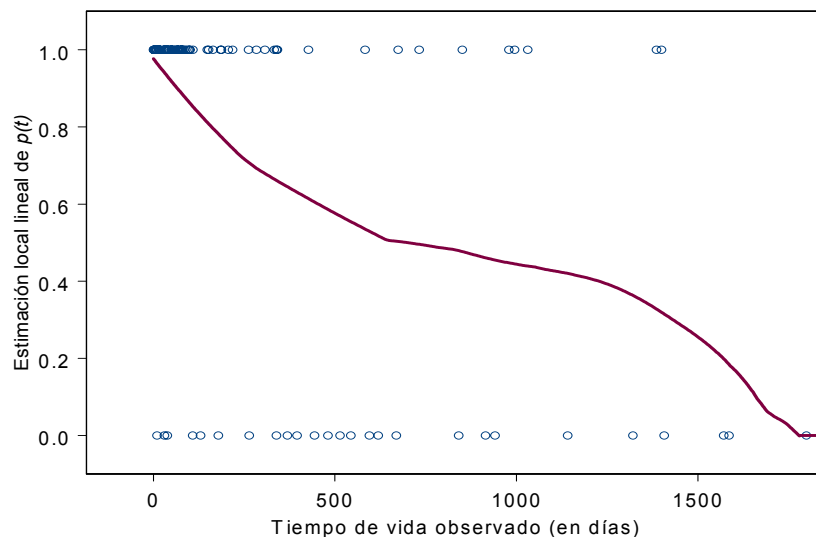


Figura 6.4. Estimación local lineal de la probabilidad condicional de no censura $p(\cdot)$. Los puntos indican el tiempo de vida observado (en días) y el estado de censura (0=no censurado, 1=censurado) para los datos de trasplante de corazón de Stanford.

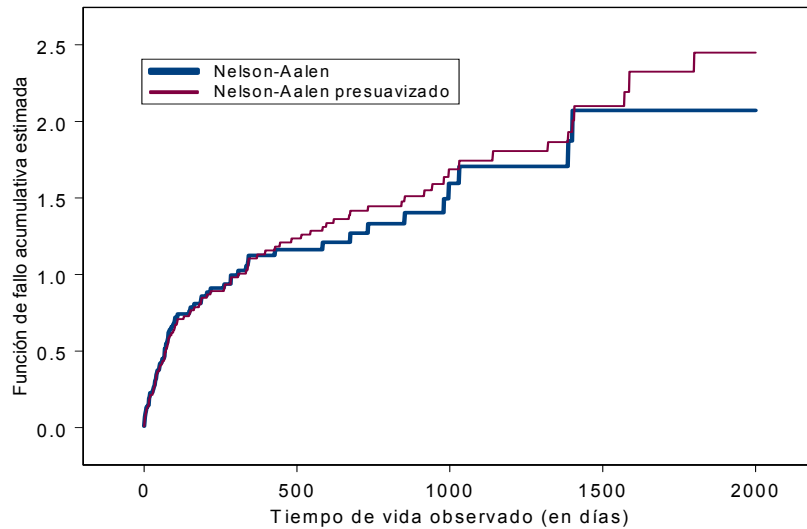


Figura 6.5. Estimaciones de Nelson-Aalen clásica y presuavizada de la función de fallo acumulativa para los datos de trasplante de corazón de Stanford.

Aunque el valor de la estimación en $t = 0$, $p_n^{LL}(0) = 0.9755$, no es exactamente 1, tanto por la tendencia general de la curva como por la posibilidad de estar ante una ligera sobresuavización (la longitud del intervalo en el que se observan datos es tan sólo 2.406 veces la ventana óptima), se puede aceptar razonablemente que la condición (p3) se verifica para este conjunto de datos.

Para calcular la ventana plug-in, \hat{b} , dada en (6.3), con la que obtener el estimador presuavizado $\Lambda_n^P(\cdot)$, las ventanas piloto necesarias, dadas en las propiedades 6.1.1 y 6.1.2, han sido estimadas suponiendo los mismos modelos paramétricos que en la estimación de la función $p(\cdot)$, es decir, asumiendo que $H(\cdot)$ es una función de distribución de Weibull, y tomando, para el ajuste de $p(\cdot)$, un modelo logístico. De esta forma, las ventanas piloto obtenidas han sido $g_1 = 101.01$ y $g_2 = 432.39$, y la ventana plug-in $\hat{b} = 217.04$.

En la figura 6.5 se representa la estimación presuavizada de la función de fallo acumulativa, junto con la obtenida empleando el estimador clásico, el estimador de Nelson-Aalen. Ambas estimaciones son parecidas, aunque el efecto de la presuavización es evidente a partir de, aproximadamente, el primer año de tomar parte en el programa de trasplante.

No hay propuesto en la literatura ningún método para seleccionar adecuadamente la ventana b óptima para el estimador presuavizado de la función de distribución $F_n^P(\cdot)$. No obstante, dado que la expresión de la ventana local asintóticamente óptima para

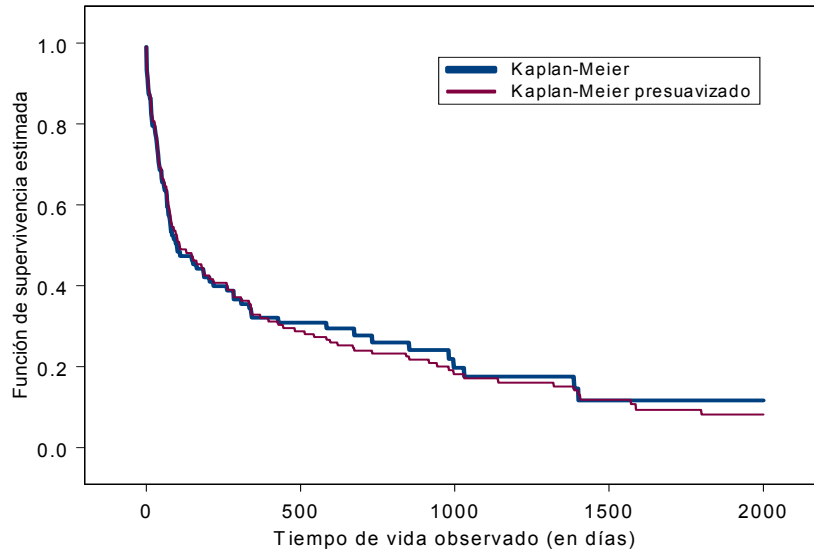


Figura 6.6. Estimaciones de Kaplan-Meier clásica y presuavizada de la función de supervivencia para los datos de transplante de corazón de Stanford.

el estimador presuavizado de la distribución, $F_n^P(\cdot)$, es la misma que para la función de fallo acumulativa presuavizada, $\Lambda_n^P(\cdot)$, la ventana global empleada ha sido (6.2).

Las estimaciones de la función de supervivencia de Kaplan-Meier y su versión presuavizada con ventana \hat{b} se representan en la figura 6.6. Aunque la forma de ambas estimaciones es muy similar, con un rápido decrecimiento en el primer año y medio de vida observado (desde la incorporación al programa de transplante), a partir del primer año se advierte mejor el carácter local del estimador presuavizado (recordemos que, a diferencia del estimador de Kaplan-Meier clásico que sólo otorga peso a los datos no censurados, el estimador presuavizado tiene saltos en todas las observaciones). Esto produce mejores estimaciones de la supervivencia en la cola derecha del intervalo de observación. A modo de ejemplo, en la tabla 6.2 se presentan los deciles 2, 4, 6 y 8 de la distribución estimada. La diferencia entre ambos estimadores aumenta, en general, con el decil buscado.

Tabla 6.2. Estimación de los deciles 2, 4, 6 y 8 de la distribución del tiempo de vida.

Deciles	Kaplan-Meier	Presuavizado
2	19.533	20.91
4	67.726	71.097
6	218.06	261.90
8	996.74	979.05

6.2. Estimación de la función de densidad

A continuación aplicamos la metodología de presuavización para estimar la función de densidad de la variable Y , tiempo de vida. Debido a la fuerte asimetría positiva de la distribución de la variable Z , tiempo de vida observado (ver figura 6.3), en esta sección trabajaremos con los datos transformados por la función logaritmo neperiano. De esta forma, los datos se encuentran ahora en del intervalo $[0, 7.5]$.

Se ha elegido como ventana, s^{KM} , para el estimador con pesos de Kaplan-Meier, $f_n^P(\cdot)$, la propuesta por Sánchez Sello *et al.* (1999), en el contexto más general de censura aleatoria por la derecha y truncamiento por la izquierda, dada en (4.74) en el capítulo 4. En su cálculo se precisa la estimación de la ventana piloto g dada en (4.75). El valor resultante fue $s^{KM} = 0.72$.

Como selectores de las ventanas s y b para el estimador presuavizado hemos considerado todos los métodos expuestos en el capítulo anterior, es decir, el método plug-in y el método bootstrap, éste último con los 5 remuestreos propuestos.

Para obtener el valor de las ventanas plug-in del estimador presuavizado, $f_n^P(\cdot)$, se han seguido los mismos pasos que en el estudio de simulación del capítulo 4. Se han considerado, por tanto, cuatro ventanas piloto. La ventana piloto b_0 , estimada por validación cruzada, se ha usado para obtener el estimador de Nadaraya-Watson $p_n(\cdot)$, sus derivadas primera, segunda y el estimador presuavizado $F_n^P(\cdot)$. Con la segunda ventana piloto, llamada b_1 y cuya expresión es (4.76), se calculó el estimador de Parzen-Rosenblatt de $h(\cdot)$ y de su derivada primera $h'(\cdot)$. Con la tercera ventana piloto, s_0 , junto con b_0 , se calculó $f_n^P(\cdot)$, y con la ventana g dada en (4.75) se estimó $f''(\cdot)$. Las funciones poblacionales y, en consecuencia, desconocidas, que se necesitan para obtener estas ventanas piloto se han estimado suponiendo modelos paramétricos. En concreto, para la función de densidad de la variable aleatoria de interés, Y , se supuso que sigue un modelo paramétrico dado por una mezcla de tres lognormales, mientras que para $h(\cdot)$ se ha considerado una distribución de Weibull. Finalmente, como ventanas piloto que se precisan en la obtención de las ventanas bootstrap se tomaron las usadas en el método plug-in.

En la tabla 6.3 se muestran los valores de las ventanas obtenidas por los diferentes métodos considerados. La notación es la misma que es usó en el capítulo 4. En la representación gráfica de las estimaciones de $f(\cdot)$, por simplicidad, se ha considerado únicamente las correspondientes a las ventanas calculadas por el método plug-in y el método bootstrap con remuestreo $SC2$.

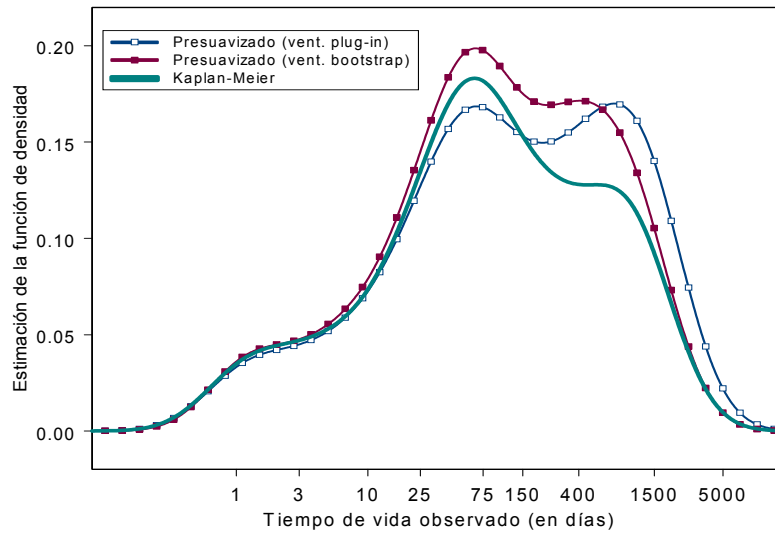


Figura 6.7. Estimación de la función de densidad mediante el estimador $f_n^{KM}(\cdot)$ con selector de ventana plug-in, y los estimadores presuavizados $f_n^P(\cdot)$ con selectores de ventana plug-in y bootstrap.

En el examen a la figura 6.7 destaca la discrepancia entre la estimación de Kaplan-Meier (línea continua) y las estimaciones presuavizadas (líneas con puntos) a partir de, aproximadamente, el día 20. El estimador $f_n^{KM}(\cdot)$ tiene una única moda, mientras que los estimadores $f_n^P(\cdot)$ sugieren que la población podría estar dividida en dos grupos, en función de alguna covariable que influye de manera notable en el tiempo de supervivencia del paciente.

Comparación de los estimadores de Nadaraya-Watson y local lineal

Previamente al análisis de las covariables presentadas al comienzo de este capítulo, veamos si para este conjunto de datos se obtienen diferentes resultados al estimar $p(\cdot)$ con un ajuste local lineal en vez de usar el estimador de Nadaraya-Watson. En el capítulo anterior se concluyó que, si $B_n^{NW}(\cdot) = B_n^{LL}(\cdot)$, entonces es mejor presuavizar

Tabla 6.3. Valores de las ventanas s y b obtenidas con los métodos plug-in y bootstrap.

Método	Ventana s	Ventana b
Plug-in	0.720	1.290
SP	0.487	0.869
OP	0.526	0.869
SC1	0.527	1.090
SC2	0.640	1.310
CS	0.487	0.869

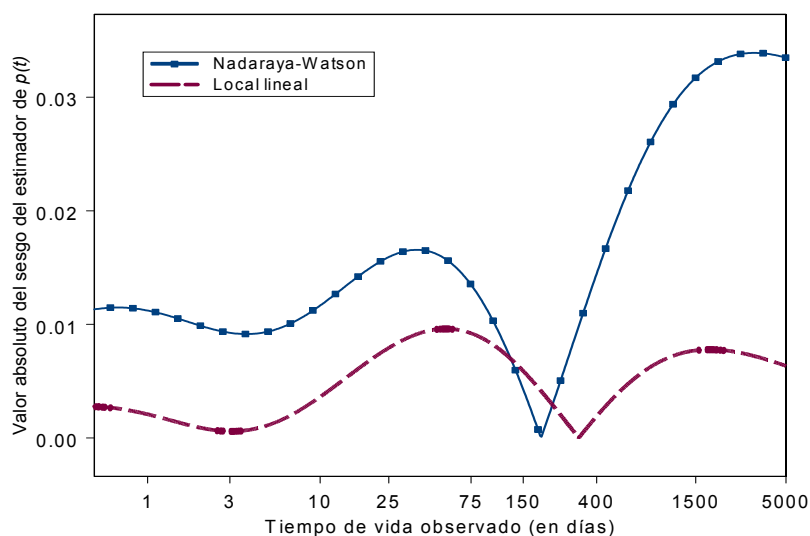


Figura 6.8. Valor absoluto del sesgo asimptótico estimado del estimador de $p(\cdot)$ de Nadaraya-Watson (línea con puntos) y del estimador local lineal (línea discontinua).

usando el estimador de Nadaraya-Watson en el ajuste de $p(\cdot)$, mientras que cuando $|B_n^{NW}(\cdot)| > |B_n^{LL}(\cdot)|$, la estimación local lineal de $p(\cdot)$ proporciona estimaciones presuavizadas más eficientes. En la figura 6.8 se representa el valor absoluto de las estimaciones de las expresiones asimptóticas de los sesgos de ambos estimadores de $p(\cdot)$, calculadas con ventana b_0 . Puesto que se tiene claramente que $|\widehat{B}_n^{NW}(\cdot)| > |\widehat{B}_n^{LL}(\cdot)|$ en la mayor parte de los puntos del intervalo de observación, la estimación presuavizada de $f(\cdot)$ parece ser más eficiente si el ajuste de la función $p(\cdot)$ es local lineal.

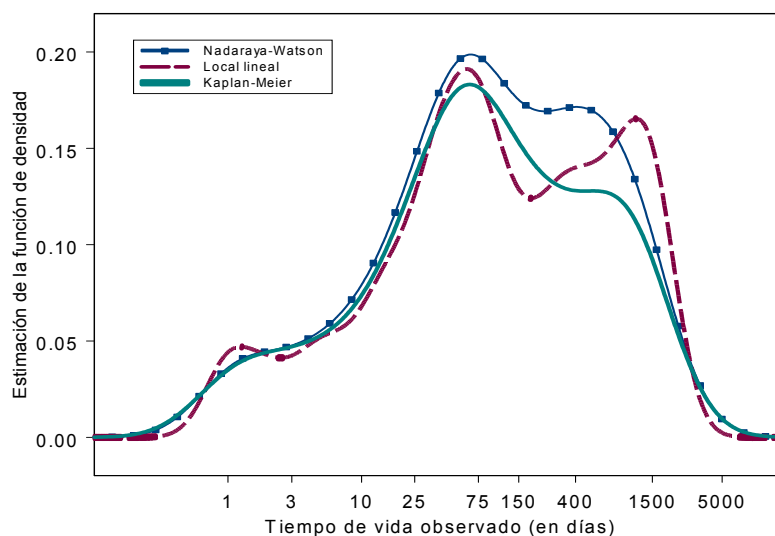


Figura 6.9. Estimaciones de la función de densidad mediante $f_n^{KM}(\cdot)$ (línea continua), $f_n^{P,NW}(\cdot)$ (línea con puntos) y $f_n^{P,LL}(\cdot)$ (línea discontinua).

Dado que la selección de las ventanas s y b para el estimador presuavizado con ajuste de $p(\cdot)$ local lineal es un problema abierto, en este ejemplo se utilizará un selector de tipo bootstrap, con remuestreo $SC2$. El valor de las ventanas ha sido $s_{\text{bootstrap}}^{LL} = 0.408$ y $b_{\text{bootstrap}}^{LL} = 4.70$. Las estimaciones presuavizadas de $f(\cdot)$ con ambos ajustes de $p(\cdot)$ se muestran en la figura 6.9 junto con el estimador de Kaplan-Meier, que sirve de referencia.

A pesar de que la estimación $f_n^{P,LL}(\cdot)$ parece menos suave que $f_n^{P,NW}(\cdot)$, su forma sigue insinuando que el tiempo de vida depende del valor de alguna covariable.

Análisis de la influencia de las covariables en el tiempo de vida

Es razonable pensar que alguna de las variables descritas en la tabla 6.1 afecten a la supervivencia de los pacientes. Con el fin de cuantificar la relación entre ellas y el tiempo de supervivencia, hemos ajustado el modelo de riesgos proporcionales de Cox (Cox (1972)). Este modelo es seguramente uno de los más utilizados en el campo de la Medicina para describir numéricamente la influencia de un conjunto de covariables en el tiempo de vida. Sean X y X^* el conjunto de covariables para dos individuos diferentes. Este modelo de regresión,

$$\lambda(t|X) = \lambda_0(t) e^{-X^T \beta_0},$$

supone que el riesgo relativo de los individuos es constante a lo largo del tiempo, es decir,

$$\frac{\lambda(t|X)}{\lambda(t|X^*)} = \frac{e^{-X^T \beta_0}}{e^{-X^{*T} \beta_0}}$$

no depende del tiempo.

En la tabla 6.4 se especifica el modelo de Cox ajustado para cada una de las variables de la tabla 6.1, así como el contraste de Wald y el p -valor asociado. Debido a la presencia de datos faltantes en la variable HLA-A2, el modelo correspondiente a dicha variable se ha construido únicamente con 65 datos.

Tabla 6.4. Modelo de Cox ajustado y tabla ANOVA correspondiente para las variables de la tabla 6.1.

Variable	n	$g.l.$	Coefficiente estimado	Error estándar	Estadístico de Wald	p -valor
Edad (<48 años)	103	1	-0.073	0.117	0.39	0.533
Operado prev.	103	1	-0.37	0.18	4.25	0.039
Transplantado	103	1	-0.662	0.122	29.5	<0.0001
HLA-A2	65	1	0.058	0.187	0.1	0.754

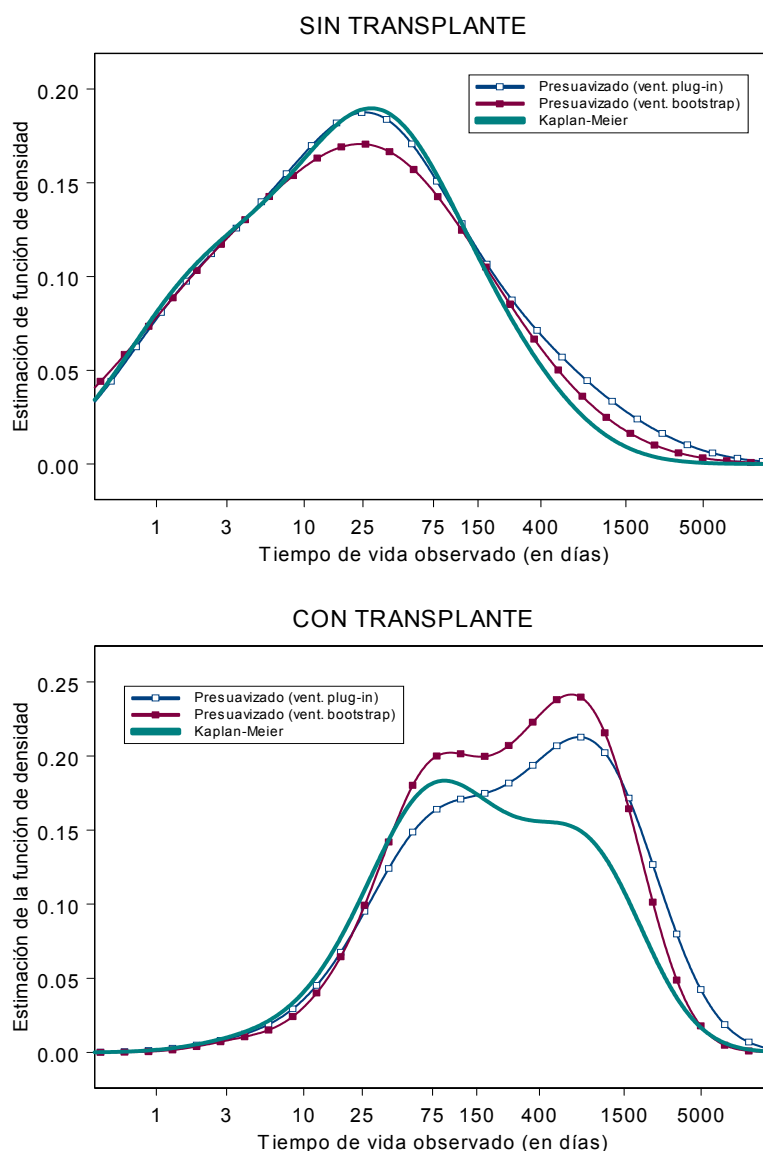


Figura 6.10. Estimación de la función de densidad del tiempo de vida mediante los estimadores $f_n^{KM}(\cdot)$ (línea continua) y $f_n^P(\cdot)$ (líneas con puntos), según el paciente haya recibido o no un transplante de corazón.

Por la tabla 6.4 se concluye que el factor que más influye en la supervivencia de los pacientes de corazón es, como se esperaba, el hecho de haber recibido un transplante de corazón. En concreto, el riesgo de morir de un paciente al que no se le ha transplantado un corazón es $\exp(0.662) = 1.94$ veces superior que el de los que sí han recibido un transplante (siempre suponiendo la validez del modelo de riesgos proporcionales de Cox, y sin tener en cuenta las restantes covariables). Los resultados son similares cuando se ajusta un modelo de Cox con los cuatro factores simultáneamente.

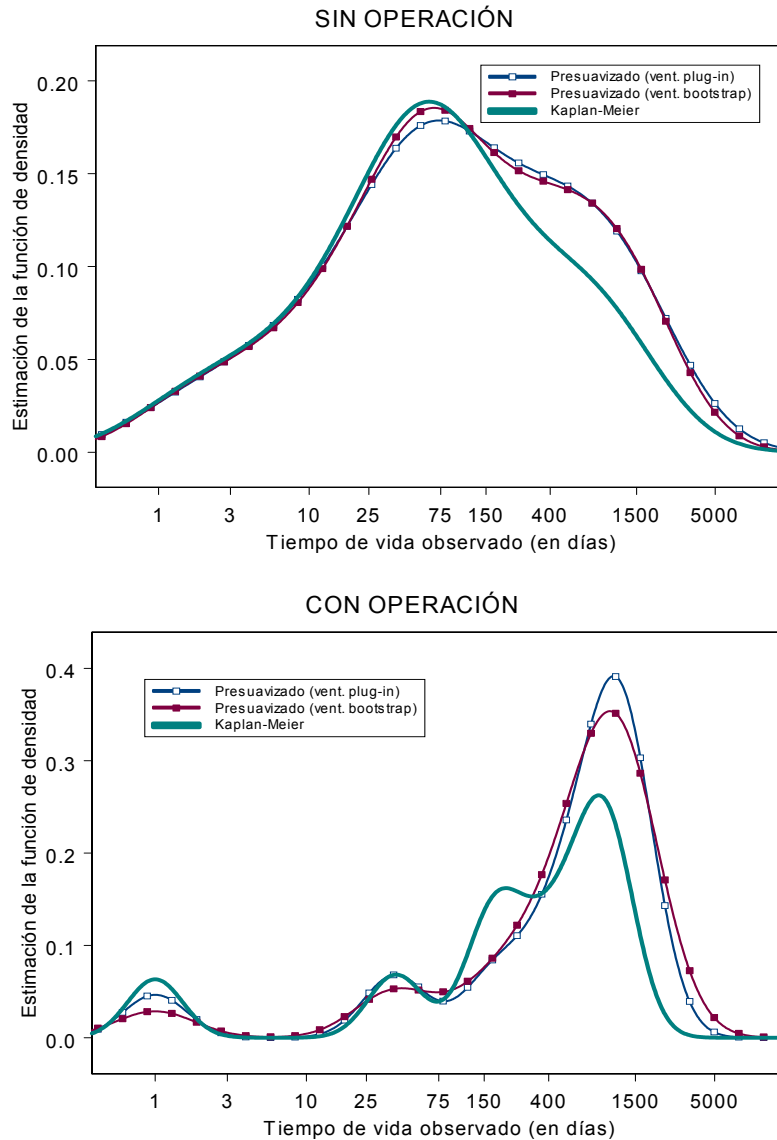


Figura 6.11. Estimación de la función de densidad del tiempo de vida mediante los estimadores $f_n^{KM}(\cdot)$ (línea continua) y $f_n^P(\cdot)$ (líneas con puntos), según el paciente haya sido operado previamente del corazón o no.

El segundo factor que más influye en la supervivencia es el haber sido objeto previamente de alguna operación a corazón abierto. En particular, los pacientes que no han sido operados previamente presentan un riesgo de morir $\exp(0.37) = 1.45$ veces mayor que los demás. La influencia de las demás covariables en el tiempo de vida es, de acuerdo con el contraste de Wald con un nivel de significación del 5 %, no significativa.

Para comprobar si haber recibido un trasplante está relacionado con el haber sido operado previamente del corazón, se ha realizado un contraste chi-cuadrado que ha resultado no significativo (p -valor= 0.187).

En las figuras 6.10 y 6.11 se muestran las estimaciones de la función de densidad con pesos de Kaplan-Meier (línea continua) y presuavizada con ventana plug-in (línea con puntos abiertos) y bootstrap con el remuestreo SC^2 (línea con puntos cerrados). Estas ventanas han sido estimadas asumiendo, en la obtención de las ventanas piloto, los mismos modelos paramétricos supuestos para las estimaciones de $f(\cdot)$ de la figura 6.7. El valor estimado de las ventanas viene dado en la tabla 6.5.

El análisis de las figuras 6.10 y 6.11 confirma que el tiempo de vida es muy distinto en función de si el paciente ha recibido o no un trasplante de corazón, o dependiendo de si al paciente se le ha practicado con anterioridad alguna operación a corazón abierto o no.

La estimación de la función de densidad del tiempo de vida, para los pacientes que no han recibido un trasplante de corazón o que no han sido objeto de operaciones previas de corazón, es parecida con las estimaciones de Kaplan-Meier clásica y presuavizada. Esto es debido al valor relativamente bajo de la ventana de presuavizado b , tanto plug-in ($b_{\text{plug-in}} = 0.065$) como bootstrap ($b_{\text{bootstrap}} = 0.198$). Como se recordará, cuando la ventana de presuavización b es cero, el estimador presuavizado $f_n^P(\cdot)$ se reduce al estimador clásico $f_n^{KM}(\cdot)$.

La diferencia entre las estimaciones de Kaplan-Meier clásica y presuavizada es mayor en los grupos de pacientes a los que sí se les ha realizado un trasplante (ver figura 6.10) y los que han sido operados del corazón con anterioridad (ver figura 6.11). El estimador de Kaplan-Meier clásico, $f_n^{KM}(\cdot)$, claramente infraestima el tiempo de vida de dichos pacientes (la elevación inicial de las estimaciones para el grupo de pacientes que sí han sido operados con anterioridad se debe a un valor aislado no censurado).

Tabla 6.5. Valores de las ventanas s y b obtenidas con los métodos plug-in y bootstrap para cada submuestra obtenida según las covariables que influyen en el tiempo de vida.

		Kaplan-Meier		Presuavizado		
		s	$s_{\text{plug-in}}$	$b_{\text{plug-in}}$	$s_{\text{bootstrap}}$	$b_{\text{bootstrap}}$
Trasplante	NO	1.065	1.081	0.065	1.323	0.198
	SÍ	0.626	0.636	2.251	0.640	3.70
Operación	NO	0.683	0.693	2.038	0.766	1.465
	SÍ	0.421	0.481	3.848	0.607	3.65

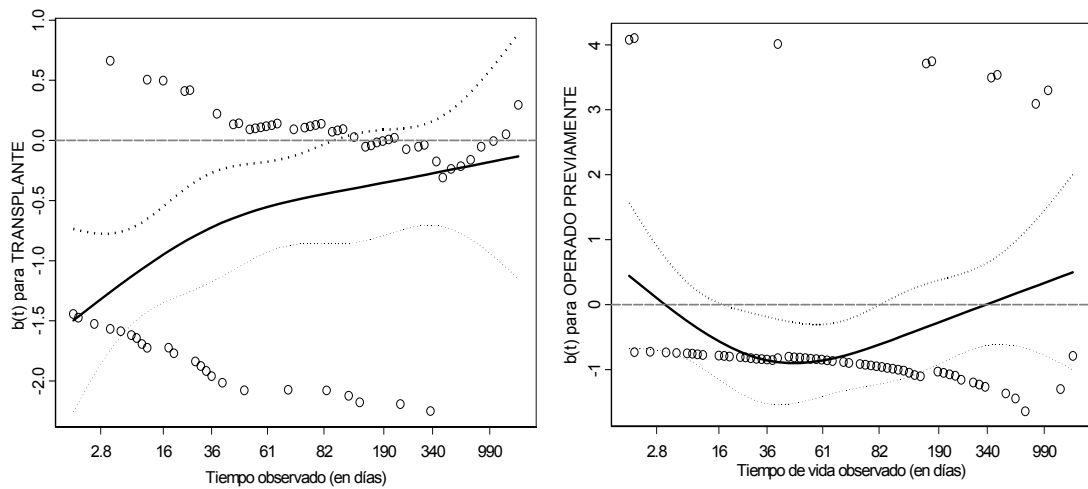


Figura 6.12. Diagrama de dispersión suavizado de los residuos de Schoenfeld escalados para los predictores Transplante y Operado previamente.

La deducción, a partir de los resultados de la tabla 6.4, de que el tiempo de supervivencia de un paciente depende, entre otros posibles factores, de si ha recibido un transplante de corazón o de si ha sido operado del corazón con antelación, está supeditado a la validez de la regresión de Cox. Esto precisa que sea cierta la hipótesis de riesgos proporcionales, es decir, el riesgo relativo de los individuos cualesquiera es constante a lo largo del tiempo.

Debido a la gran aceptación de este modelo de regresión en el análisis de la relación entre el tiempo de supervivencia y un conjunto de covariables, sobre todo en Medicina, es copiosa la literatura sobre contrastes de bondad de ajuste del modelo de Cox (Andersen (1982), Arjas (1988), Lin y Wei (1991) o Burke y Yuen (1995), entre otros). En esta aplicación a datos reales, se ha elegido un contraste gráfico con el fin de facilitar su interpretación.

Uno de los contrastes gráficos más populares es el diagrama de dispersión suavizado de los residuos de Schoenfeld escalados (Grambsch y Therneau (1994)). En la figura 6.12 se muestra, además del diagrama, una estimación del correspondiente coeficiente para cada covariable, en principio dependiente del tiempo, $\beta(t)$, y su intervalo de confianza al 90%. Si la suposición de riesgos proporcionales es cierta, $\beta(t)$ será una línea horizontal. La manifiesta no horizontalidad del parámetro $\beta(t)$, para las dos covariables consideradas, establece una evidencia de la violación de la hipótesis de riesgos proporcionales en la que se basa del modelo de Cox.

Respecto al coeficiente $\beta(t)$ para la variable Transplante, la pendiente positiva y su aproximación a cero indica que el efecto beneficioso, en la supervivencia, de haber recibido un transplante se disipa a medida que pasa el tiempo. Razonablemente, conforme aumenta el periodo de pertenencia al programa de transplante de corazón de Stanford, el desarrollo de la enfermedad avanza y el efecto ventajoso del transplante desaparece con el tiempo.

Por su parte, el beneficio de haber sido operado del corazón con anterioridad se presenta a partir del cuarto día de haber sido aceptado en el programa de Stanford, y únicamente durante algo menos de un año.

Bibliografía

- [1] Aalen, O.O. (1978). Nonparametric inference for a family of counting process. *The Annals of Statistics*, **6**, 701-726.
- [2] Abdushukurov, A.A. (1987). Estimation of the probability density and intensity function of the Koziol-Green model of random censoring. *Sankhya Series A*, **48**, 150-168.
- [3] Akritas, M.G. (1986). Bootstrapping the Kaplan-Meier estimator. *Journal of the American Statistical Association*, **81**, 1032-1038.
- [4] Andersen, P.K. (1982). Testing goodness-of-fit of Cox's regression and life model. *Biometrics*, **38**, 67-77.
- [5] Andersen, P.K., Borgan, O., Gill, R.D. y Keiding, N. (1993). *Statistical models based on counting process*. Springer-Verlag, New York.
- [6] Arjas, E. (1988). A graphical method for assessing goodness-of-fit in Cox's proportional hazards model. *Journal of the American Statistical Association*, **83**, 204-212.
- [7] Bae, J. y Kim, S. (2003a). The uniform law of large numbers for the Kaplan-Meier integral process. *Bulletin of the Australian Mathematical Society*, **67**, 459-465.
- [8] Bae, J. y Kim, S. (2003b). The uniform central limit theorem for the Kaplan-Meier integral process. *Bulletin of the Australian Mathematical Society*, **67**, 467-480.
- [9] Bae, J. y Kim, S. (2003c). The empirical LIL for the Kaplan-Meier integral process. *Bulletin of the Korean Mathematical Society*, **40**, 269-279.
- [10] Billingsley, P. (1968). *Convergence of probability measures*. John Wiley & Sons, New York-London-Sydney.

-
- [11] Blum, J.R. y Susarla, V. (1980). Maximal deviation theory of density and failure rate function estimates based on censored data. *Journal of Multivariate Analysis*, **5**, 213-222.
- [12] Bowman, A.W. (1984). An alternative method of cross-validation for the smoothing of density estimates. *Biometrika*, **71**, 353-360.
- [13] Breslow, N. y Crowley, J. (1974). A large sample study of the life table and product limit estimates under random censorship. *The Annals of Statistics*, **2**, 437-453.
- [14] Burke, M.D. y Yuen, K.C. (1995). Goodness-of-fit tests for the Cox model via bootstrap method. *Journal of Statistical Planning and Inference*, **47**, 237-256.
- [15] Cao, R. (1988). *El bootstrap con datos censurados*. Memoria de Licenciatura. Universidad de Santiago de Compostela.
- [16] Cao, R. (1993). Bootstrapping the mean integrated squared error. *Journal of Multivariate Analysis*, **45**, 137-160.
- [17] Cao, R., Cuevas, A. y González Manteiga, W. (1994). A comparative study of several smoothing methods in density estimation. *Computational Statistics and Data Analysis*, **17**, 153-176.
- [18] Cao, R., López de Ullibarri, I., Janssen, P. y Veraverbeke, N. (2005) Presmoothed Kaplan-Meier and Nelson-Aalen estimators. *Journal of Nonparametric Statistics*, **17**, 31-56.
- [19] Chen, Y.Y., Hollander, M. y Langberg, N.A. (1982). Small-sample results for the Kaplan-Meier estimator. *Journal of the American Statistical Association*, **77**, 141-144.
- [20] Cheng, P.E. y Lin, G.D. (1987). Maximum likelihood estimation of a survival function under the Koziol-Green proportional hazards model. *Statistics & Probability Letters*, **5**, 75-80.
- [21] Cox, D.R. (1972). Regression models and life tables (with discussion). *Journal of the Royal Statistical Society Series B*, **34**, 187-220.
- [22] Crowley, J. y Hu, M. (1977). Covariance analysis of heart transplant survival data. *Journal of the American Statistical Association*, **72**, 27-36.

-
- [23] Csörgö, M., Gombay, E. y Horváth, L. (1991). Central limit theorems for L_p distances of kernel estimators of densities under random censorship. *Annals of Statistics*, **19**, 1813–1831.
- [24] Csörgö, S. y Horváth, L. (1983). The rate of the strong uniform consistency for the product-limit estimator. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, **62**, 411-426.
- [25] Deheuvels, P. y Einmahl, J.H. (2000). Functional limit laws for the increments of Kaplan-Meier product-limit processes and applications. *The Annals of Probability*, **28**, 1301–1335.
- [26] De la Peña, V. y Giné, E. (1999). *Decoupling. From dependence to independence*. Probability and its Applications, New York.
- [27] Devroye, L. (1983). The equivalence of weak, strong and complete convergence in L_1 for kernel density estimates. *The Annals of Statistics*, **11**, 896–904.
- [28] Devroye, L. (1997). Universal smoothing factor selection in density estimation: theory and practice. With discussion and a rejoinder by the author. *Test*, **6**, 223–320.
- [29] Diehl, S. y Stute, W. (1988). Kernel density and hazard function estimation in the presence of censoring. *Journal of Multivariate Analysis*, **25**, 299-310.
- [30] Dikta, G. (1998). On semiparametric random censorship models. *Journal of Statistical Planning and Inference*, **66**, 253-279.
- [31] Dikta, G. (2000). The strong law under semiparametric random censorship models. *Journal of Statistical Planning and Inference*, **83**, 1-10.
- [32] Dikta, G. (2001). Weak representation of the cumulative hazard function under semiparametric random censorship models. *Statistics*, **35**, 395–409.
- [33] Dvoretzky, A., Kiefer, J. y Wolfowitz, J. (1956). Asymptotic minimax character of the sample distribution function and of the classical multinomial estimator. *The Annals of Mathematical Statistics*, **27**, 642–669.
- [34] Efron, B. (1967). The two-sample problem with censored data. *Proceedings Fifth Berkeley Symposium Math. Statistics & Probability*, **4**, 831-853.
- [35] Efron, B. (1979). Bootstrap methods: another look at the jackknife. *The Annals of Statistics* **7**, 1–26.

-
- [36] Efron, B. (1981). Censored data and the bootstrap. *Journal of the American Statistical Association*, **76**, 312–319.
- [37] Efron, B. y Tibshirani, R.J. (1993). *An introduction to the bootstrap*. Monographs on Statistics and Applied Probability, 57. Chapman and Hall, New York.
- [38] Fan, J. (1992). Design-adaptive nonparametric regression. *Journal of the American Statistical Association*, **87**, 998–1004.
- [39] Fan, J. (1993). Local linear regression smoothers and their minimax efficiency. *The Annals of Statistics*, **21**, 196–216.
- [40] Fan, J. y Gijbels, I. (1992). Variable bandwidth and local linear regression smoothers. *The Annals of Statistics*, **20**, 2008–2036.
- [41] Fan, J. y Gijbels, I. (1996). *Local polynomial modelling and its applications*. Monographs on Statistics and Applied Probability. 66. Chapman and Hall, London.
- [42] Fleming, T.R. y Harrington, D.P. (1991). *Counting process and survival analysis*. Wiley, New York.
- [43] Földes, A. y Retjö, L. (1980). Strong uniform consistency for nonparametric estimation based on censored data. *The Annals of Statistics*, **9**, 122–129.
- [44] Földes, A. y Retjö, L. (1981). A LIL type result for the product limit estimator. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, **56**, 75–86.
- [45] Földes, A., Retjö, L. y Winter, B.B. (1981). Strong consistency properties of nonparametric estimators for randomly censored data. II Estimation of density and failure rate. *Periodica Mathematica Hungarica*, **12**, 15–29.
- [46] Gannoun, A. y Saracco, J. (2002). A new proof of strong consistency of kernel estimation of density function and mode under random censorship. *Statistics & Probability Letters*, **59**, 61–66.
- [47] Ghorai, J.K. y Pattanaik, L.M. (1990). L_1 -consistency of the kernel density estimators based on randomly right censored data. *Communications in Statistics. Theory and Methods*, **19**, 2853–2870.
- [48] Gijbels, I. y Wang, J.L. (1993). Strong representations of the survival function for truncated and censored data with applications. *Journal of Multivariate Analysis*, **47**, 210–229.

-
- [49] Gill, R.D. (1980). Censoring and Stochastic Integrals. *Mathematical Centre Tracts*. Amsterdam: Mathematisch Centrum, 124.
- [50] Gill, R.D. (1983). Large sample behaviour of the product-limit estimator on the whole line. *The Annals of Statistics*, **11**, 19-58.
- [51] Giné, E. y Guillou, A. (1999). Laws of the iterated logarithm for censored data. *The Annals of Probability*, **27**, 2042-2067.
- [52] González Manteiga, W., Cao, R. y Marron, J.S. (1996). Bootstrap selection of the smoothing parameter in nonparametric hazard rate estimation. *Journal of the American Statistical Association*, **91**, 1130-1140.
- [53] Grambsch, P. y Therneau, T. (1994). Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika*, **81**, 515-26.
- [54] Guerts, J.H.L. (1985). Some small-sample non-proportional hazards results for the Kaplan-Meier estimator. *Statistica Neerlandica*, **39**, 1-13.
- [55] Guerts, J.H.L. (1987). On the small-sample performance of Efron's and Gill's version of the product limit estimator under proportional hazards. *Biometrics*, **43**, 683-692.
- [56] Hájek, J. y Sidák, Z. (1967). *Theory of rank tests*. Academic Press, New York-London.
- [57] Hall, W.J. y Wellner, J.A. (1980). Confidence bands for a survival curve from censored data. *Biometrika*, **67**, 133-143.
- [58] Hall, P. (1992). *The bootstrap and Edgeworth expansion*. Springer Series in Statistics. Springer-Verlag, New York.
- [59] Härdle, W. (1990). *Applied Nonparametric Regression*. Cambridge University Press, Boston.
- [60] Johansen, S. (1978). The product limit estimator as maximum likelihood estimator. *Scandinavian Journal of Statistics*, **5**, 195-199.
- [61] Jones, M.C., Marron, J.S. y Sheather, S.J. (1996). A brief survey of bandwidth selection for density estimation. *Journal of the American Statistical Association*, **91**, 401-407.
- [62] Kaplan, E.L. y Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*, **53**, 457-481.

- [63] Karunamuni, R.J. y Yang, S. (1991). Weak and strong uniform consistency rates of kernel density estimates for randomly censored data. *Canadian Journal of Statistics*, **19**, 349-359.
- [64] Klein, J.P. (1991). Small-sample moments of some estimators of the variance of the Kaplan-Meier and Nelson-Aalen estimators. *Scandinavian Journal of Statistics*, **18**, 333-340.
- [65] Kuhn, J.W. y Padgett, W.J (1997). Local bandwidth selection for kernel density estimation from right-censored data based on asymptotic mean absolute error. Proceedings of the Second World Congress of Nonlinear Analysis, Part 7 (Athens, 1996). *Nonlinear Analysis*, **7**, 4375-4384.
- [66] Lin, D.Y. y Wei, L.J. (1991). Goodness-of-fit tests for the general Cox regression model. *Statistica Sinica*, **1**, 1-17.
- [67] Lo, S.H., Mack, Y.P. y Wand, J.L (1989). Density and hazard rate estimation for censored data via strong representation of the Kaplan-Meier estimator. *Probability Theory and Related Fields*, **80**, 461-473.
- [68] Lo, S.H. y Singh, K. (1986). The product-limit estimator and the bootstrap: some asymptotic representations. *Probability Theory and Related Fields*, **71**, 455-465.
- [69] López de Ullibarri, I. (2004). *Estimación presuavizada de las funciones de riesgo acumulativa y no acumulativa con datos censurados*. Tesis de doctorado. Universidad de Santiago de Compostela.
- [70] Louani, D. (1998). On the asymptotic normality of the kernel estimators of the density function and its derivatives under censoring. *Communications in Statistics. Theory and Methods*, **27**, 2909-2924.
- [71] Mack, Y.P. y Silverman, B.W. (1982). Weak and strong uniform consistency of kernel regression estimates. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, **61**, 405-415.
- [72] Marron, J.S. (1992). *Bootstrap bandwidth selection. Exploring the limits of bootstrap*. Wiley, New York.
- [73] Marron, J.S. y Padgett, W.J. (1987). Asymptotically optimal bandwidth selection for kernel density estimators from randomly right-censored samples. *The Annals of Statistics*, **15**, 1520-1535.

-
- [74] Meier, P. (1975). *Estimation of a distribution function from incomplete observations*. Perspectives in Probability and Statistics, ed. J. Gani, 67-87, Sheffield, England: Applied Probability Trust.
- [75] Mielniczik, J. (1986). Some asymptotic properties of kernel estimators of a density function in case of censored data. *The Annals of Statistics*, **14**, 766-773.
- [76] Nadaraya, È.A. (1964). On estimating regression. *Theory of Probability and its Applications*, **10**, 186-190.
- [77] Nadaraya, È.A. (1989). *Nonparametric estimation of probability densities and regression curves*. Mathematics and its Applications (Soviet Series), 20. Kluwer Academic Publishers Group, Dordrecht.
- [78] Nelson, W. (1972). Theory and applications of hazard plotting for censored failure data. *Technometrics*, **14**, 945-965.
- [79] Parzen, E. (1962). On estimation of a probability density function and mode. *The Annals of Mathematical Statistics*, **33**, 1065-1076.
- [80] Patil, P.N. (1993). On the least squares cross-validation bandwidth in hazard rate estimation. *The Annals of Statistics*, **21**, 1792-1810.
- [81] Petrov, V.V. (1995). *Limit Theorems of Probability Theory*. Oxford University Press.
- [82] Phadia, E.G. y Shao, P.Y. (1999). Exact moments of the product limit estimator. *Statistics & Probability Letters*, **41**, 277-286.
- [83] Reid, N. (1981). Estimating the median survival time. *Biometrika*, **68**, 601-608.
- [84] Rosenblatt, M. (1956). Remarks on some nonparametric estimates of a density function. *The Annals of Mathematical Statistics*, **27**, 832-837.
- [85] Rudemo, M. (1982). Empirical choice of histograms and kernel density estimators. *Scandinavian Journal of Statistics*, **9**, 65-78.
- [86] Ruppert, D., Sheather, S.J. y Wand, M.P. (1995). An effective bandwidth selector for local least squares regression. *Journal of the American Statistical Association*, **90**, 1257-1270.
- [87] Ruppert, D. y Wand, M.P. (1994). Multivariate weighted least squares regression. *The Annals of Statistics*, **22**, 1346-1370.

- [88] Sánchez Sellero, C. (2001). *Inferencia Estadística en Datos con Censura y/o Truncamiento*. Tesis de doctorado, Universidad de Santiago de Compostela.
- [89] Sánchez Sellero, C., González Manteiga, W. y Cao, R. (1999). Bandwidth selection in density estimation with truncated and censored data. *Annals of the Institute of Statistical Mathematics*, **51**, 51–70.
- [90] Sánchez Sellero, C., González Manteiga, W. y Van Keilegom, I. (2005). Uniform representation of product-limit integrals with applications. Próxima publicación en *Scandinavian Journal of Statistics*.
- [91] Scholz, F.W. (1980). Towards a unified definition of maximum likelihood. *Canadian Journal of Statistics*, **8**, 193-203.
- [92] Schuster, E.F. (1969). Estimation of a probability density function and its derivatives. *The Annals of Mathematical Statistics*, **40**, 1187–1195.
- [93] Schuster, E.F. (1985). Incorporating support constraints into nonparametric estimators of densities. *Communications in Statistics - Theory and Methods*, **14**, 1123-1136.
- [94] Scott, D.W. y Wand, M.P. (1991). Feasibility of multivariate density estimates. *Biometrika*, **78**, 197–205.
- [95] Shorack, G.R. y Wellner, J.A. (1986). *Empirical processes with applications to statistics*. Wiley, New York.
- [96] Silverman, B.M. (1978). Weak and strong uniform consistency of the kernel estimate of a density and its derivatives. *The Annals of Statistics*, **6**, 177-184.
- [97] Silverman, B.M. (1986). *Density Estimation for Statistics and Data Analysis*. Chapman and Hall, London.
- [98] Stone, M. (1974). Cross-validatory choice and assessment of statistical predictions (with discussion). *Journal of the Royal Statistical Society, Series B*, **36**, 111–147.
- [99] Stute, W. (1985). Optimal bandwidth selection in pointwise density and hazard rate estimation when censoring is present. *Technical report*, Universität Gießen.
- [100] Stute, W. (1994). The bias of Kaplan-Meier integrals. *Scandinavian Journal of Statistics*, **21**, 475-484.

-
- [101] Stute, W. (1995). The central limit theorem under random censorship. *The Annals of Statistics*, **23**, 422-439.
- [102] Stute, W. (1996). The jackknife estimate of variance of a Kaplan-Meier integral. *The Annals of Statistics*, **24**, 2679-2704.
- [103] Stute, W. y Wang, J. L. (1993). A strong law under random censorship. *The Annals of Statistics*, **21**, 1591-1607.
- [104] Wang, M.C. (1987). Product-limit estimates: A generalized maximum likelihood study. *Communications in Statistics - Theory and Methods*, **16**, 3117-3132.
- [105] Watson, G.S. (1964). Smooth regression analysis. *Shankya Series A*, **26**, 359-372.
- [106] Xiang, X. (1994). Law of the logarithm for density and hazard rate estimation for censored data. *Journal of Multivariate Analysis*, **49**, 278-286.
- [107] Zhang, B. (1996). Some asymptotic results for kernel density estimation under random censorship. *Bernoulli*, **2**, 183-198.
- [108] Zhang, B. (1998). A note on the integrated square errors of kernel density estimators under random censorship. *Stochastic Processes and their Applications*, **75**, 225-234.
- [109] Zhang, B. (1999). A law of the iterated logarithm for kernel density estimators under random censorship. *Far East Journal of Theoretical Statistics*, **3**, 171-185.
- [110] Zhu, L.X., Yuen, K.C. y Tang, N.Y. (2002). Resampling methods for testing a semiparametric random censorship model. *Scandinavian Journal of Statistics*, **29**, 111-123.
- [111] Ziegler, S. (1995). *Ein modifizierter Kaplan-Meier Schätzer*. Dissertation, Mathematisches Institut, Justus-Liebig-Universität Gießen.

