

Algunos problemas en Inferencia no Paramétrica:
Estimación de Distribuciones ordenadas por Picudez
y Supervivencia Bivariada con Censura

Tesis que para obtener el grado de Doctor en Ciencias con orientación en
Probabilidad y Estadística, presenta

José Luis Batún Cutz

Sinodales:

Dr Victor Manuel Pérez-Abreu Carrión (Cimat)

Dra. Graciela María González Farías (Cimat)

Dr Javier Rojo Jiménez, Asesor (Rice University)

Dr José Aurelio Villaseñor Alva (Colegio de Posgraduados)

Dr Joaquín Ortega Sánchez (Cimat)

Guanajuato, Gto, 20 de Mayo de 2005.

Contenido

1	Introducción	5
1.1	Orden Estocástico	6
1.2	Función de Supervivencia Bivariada	9
2	Estimación bajo restricciones de picudez	14
2.1	El estimador: Una muestra	15
2.1.1	Propiedades asintóticas del estimador F_n^* .	19
2.2	El estimador: Dos muestras	21
2.2.1	Propiedades asintóticas del estimador $F_{nm}^*(x)$	23
2.3	Estimación bajo Censura.	24
2.4	Caso simétrico	24
2.4.1	Definición del estimador	26
2.4.2	Propiedades asintóticas de los estimadores	

	$F_{n,1}$ y $F_{n,2}$	35
2.4.3	El caso dos muestras.	38
2.4.4	El caso de una muestra con censura	39
2.5	Simulaciones	40

3 Estimación de la Función de Supervivencia Bi-

	variada	45
3.1	El estimador de Dabrowska	47
3.2	El estimador de Prentice y Cai	51
3.3	Nuevos estimadores de la función de supervivencia S	52
3.4	Estudios de Simulación comparando el MSE de los nuevos estimadores con los existentes.	55

Agradecimientos

A mi Asesor Dr Javier Rojo Jiménez, por su acertada dirección, consejos y paciencia durante todo el tiempo en que se realizó este trabajo de investigación y que dió como resultado esta tesis.

A mis sinodales por el trabajo que hicieron en la revisión y sus sugerencias que ayudaron a mejorar el trabajo.

Al Cimat, a través del Dr Victor Pérez-Abreu Carrión, por su apoyo en todo momento a lo largo del Doctorado.

A la Dra Graciela González Farías, quien a través de su proyecto de Investigación No 39017E de Conacyt, "Inferencia y Aplicaciones en Análisis Multivariado bajo Distribuciones No-Estándar ", me apoyó en mi estancia en Rice University.

Al Departamento de Estadística de Rice University, por las facilidades dadas durante mi estancia.

Capítulo 1

Introducción

En esta tesis se consideran dos problemas de estimación no paramétrica: la estimación de la función de supervivencia bivariada bajo censura aleatoria independiente y la estimación de funciones de distribución bajo restricciones de picudez, una relación de orden que se definirá mas adelante. En primer lugar se discute el problema de la estimación de la función de distribución cuando hay restricciones de picudez, proponiéndose un estimador que satisface las condiciones de picudez y que además es uniformemente consistente, además de satisfacer teoremas del límite central. En segundo término, si además de restricciones de picudez, se incorpora simetría con respecto a un punto conocido, se presentan 2 nuevos estimadores con estas características, que son uniformemente consistentes, mostrándose su distribución asintótica, así como resultados sobre convergencia débil del proceso asociado. Estudios de simulación avalan que estos nuevos estimadores son mejores que la distribución empírica o el estimador Kaplan-Meier, según sea el caso.

Por último, dando un salto al mundo bivariado, se describe el problema de la estimación de la función de supervivencia bivariada bajo censura independiente. Este problema ha sido estudiado por varios autores y por mucho tiempo, sin que haya una solución satisfactoria. Aquí se presentan dos nuevos estimadores de la función de supervivencia bivariada, cuyo desempeño es mejor que los hasta ahora propuestos, y siempre son funciones de supervivencia propias, es decir, asignan masa positiva y son monótonos no crecientes en cada componente, características que los estimadores propuestos hasta ahora en la literatura no poseen.

En cada caso se presentan estudios de simulación comparando mediante el error cuadrático medio, el desempeño de los estimadores nuevos propuestos en esta tesis, con los respectivos estimadores hasta ahora dados en la literatura correspondiente.

El trabajo está organizado como sigue: en este capítulo introductorio, se presentan las definiciones de orden estocástico y los conceptos básicos relacionados con la función de supervivencia bivariada. En el Capítulo 2, se considera el problema de estimación de una función de distribución F más picuda que otra distribución G , proponiéndose estimadores que satisfacen las restricciones de picudez, bajo las condiciones de G conocida, G no conocida, así como bajo censura en los datos provenientes de F . Propiedades de consistencia, así como distribuciones asintóticas de los estimadores son presentados. En la sección 2.4 se agrega una restricción a las funciones de distribución de interés, suponiendo que son simétricas alrededor de un punto conocido. En este caso se presentan dos estimadores que son funciones de distribución simétricas y que satisfacen las restricciones de picudez. Ambos estimadores son uniformemente consistentes, y son proyección de la función de distribución empírica bajo la métrica uniforme, sobre ciertos conjuntos convexos. Se discuten sus distribuciones asintóticas y se prueba la convergencia de los procesos obtenidos con estos estimadores.

En el Capítulo 3 se estudian dos estimadores de la función de supervivencia bivariada bajo censura independiente, así como algunas representaciones de la función de supervivencia bivariada en términos de funciones de riesgo bivariadas y las marginales. Se proponen dos nuevos estimadores, basados en correcciones al estimador de Dabrowska, pues este último no siempre asigna masa positiva a todo rectángulo o no es monótono decreciente en cada componente. Estudios de simulación comparando estos nuevos estimadores con respecto a los ya presentes en la literatura son presentados, mostrando un mejor desempeño de los nuevos estimadores en cuanto al error cuadrático medio.

1.1 Orden Estocástico

Un orden parcial estocástico se entiende como una relación de orden parcial entre funciones de distribución. Es decir una relación binaria definida en un conjunto de funciones de distribución, que es reflexiva, antisimétrica y transitiva. Actualmente diversos problemas en Ingeniería,

Medicina, Finanzas, Genética (entre otros) han utilizado los distintos tipos de orden estocástico hasta ahora estudiados. En Oja (1981) se describen conceptos de localización, escala, sesgo y curtosis, mencionando entre otras cosas su relación con el orden estocástico. Cabe aclarar que los primeros tres ordenes que se definirán mas adelante corresponden a cambios de localización, mientras que los dos últimos (y en particular el de picudez) corresponden a cambios de escala (o dispersión). En Lehmann y Rojo (1992) se presentan caracterizaciones de distintos ordenes en términos del maximal invariante bajo un grupo de transformaciones monótonas. Resultados sobre estimación, pruebas de hipótesis, intervalos de confianza, y distribuciones asintóticas, han sido el común de los trabajos realizados y que se pueden encontrar en la literatura. A continuación describiremos algunos ejemplos de relaciones de orden parcial estocástico:

Definición 1 *Orden Estocástico Usual.* Sean F y G funciones de distribución. Entonces $F \geq_{st} G$ (se lee F es estocásticamente mayor que G) si y solo si,

$$F(x) \leq G(x) \text{ para todo } x \in \mathbb{R}$$

Algunos trabajos sobre estimación de la función de distribución F cuando se tiene esta restricción de orden son Dykstra (1982) y Dykstra y Feltz (1989), donde se presenta el estimador de máxima verosimilitud no paramétrico y su consistencia, así como generalizaciones a cuando se tienen mas de dos funciones de distribuciones ordenadas. El Barmi y Rojo (1997) consideran el caso cuando las funciones de distribución son multinomiales. En Rojo y Ma (1996) y Rojo (1995) se aborda un enfoque diferente en la construcción del estimador, dando origen a un nuevo y mejor estimador (en cuanto al error cuadrático medio) que el de máxima verosimilitud. Este estimador también es uniformemente consistente y su distribución asintótica es de tipo normal.

Definición 2 *Orden Tasa de Riesgo.* La función de distribución F es mayor que G con respecto a la tasa de riesgo ($F \geq_{hr} G$), si y sólo si la función

$$t \mapsto \frac{\overline{F}(t)}{\overline{G}(t)} \text{ es creciente,}$$

con $\overline{F} = 1 - F$ y $\overline{G} = 1 - G$, las funciones de supervivencia respectivas.

Definición 3 *Orden Razón de Verosimilitud.* Una función de distribución F es mayor en razón de verosimilitud que una distribución G , ($F \geq_{lr} G$), si y sólo si

$$f(s)g(t) \leq f(t)g(s) \quad \text{para todo } s \leq t.$$

con f y g las densidades de F y G respectivamente.

Esta definición solo aplica para funciones de distribución que tienen una densidad.

Definición 4 *Orden de Picudez.* Definida en el conjunto de las funciones de distribución, se dice que F (en a) es más picuda que G (en b), si y sólo si

$$F(x+a) - F(-x+a) \geq G(x+b) - G(-x+b) \quad \text{para todo } x \geq 0. \quad (1.1)$$

Esto se denota por $F \geq_p G$. La condición (1.1) se expresa también como $P[|X-a| < x] \geq P[|Y-b| < x]$ (equivalentemente, $P[|X-a| \geq x] \leq P[|Y-b| \geq x]$) para $X \sim F$, $Y \sim G$.

Este orden fue introducido por primera vez en Birnbaum (1948). En Bickel & Lehmann (1976) se menciona este orden con F y G simétricas, y a y b los respectivos centros de simetría. En este caso, la condición (1.1) es equivalente a

$$G^{-1}(v) - G^{-1}\left(\frac{1}{2}\right) \underset{\geq}{\overset{\leq}{\leq}} F^{-1}(v) - F\left(\frac{1}{2}\right) \quad \text{para } v \underset{>}{<} \frac{1}{2}$$

Definición 5 *Orden Dispersivo.* Una función de distribución F se dice que es mayor que G con respecto al orden dispersivo ($F \geq_{disp} G$), si y sólo si

$$F^{-1}(t) - F^{-1}(s) \geq G^{-1}(t) - G^{-1}(s) \quad \text{para todo } 0 < s < t < 1.$$

Este orden fue introducido por primera vez en Fraser (1976) y en Bickel & Lehmann(1976), y se interpreta como dos cuantiles de la distribución F están separados en al menos la misma distancia que los correspondientes cuantiles de la distribución G . Una caracterización de este orden es que $F \geq_{disp} G$ si y solo si, $G^{-1}F - I$ es creciente (Doksum, 1969), donde I es la función identidad.

En este trabajo el orden de picudez es el más importante y más adelante se darán ejemplos de distribuciones con estas características, así como sus propiedades. Ahora se hablará brevemente de la teoría de la función de supervivencia bivariada.

1.2 Función de Supervivencia Bivariada

Sea $\mathbf{T} = (T_1, T_2)$ un vector de tiempos de vida, es decir, con componentes variables aleatorias no negativas, función de distribución conjunta F y funciones de distribución marginales F_1 , y F_2 . La función de supervivencia S de \mathbf{T} , se define como $S(s, t) = P [T_1 > s, T_2 > t]$ para $0 \leq s, t$. Estas funciones están relacionadas mediante

$$\begin{aligned} S(s, t) &= F(s, t) - F_1(s) - F_2(t) + 1, \\ F(s, t) &= S(s, t) - S_1(s) - S_2(t) + 1 \end{aligned}$$

donde S_1 y S_2 son las funciones de supervivencia marginales. Así, la estimación de la función de distribución bivariada o la función de supervivencia bivariada se pueden considerar como problemas equivalentes. En este capítulo presentamos teoría relacionada con funciones de supervivencia bivariada, y que ha sido utilizada para desarrollar estimadores para $S(s, t)$, específicamente, mostraremos algunas de las representaciones de la función de supervivencia en términos de las marginales y funciones que incorporan la dependencia entre las variables.

Así como en el caso unidimensional la función de supervivencia está relacionada con la función de riesgo acumulada Λ mediante la igualdad $\Lambda_i(t) = -S_i(dt)/S_i(t-)$, en el caso bivariado, también se tiene una relación de ese tipo. Siguiendo Dabrowska (1988) tenemos la siguiente definición:

Definición 6 Una función de riesgo acumulada bivariada es un vector $\mathbf{\Lambda} = (\Lambda_{10}, \Lambda_{01}, \Lambda_{11})$ con componentes definidos por

$$\begin{aligned} \Lambda_{11}(ds, dt) &= \frac{P [T_1 \in [s, s + ds], T_2 \in [t, t + dt]]}{P [T_1 \geq s, T_2 \geq t]} = \frac{S(ds, dt)}{S(s-, t-)} \\ \Lambda_{10}(ds, t) &= \frac{P [T_1 \in [s, s + ds], T_2 > t]}{P [T_1 \geq s, T_2 > t]} = -\frac{S(ds, t)}{S(s-, t)} \\ \Lambda_{01}(s, dt) &= \frac{P [T_1 > s, T_2 \in [t, t + dt]]}{P [T_1 > s, T_2 \geq t]} = -\frac{S(s, dt)}{S(s, t-)} \end{aligned}$$

y

$$\Lambda_{10}(0, t) = \Lambda_{01}(s, 0) = \Lambda_{11}(0, 0) = 0.$$

Note que $\Lambda_{10}(ds, 0) = \Lambda_1(ds)$ y $\Lambda_{01}(s, dt) = \Lambda_2(dt)$, es decir, son las funciones de riesgo acumuladas univariadas. Como caso particular, si F tiene densidad f , entonces, $\Lambda_{11}(ds, dt) = \lambda_{11}(s, t)dsdt$, $\Lambda_{10}(ds, t) = \lambda_{10}(s, t)ds$ y $\Lambda_{01}(s, dt) = \lambda_{01}(s, t)dt$, con

$$\begin{aligned}\lambda_{11}(s, t) &= \frac{f(s, t)}{S(s, t)}, \\ \lambda_{10}(s, t) &= \frac{1}{S(s, t)} \int_s^\infty f(s, v)dv \\ \lambda_{01}(s, t) &= \frac{1}{S(s, t)} \int_s^t f(u, t)du.\end{aligned}$$

En este caso, tenemos las siguientes expresiones para Λ :

$$\begin{aligned}\Lambda_{11}(s, t) &= \frac{1}{S(s, t)} \int_0^s \int_0^t f(u, v)dudv, \\ \Lambda_{10}(s, t) &= \frac{1}{S(s, t)} \int_0^s \int_s^\infty f(u, v)dvdu, \\ \Lambda_{01}(s, t) &= \frac{1}{S(s, t)} \int_0^t \int_s^\infty f(u, v)dvdu.\end{aligned}$$

En adelante, utilizaremos la notación $h(\Delta s, t) = h(s, t) - h(s-, t)$, $h(s, \Delta t) = h(s, t) - h(s, t-)$ y $h(\Delta s, \Delta t) = h(s, t) - h(s-, t) - h(s, t-) + h(s-, t-)$, para cualquier función h bivariada.

En Dabrowska (1988) se presenta una representación de la función de supervivencia bivariada en términos de la función de riesgo acumulada definida anteriormente. Esta representación se basa en una "descomposición" de la función $A(s, t) = \ln S(s, t)$ en los siguientes conjuntos:

$$\begin{aligned}E_1 &= \{(s, t) : A(s, t) < 0, A(\Delta s, t) = A(s, \Delta t) = 0\}, \\ E_2 &= \{(s, t) : A(s, t) < 0, A(\Delta s, t) < 0, A(\Delta s, \Delta t) = 0\}, \\ E_3 &= \{(s, t) : A(s, t) < 0, A(s, \Delta t) < 0, A(\Delta s, \Delta t) = 0\}, \\ E_4 &= \{(s, t) : A(s, t) < 0, A(\Delta s, \Delta t) > 0\},\end{aligned}$$

que corresponden a el soporte de la componente puramente continua, los soportes de las componentes que tienen discontinuidades que viven a lo largo de líneas ortogonales a los ejes coordenados y el soporte de la componente puramente discreta para la función A .

Proposición 7 (Dabrowska, 1988) Para (s, t) tales que $S(s, t) > 0$, tenemos

$$S(s, t) = S(s, 0), S(0, t) \prod_{i=1}^4 B_i(s, t),$$

con

$$B_i(s, t) = \exp \left\{ - \int_0^s \int_0^t I[(u, v) \in E_i] L(du, dv) \right\} \quad i = 1, 2, 3,$$

$$B_4(s, t) = \prod_{u \leq s} \prod_{v \leq t} I[(u, v) \in E_4] (1 - L(\Delta u, \Delta v)),$$

y la función L dada por

$$L(ds, dt) = \frac{\Lambda_{10}(ds, t-) \Lambda_{01}(s-, dt) - \Lambda_{11}(ds, dt)}{[1 - \Lambda_{10}(\Delta s, t-)] [1 - \Lambda_{01}(s-, \Delta t)]} \quad (1.2)$$

Demostración. Ver prueba en Dabrowska (1988). ■

La función L incorpora la relación de dependencia que existe entre los tiempos de falla, al conjuntar las distintas funciones de riesgo acumuladas (univariadas y bivariada). Veamos el siguiente ejemplo.

Ejemplo 8 Oakes (1982). Sean F y H funciones de distribución con densidades f y h . Una función de supervivencia bivariada está dada por $S(s, t) = (\overline{F}(s)^{1-\theta} + \overline{H}(t)^{1-\theta} - 1)^{-1/(\theta-1)}$, $\theta > 1$. Su densidad conjunta es $f(s, t) = \theta f(s)h(t) (\overline{F}(s)\overline{H}(t))^{-\theta} D^{-(2+1/(\theta-1))}(s, t)$, con $D(s, t) = \overline{F}(s)^{1-\theta} + \overline{H}(t)^{1-\theta} - 1$. Es fácil ver que las expresiones para las funciones de riesgo están dadas por :

$$\lambda_{11}(s, t) = \theta f(s)h(t) (\overline{F}(s)\overline{H}(t))^{-\theta} D^{-2}(s, t),$$

$$\lambda_{10}(s, t) = f(s)\overline{F}(s)^{-\theta} D^{-1}(s, t),$$

$$\lambda_{01}(s, t) = h(t)\overline{H}(t)^{-\theta} D^{-1}(s, t),$$

y la función L resultante es:

$$\begin{aligned} L(s, t) &= \ln \left[\overline{F}(s) \overline{H}(t) D^{1/(\theta-1)}(s, t) \right] \\ &= \ln \left[\frac{\overline{F}(s) \overline{H}(t)}{S(s, t)} \right]. \end{aligned}$$

Observación 9 En general, si la función de supervivencia S tiene una densidad conjunta f y funciones de supervivencia marginales \overline{F} y \overline{H} , entonces

$$L(s, t) = \ln \left[\frac{\overline{F}(s) \overline{H}(t)}{S(s, t)} \right].$$

En el caso univariado, existe una relación entre la función de supervivencia y la función de riesgo acumulada, expresada en términos de la integral producto (ver Gill y Johansen, 1987, Gill 1992a y Gill 1992b). También en el caso bivariado (y multivariado), es posible tener una representación en términos de la integral producto como sigue:

$$S(s, t) = \prod_{u \leq s} (1 - \Lambda_{10}(du, 0)) \prod_{v \leq t} (1 - \Lambda_{01}(0, dv)) \prod_{u \leq s, v \leq t} (1 - L(du, dv)) \quad (1.3)$$

donde los primeros dos factores corresponden a las funciones de supervivencia marginales expresadas como integral producto de sus respectivas funciones de riesgo acumuladas, y el tercer factor está definido como

$$\prod_{u \leq s, v \leq t} (1 - L(du, dv)) = \lim_{\substack{\max |u_i - u_{i-1}| \rightarrow 0 \\ \max |v_j - v_{j-1}| \rightarrow 0}} \prod_{i, j} (1 - L((u_{i-1}, u_i] \times (v_{j-1}, v_j])),$$

donde $\{u_i\}$ es una partición finita del intervalo $[0, s]$ y $\{v_j\}$ es una partición del intervalo $[0, t]$ y $L((u_{i-1}, u_i] \times (v_{j-1}, v_j]) = L(u_i, v_j) - L(u_{i-1}, v_j) - L(u_i, v_{j-1}) + L(u_{i-1}, v_{j-1})$.

Cuando S es una función continua, de la representación dada en la Proposición 7, se obtienen

las siguientes ecuaciones:

$$\begin{aligned} S(s, t) &= S(s, 0)S(0, t) \exp \{-L(s, t) + L(s, 0) + L(0, t) - L(0, 0)\} \\ L(s, t) &= L(s, 0) + L(0, t) - L(0, 0) - \ln \left\{ \frac{S(s, t)}{S(s, 0)S(0, t)} \right\}. \end{aligned}$$

Otra representación de la función de supervivencia es mediante la ecuación integral

$$S(s, t) = S(s, 0) + S(0, t) - 1 + \int_0^s \int_0^t S(u-, v-) \Lambda_{11}(du, dv)$$

con condiciones iniciales, $S(s, 0) = \prod_{u \leq s} (1 - \Lambda_{10}(du, 0))$, y $S(0, t) = \prod_{v \leq t} (1 - \Lambda_{01}(0, dv))$, cuya solución para S se expresa en términos de series de Peano. Como se ha visto, existen varias representaciones de la función de supervivencia bivariada en términos de las funciones de riesgo acumuladas y las marginales. Ahora veamos una representación más.

En Prentice-Cai (1992) se presenta la representación en términos de la función de covarianza de los procesos de conteo $N_i(u) = I[T_i \leq u]$, $i = 1, 2$, definida por

$$C(s, t) = E [dN_1(s) - \Lambda_{10}(T_1 \wedge s)] [dN_2(t) - \Lambda_{01}(T_2 \wedge t)], \quad (1.4)$$

al demostrar que ésta, junto con las marginales, caracterizan a la función de supervivencia bivariada mediante la ecuación:

$$S(s, t) = S(s, 0)S(0, t) \left[1 + \int_0^s \int_0^t (S(u, 0)S(0, v))^{-1} C(du, dv) \right] \quad (1.5)$$

Esta representación es utilizada para construir un estimador de la función de supervivencia bivariada, el cual se presentará más adelante. Cabe comentar que las representaciones dadas en la proposición (7) y en (1.5) son solo dos de tantas, pero son las más interesantes y su obtención involucra conocimiento de herramientas tales como integral producto, descomposición de Jordan para funciones acotadas, así como teoría de martingalas y procesos de conteo.

Capítulo 2

Estimación bajo restricciones de picudez

Bajo el esquema de orden por picudez entre dos funciones de distribución F y G , consideremos el problema de la estimación de una función de distribución F , con $F \geq_p G$ ($a = b = 0$), considerando distintos contextos. En este Capítulo se proponen estimadores para F cuando G es conocida, cuando G no es conocida, así como cuando ambas F y G son funciones de distribución simétricas. También se considera el caso cuando se tienen muestras censuradas. En cada caso, se presentan las propiedades asintóticas de los estimadores, así como estudios de simulación comparativos con otros estimadores propuestos en la literatura.

Algunos ejemplos de funciones de distribución ordenadas por picudez, son los siguientes:

Ejemplo 10 *En la familia de las distribuciones normales, tenemos $N(\mu, \sigma_1^2) \geq_p N(\nu, \sigma_2^2)$ con $a = \mu$ y $b = \nu$, si y solo si $\sigma_1 \leq \sigma_2$.*

Ejemplo 11 *La función de distribución Logística, denotada por $LOG(\sigma)$ está dada por*

$$F(x) = \frac{1}{1 + \exp\left(-\frac{x}{\sigma}\right)}$$

y $LOG(\sigma_1) \geq_p LOG(\sigma_2)$ con $a = b = 0$, si y solo si $\sigma_1 \leq \sigma_2$.

Ejemplo 12 La función de distribución doble exponencial ($DEXP(\sigma)$) está definida como

$$F(x) = \begin{cases} 1 - \frac{1}{2} \exp(-\frac{x}{\sigma}) & x \geq 0 \\ \frac{1}{2} \exp(\frac{x}{\sigma}) & x < 0 \end{cases}$$

y $DEXP(\sigma_1) \geq_p DEXP(\sigma_2)$ con $a = b = 0$, si y solo si $\sigma_1 \leq \sigma_2$.

Ejemplo 13 Este es un ejemplo con distribuciones no simétricas. Sea F la función de distribución definida por

$$F(x) = \begin{cases} \frac{1}{1 + \exp(-x)} & x \leq 0 \\ \frac{1}{1 + \exp(-\frac{x}{2})} & x \geq 0 \end{cases}$$

es decir, $F = LOG(1)I[x < 0] + LOG(2)I[x \geq 0]$. Si $G = LOG(\sigma_1)I[x < 0] + LOG(\sigma_2)I[x \geq 0]$, entonces $F \geq_p G$ si $\sigma_1 \geq 1$ y $\sigma_2 \geq 2$.

Como se mencionó anteriormente, la estimación de una función de distribución F más picuda que una distribución G , ha sido estudiada por ejemplo en El Barmi y Rojo (1997), Dykstra (1982), Dykstra y Feltz (1989), Rojo y Ma (1996), Rojo (1995), asumiendo que los puntos a y b alrededor de los cuales se presenta la picudez, son conocidos. En este trabajo mediante modificaciones a la distribución empírica, y suponiendo también que los puntos a y b son conocidos, se proponen estimadores para la función F cuando la función de distribución G es conocida, cuando no lo es y cuando la muestra de la función F es censurada. Este es el tema de la siguiente sección.

2.1 El estimador: Una muestra

En el problema de la estimación de una función de distribución F , con $F \geq_p G$ (con $a = b = 0$) y G conocida, sea X_1, \dots, X_n una muestra de la distribución F y F_n la función de distribución empírica correspondiente. Aun cuando F_n es el estimador no paramétrico natural para F , F_n

podría no satisfacer la condición (1.1), y una modificación de F_n en la forma apropiada, podría tener un mejor desempeño. Supongamos que F_n no satisface (1.1). En este caso podemos particionar el intervalo $[0, \infty)$ de tal forma que en algunos intervalos la desigualdad (1.1) es válida y en otros no lo sea. Sea $\mathcal{T} = \{|X_i| : i = 1, 2, \dots, n\}$ el conjunto de los valores absolutos de la muestra. Entonces para el máximo del conjunto \mathcal{T} , $T_{(n)}$ se satisface la relación $F_n(T_{(n)}) - F_n(-T_{(n)}) = 1 \geq G(T_{(n)}) - G(-T_{(n)})$, por lo que siempre existe un número $a_{2k-1} \leq T_{(n)}$, tal que $F_n(y) - F_n(-y) \geq G(y) - G(-y)$ para $y > a_{2k-1}$. Si hacemos $a_{2k} = \infty$ y

$$\begin{aligned} a_0 &= 0 \\ a_{2i-1} &= \inf \{y > a_{2i-2} : F_n(y) - F_n(-y) \geq G(y) - G(-y)\} \\ a_{2i} &= \inf \{y > a_{2i-1} : F_n(y) - F_n(-y) < G(y) - G(-y)\} \end{aligned} \quad (2.1)$$

para $i = 1, 2, \dots, k-1$, se obtiene la partición deseada, es decir, $F_n(x) - F_n(-x) \geq G(x) - G(-x)$ para $x \in (a_{2i+1}, a_{2i+2})$ y $F_n(x) - F_n(-x) < G(x) - G(-x)$ para $x \in (a_{2i}, a_{2i+1})$, $i = 0, 1, \dots, k-1$.

El conjunto $E_n = \{a_0, \dots, a_{2k}\}$ define una partición finita del intervalo $[0, +\infty)$ y está definido en función de $\omega \in \Omega$ y n , es decir, los valores a_i son de la forma $a_i(\omega, n)$, pero omitiremos la dependencia de a_i en ω y n (salvo cuando sea necesario) para facilitar la notación. El último elemento de E_n , es $a_{2k} = \infty$ y además, para todo $x \in (a_{2k-1}, \infty)$ se satisface la relación $F_n(x) - F_n(-x) = 1 \geq G(x) - G(-x)$.

Dado que F_n es una función escalonada, si $Z_i = \inf \{z : G(z) - G(-z) \geq i/n\}$, $i = 1, 2, \dots, n-1$, entonces, $E_n \subset \mathcal{T} \cup \{Z_i : i = 1, 2, \dots, n-1\} \cup \{\infty\}$.

Un algoritmo para encontrar los valores a_i es el siguiente. Sea $T_{(i)}$ el i -ésimo estadístico de orden del conjunto \mathcal{T} . Entonces:

- Para $i = 1, 2, \dots, n-1$, si $T_{(i)} \leq Z_i \leq T_{(i+1)}$, entonces $Z_i \in E_n$,
- Si $T_{(1)} \leq Z_1$ entonces $T_{(1)} \in E_n$,
- Para $i = 2, \dots, n-1$, si $Z_{i-1} \leq T_{(i)} \leq Z_i$ entonces $T_{(i)} \in E_n$,
- Por último, si $Z_{n-1} \in E_n$, entonces $T_{(n)} \in E_n$.

Una forma natural de corregir la función de distribución empírica F_n de tal forma que sea

válida la condición (1.1) es la siguiente:

$$\tilde{F}_n(x) = \begin{cases} \sum_{j=0}^{\infty} G(x)I_{[a_{2j}, a_{2j+1})}(x) + \sum_{j=0}^{\infty} F_n(x)I_{[a_{2j+1}, a_{2j+2})}(x) & x \geq 0 \\ \sum_{j=0}^{\infty} G(x)I_{[-a_{2j+1}, -a_{2j})}(x) + \sum_{j=0}^{\infty} F_n(x)I_{[-a_{2j+2}, -a_{2j+1})}(x) & x < 0 \end{cases}$$

Esta idea fue utilizada en El Barmi y Rojo (2003) para la estimación de funciones de distribución ordenadas de acuerdo a una dominancia estocástica de segundo orden, con aplicaciones en Finanzas.

Sin embargo, la función \tilde{F}_n no tiene un comportamiento asintótico apropiado para nuestros propósitos, ya que por ejemplo, $\tilde{F}_n(0) = G(0)$, sin importar el valor de n , lo cual influye en el comportamiento asintótico de \tilde{F}_n alrededor del 0.

Otra forma de modificar la función de distribución empírica para obtener un estimador de F , conservando las restricciones de picudez, es la siguiente: sea $\varepsilon_n \geq 0$, una sucesión tal que $\sum \exp(-n\varepsilon_n^2/2) < \infty$, $\varepsilon_n \rightarrow 0$. Definamos

$$\begin{aligned} b_0 &= 0 \\ b_1 &= \inf \{t > 0 : \min(F_n(t) + \frac{\varepsilon_n}{2}, 1) - \max(F_n(t) - \frac{\varepsilon_n}{2}, 0) < G(t) - G(-t)\} \\ b_2 &= \inf \{t > b_1 : F_n(t) - F_n(-t) \geq G(t) - G(-t)\} \\ b_{2i-1} &= \inf \{t > b_{2i-2} : F_n(t) - F_n(-t) < G(t) - G(-t)\} \\ b_{2i} &= \inf \{t > b_{2i-1} : F_n(t) - F_n(-t) \geq G(t) - G(-t)\} \end{aligned} \tag{2.2}$$

para $i = 2, \dots, k < \infty$.

Ahora definamos la función F'_n como sigue:

$$F'_n(x) = \begin{cases} \min \left\{ F_n(x) + \frac{\varepsilon_n}{2}, 1 \right\} I_{[0, b_1)}(x) + \sum_{j=1}^{\infty} G(x) I_{[b_{2j}, b_{2j+1})}(x) \\ + \sum_{j=0}^{\infty} F_n(x) I_{[b_{2j+1}, b_{2j+2})}(x) & x \geq 0 \\ \max \left\{ F_n(x) - \frac{\varepsilon_n}{2}, 0 \right\} I_{[-b_1, 0)}(x) + \sum_{j=1}^{\infty} G(x) I_{[-b_{2j+1}, -b_{2j})}(x) \\ + \sum_{j=0}^{\infty} F_n(x) I_{[-b_{2j+2}, -b_{2j+1})}(x) & x < 0 \end{cases} \quad (2.3)$$

Es fácil demostrar que la función F'_n es nonegativa, con $\lim_{x \rightarrow -\infty} F'_n(x) = 0$, $\lim_{x \rightarrow \infty} F'_n(x) = 1$ y que es continua por la derecha cuando G es continua. Pero no es una función de distribución dado que no necesariamente es no decreciente. El siguiente Lema nos garantiza que $F'_n \geq_p G$.

Lema 14 Sean F y G funciones de distribución con $F \geq_p G$ ($a = b = 0$), y F_n la función de distribución empírica correspondiente a una muestra de tamaño n de F . Entonces F'_n definida por (2.2) y (2.3) satisface la relación (1.1) con $a = b = 0$.

Demostración. Si $x \in (b_{2j}, b_{2j+1})$, $j \geq 1$, entonces $F'_n(x) - F'_n(-x) = G(x) - G(-x)$. Si $x = b_{2j}$, $j \geq 1$, entonces para toda $h > 0$ suficientemente pequeña, se tiene $F_n(x-h) - F_n(-x+h) \geq G(x-h) - G(-x+h)$. Si $h \downarrow 0$, se obtiene $F'_n(x^-) - F'_n(-x) \geq G(x^-) - G(-x)$.

Si $x \in [0, b_1)$, entonces $F'_n(x^-) - F'_n(-x) = \min \{ F_n(b_1^-) + \varepsilon_n/2, 1 \} - \max \{ F_n(-b_1) - \varepsilon_n/2, 0 \} \geq G(x^-) - G(-x)$, por definición de b_1 .

Ahora sea $x \in (b_{2j+1}, b_{2j+2})$, entonces es inmediato que $F'_n(x^-) - F'_n(-x) \geq G(x^-) - G(-x)$. Por último, para $x = b_{2j+1}$, se tiene la igualdad $F'_n(x^-) - F'_n(-x) = G(x^-) - G(-x)$. ■

Dado que la función (2.3) no es monótona, una forma inmediata de “monotonizarla” es mediante la idea dada en Rojo y Ma (1996):

$$F_n^*(x) = \begin{cases} \sup_{0 \leq y \leq x} F'_n(y) & x \geq 0 \\ \inf_{x \leq y \leq 0} F'_n(y) & x \leq 0 \end{cases} \quad (2.4)$$

De esta forma obtenemos una función monótona no decreciente que conserva las propiedades de F'_n , por lo que:

Lema 15 F_n^* es una función de distribución tal que $F_n^* \geq_p G$.

Demostración. Sea $x > 0$, entonces, por definición, $F'_n(-x) \geq F_n^*(-x)$ y $F_n^*(x^-) = \lim_{h \downarrow 0} F_n^*(x-h) \geq \lim_{h \downarrow 0} F'_n(x-h) = F'_n(x^-)$. Combinando las dos desigualdades se llega al resultado deseado. ■

2.1.1 Propiedades asintóticas del estimador F_n^*

Ahora investigaremos el comportamiento asintótico del estimador (2.4). Para esto, definamos el conjunto $A = \{x > 0 : F(x) - F(-x) = G(x) - G(-x)\}$. De ahora en adelante supondremos $A = \phi$, es decir, $F(x) - F(-x) > G(x) - G(-x)$ para todo $x > 0$. Primero un Lema que muestra el comportamiento asintótico de b_1 .

Lema 16 Sean $F \geq_p G$ funciones de distribución continuas, con $A = \phi$. Si $\varepsilon_n > 0$ es una sucesión tal que $\sum \exp(-n\varepsilon_n^2/2) < \infty$ y $\varepsilon_n \rightarrow 0$, entonces para todo $\delta > 0$,

$$P[b_1 < \delta \text{ i.o.}] = 0$$

Demostración. Sea $B_n = [b_1 < \delta]$, es decir,

$$\begin{aligned} B_n &= \left\{ \omega : \exists 0 < t < \delta \text{ tal que } \min\{F_n(t) + \frac{\varepsilon_n}{2}, 1\} - \max\{0, F_n(-t) - \frac{\varepsilon_n}{2}\} < G(t) - G(-t) \right\} \\ &= \left\{ \omega : \exists 0 < t < \delta \text{ tal que } \min\{F_n(t), 1 - \frac{\varepsilon_n}{2}\} - \max\{\frac{\varepsilon_n}{2}, F_n(-t)\} + \varepsilon_n < G(t) - G(-t) \right\} \\ &= \left\{ \omega : \exists 0 < t < \delta \text{ tal que } \min\{F_n(t), 1 - \frac{\varepsilon_n}{2}\} - F(t) \right. \\ &\quad \left. - \max\{\frac{\varepsilon_n}{2}, F_n(-t)\} + F(-t) + \varepsilon_n < G(t) - G(-t) - F(t) + F(-t) \right\} \end{aligned}$$

dado que $\varepsilon_n \downarrow 0$, entonces para un $\varepsilon > 0$ suficientemente pequeño tal que $1 - \frac{\varepsilon}{2} > F(\delta)$, existe un $N > 0$ tal que si $n \geq N$, entonces $\varepsilon_n \leq \varepsilon$ y por tanto $1 - \frac{\varepsilon_n}{2} - F(t) > 1 - \frac{\varepsilon}{2} - F(\delta)$. Así, para n suficientemente grande, $\min\{F_n(t), 1 - \frac{\varepsilon_n}{2}\} - F(t) = F_n(t) - F(t)$. De manera similar, se

prueba que $\max\{\frac{\varepsilon_n}{2}, F_n(-t)\} + F(-t) = F_n(-t) - F(-t)$ para n grande. Por tanto,

$$\begin{aligned} B_n &\subset \{\omega : \forall 0 < t < \delta, F_n(t) - F(t) - F_n(-t) + F(-t) + \varepsilon_n < G(t) - G(-t) - F(t) + F(-t)\} \\ &\subset \{\omega : \forall 0 < t < \delta, F_n(t) - F(t) - F_n(-t) + F(-t) < -\varepsilon_n\} \\ &\subset \{\omega : \forall 0 < t < \delta, |F_n(t) - F(t) - F_n(-t) + F(-t)| > \varepsilon_n\} \end{aligned}$$

y por tanto, para n suficientemente grande, $P(B_n) \leq \exp(-n\varepsilon_n^2/2)$, debido a la desigualdad dada en Dvoretzky, *et. al.*(1956). Bajo la condición $\sum \exp(-n\varepsilon_n^2/2) < \infty$, podemos afirmar que $\sum P(B_n) < \infty$ y por tanto, $P[\limsup B_n] = P[b_1 < \delta \text{ i.o}] = 0$. ■

Lema 17 Sean $F \geq_p G$ ($a = b = 0$) funciones de distribución continuas. Sea F'_n definida por (2.3), con $\varepsilon_n > 0$, una sucesión de números tales que $\sum \exp(-n\varepsilon_n^2/2) < \infty$ y $\varepsilon_n \rightarrow 0$. Entonces

$$\sup_x |F'_n(x) - F(x)| \rightarrow 0$$

con probabilidad 1.

Demostración. Sea $M > 0$ suficientemente grande, tal que $|F(x) - G(x)| \leq \varepsilon/2$ para todo $x \geq M$, con $\varepsilon > 0$. Del Lema anterior sabemos que para n grande, con probabilidad 1, $b_1 > M$. Así,

$$\begin{aligned} \sup_{0 \leq x < \infty} |F'_n(x) - F(x)| &= \max \left\{ \sup_{0 \leq x < b_1} |F'_n(x) - F(x)|, \sup_{M < x < \infty} |F'_n(x) - F(x)| \right\} \\ &\leq \max \left\{ \sup_{0 \leq x < b_1} |F'_n(x) - F(x)|, \frac{\varepsilon}{2} \right\}. \end{aligned}$$

Dado que $\sup_{0 \leq x < \delta} |F'_n(x) - F(x)| = \sup_{0 \leq x < \delta} |F_n(x) + \frac{\varepsilon_n}{2} - F(x)| \leq \frac{\varepsilon}{3}$, entonces concluimos que $\sup_{0 \leq x < \infty} |F'_n(x) - F(x)| \rightarrow 0$, cuando $n \rightarrow \infty$, con probabilidad 1. Análogamente se prueba que $\sup_{-\infty \leq x < 0} |F'_n(x) - F(x)| \rightarrow 0$ cuando $n \rightarrow \infty$ con probabilidad 1. ■

Corolario 18 Bajo las condiciones del Lema anterior,

$$\sup_{-\infty < x < \infty} |F_n^*(x) - F(x)| \rightarrow 0$$

Demostración. Se sigue directamente del teorema anterior y la siguiente desigualdad dada en Rojo (1999),

$$\sup_{-\infty < x < \infty} |F_n^*(x) - F(x)| \leq \sup_{-\infty < x < \infty} |F_n'(x) - F(x)|.$$

■

Una vez obtenida la convergencia fuerte uniforme, nuestro interés se enfoca en determinar la rapidez con que $F_n^* \rightarrow F$. Con este propósito ahora consideremos la convergencia débil del estimador.

Teorema 19 Sean F y G funciones de distribución continuas que satisfacen (1.1) con $a = b = 0$. Si $A = \phi$, entonces

$$\sqrt{n} (F_n^*(x) - F(x)) \xrightarrow{D} N(0, F(x)(1 - F(x)))$$

para todo x .

Demostración. Para $x = 0$, dado que $\varepsilon_n \rightarrow 0$, entonces $\sqrt{n} (F_n^*(0) - F(0)) \rightarrow N(0, F(0)(1 - F(0)))$.

Ahora sea $x > 0$. Por el Lema anterior, con probabilidad 1 para n grande, $b_1 > x$. Entonces $F_n^*(x) = F_n(x) + \frac{\varepsilon_n}{2}$, y la convergencia en distribución se sigue de forma inmediata. Análogamente se prueba para $x < 0$. ■

A continuación se presenta un método para estimar F , con $F \geq_p G$ y G desconocida. Llamaremos a este caso, el de dos muestras.

2.2 El estimador: Dos muestras

Ahora supongamos que la distribución G no es conocida y que tenemos las muestras X_1, X_2, \dots, X_n de F y Y_1, Y_2, \dots, Y_m de G , con $F \geq_p G$ ($a = b = 0$). De manera similar al caso una muestra, para F_n y G_m las distribuciones empíricas correspondientes, definamos:

$$\begin{aligned} b_0 &= 0 \\ b_{2i-1} &= \inf\{y > b_{2i-2} : F_n(y) - F_n(-y) < G_m(y) - G_m(-y)\} \\ b_{2i} &= \inf\{y > b_{2i-1} : F_n(y) - F_n(-y) \geq G_m(y) - G_m(-y)\} \end{aligned} \quad (2.5)$$

Si $E_{nm} = \{b_0, b_1, \dots, b_{2k+1}\}$, $\mathcal{T} = \{|X_i| : i = 1, 2, \dots, n\}$ y $\mathcal{S} = \{|Y_i| : i = 1, 2, \dots, m\}$, entonces $E_{nm} \subset \mathcal{T} \cup \mathcal{S} \cup \{\infty\}$. Un algoritmo para determinar los valores b_i es el siguiente: sean $T_{(i)}$, $S_{(i)}$ los estadísticos de orden de los conjuntos T y S respectivamente, y $p = \min\{n, m\}$,

- *) Si $S_{(1)} \leq T_{(1)}$, entonces $S_{(1)} \in E_{nm}$,
- *) Si $S_{(1)} \leq T_{(1)} \leq S_{(2)}$, entonces $T_{(1)} \in E_{nm}$,
- *) Para $i = 1, 2, \dots, p-1$, si $S_{(i)} \leq T_{(i)} \leq S_{(i+1)}$, entonces $T_{(i)} \in E_{nm}$, y si $T_{(i-1)} \leq S_{(i)} \leq T_{(i)}$, entonces $S_{(i)} \in E_{nm}$,
- *) Si $p = m$ y $S_{(m)} \leq T_{(n)}$, entonces $S_{(m)} \in E_{nm}$, $T_{(n)} \in E_{nm}$ y $T_{(i)} \notin E_{nm}$ para $i = m, \dots, n-1$,
- *) Si $p = n$, entonces $T_{(n)} \in E_{nm}$ y $S_{(i)} \notin E_{nm}$ para $i = n, \dots, m$.

Sea F'_{nm} la función definida por:

$$F'_{nm}(x) = \begin{cases} \sum_{j=0}^{\infty} F_n(x)I[b_{2j}, b_{2j+1}) + \sum_{j=0}^{\infty} G_m(x)I[b_{2j+1}, b_{2j+2}) & x \geq 0 \\ \sum_{j=0}^{\infty} F_n(x)I[-b_{2j+1}, -b_{2j}) + \sum_{j=0}^{\infty} G_m(x)I[-b_{2j+2}, -b_{2j+1}) & x < 0 \end{cases} \quad (2.6)$$

entonces, como en el caso de una muestra, esta función es continua por la derecha, no es no-decreciente necesariamente, y es igual a F_n cuando $F_n \geq_p G_m$. También de forma directa se puede ver que $\lim_{x \rightarrow \infty} F'_{nm}(x) = 1$ cuando $x \rightarrow \infty$, y $\lim_{x \rightarrow -\infty} F'_{nm}(x) = 0$ cuando $x \rightarrow -\infty$, por lo que sólo le falta ser monótona no decreciente para ser función de distribución. Además, por construcción, satisface las restricciones de picudez, según el siguiente Lema:

Lema 20 Sean $F \geq_p G$ funciones de distribución continuas, entonces la función F'_{nm} definida por (2.6) satisface la relación $F'_{nm} \geq_p G_m$.

Demostración. Sea $x > 0$. Si $x \in [b_{2j}, b_{2j+1})$, entonces $F'_{nm}(x) - F'_{nm}(-x) \geq G_m(x) - G_m(-x)$. Si $x \in [b_{2j+1}, b_{2j+2})$, entonces $F'_{nm}(x) - F'_{nm}(-x) = G_m(x) - G_m(-x)$, y por tanto se tiene el resultado deseado. ■

Como en el caso de una muestra, procedemos a monotonicar la función F'_{nm} obteniendo un estimador de F para el caso de dos muestras:

$$F_{nm}^*(x) = \begin{cases} \inf_{x \leq y \leq 0} F'_{nm}(y) & x < 0 \\ \sup_{0 \leq y \leq x} F'_{nm}(y) & x \geq 0 \end{cases} \quad (2.7)$$

y es fácil demostrar que:

Lema 21 *La función de distribución definida por (2.7) satisface la relación $F_{nm}^* \geq_p G_m$.*

Demostración. La prueba es similar al caso de una muestra. ■

2.2.1 Propiedades asintóticas del estimador $F_{nm}^*(x)$

Teorema 22 *Si $A = \emptyset$, y $F \geq_p G$ son funciones de distribución continuas, entonces*

$$\sup_{-\infty < x < \infty} |F_{nm}^*(x) - F(x)| \rightarrow 0 \text{ cuando } m, n \rightarrow \infty$$

con probabilidad 1.

Demostración. Primero se probará para F'_{nm} . Para $x = 0$, el resultado es inmediato, dado que $F'_{nm}(0) = F'_n(0)$. Ahora sea $x > 0$, por la consistencia uniforme de la función de distribución empírica, eventualmente y con probabilidad 1, $F'_{nm}(x) = F'_n(x)$, y por tanto se tiene el resultado. De forma análoga se prueba para $x < 0$. Ahora aplicando el Lema dado en Chung (1974) página 133, se obtiene $\sup_{-\infty < x < \infty} |F'_{nm}(x) - F'(x)| \rightarrow 0$ cuando $m, n \rightarrow \infty$. Y por último, aplicando la desigualdad dada en Rojo (1999), $\sup_{-\infty < x < \infty} |F_{nm}^*(x) - F(x)| \leq \sup_{-\infty < x < \infty} |F'_{nm}(x) - F'(x)|$ y se llega al resultado. ■

2.3 Estimación bajo Censura.

Si la muestra proveniente de la función de distribución F es censurada por la derecha, G es conocida, y $F \geq_p G$ (con $a = b = 0$), sustituyendo la función de distribución empírica por el estimador Kaplan-Meier \hat{F}_n en (2.2), (2.3) y (2.4), se obtiene un estimador para F , digamos F_n^c .

El estimador Kaplan-Meier está definido mediante:

$$1 - \hat{F}_n(x) = \prod_{s_i \leq x} \left(\frac{n_i - d_i}{n_i} \right)$$

donde los s_i son tiempos de falla observados y n_i, d_i son las observaciones en riesgo justo antes de s_i , y las muertes en s_i respectivamente.

Las propiedades del estimador Kaplan-Meier han sido estudiados por varios autores, y a continuación una breve descripción de algunos de esos resultados. Antes un poco de notación. Para cualquier función de distribución H , se define $T_H = \inf\{t : H(t) = 1\}$. Sea $T^* = \min(T_F, T_L)$, donde L es la función de distribución que censura. Stute y Wang (1993) demuestran la convergencia fuerte uniforme del estimador Kaplan-Meier en $(-\infty, T^*]$, cuando F y L no tienen saltos en común. Mas aún, el resultado se sigue cumpliendo si y solo si, $F(T^*) = 0$ o $F(T^*) > 0$ y $L(T^* -) < 1$. Sobre la convergencia débil del proceso asociado con el estimador Kaplan-Meier, en Breslow y Crowley (1974) se presenta la solución, mostrando la convergencia débil a un proceso Gaussiano.

Las propiedades del estimador Kaplan-Meier mencionadas anteriormente, nos permiten afirmar que el estimador F_n^c satisface las mismas propiedades asintóticas que el estimador (2.4).

Ahora analizaremos la situación en que además de restricciones de picudez, se asume que las funciones son simétricas.

2.4 Caso simétrico

En esta sección, F y G representan a dos funciones de distribución simétricas con respecto a los puntos u y v respectivamente. Bajo la suposición de que u, v son conocidos se puede suponer $u = v = 0$. Así F y G son funciones de distribución que satisfacen la igualdad

$F(x) + F(-x^-) = 1 = G(x) + G(-x^-)$ para todo x . La relación de orden definida por (1.1), en este caso, se reduce a:

Definición 23 Para F y G funciones de distribución simétricas con respecto al 0, entonces $F \geq_p G$ si y solo si

$$\begin{aligned} F(x^-) &\geq G(x^-) && \text{para } x \geq 0, \\ F(x) &\leq G(x) && \text{para } x < 0. \end{aligned} \tag{2.8}$$

La estimación de una función de distribución F simétrica que satisface (2.8) (para alguna distribución G simétrica) mediante un estimador que sea simétrico y que también satisfaga (2.8) (para la misma distribución G) se puede resolver al menos de dos maneras, utilizando los siguientes operadores:

Para f una función de distribución arbitraria. Siguiendo el método dado en Schuster (1987), el operador Φ_1 definido por

$$\Phi_1(f(x)) = \frac{1}{2} (f(x) + 1 - f(-x^-))$$

transforma una función de distribución f , en una función de distribución simétrica alrededor del 0.

Utilizando la teoría dada en Rojo y Ma (1996), definimos el operador Φ_2 mediante:

$$\Phi_2(f(x), G) = \begin{cases} \max\{f(x), G(x)\} & x \geq 0 \\ \min\{f(x), G(x)\} & x < 0 \end{cases}$$

el cual transforma la función de distribución f en una función de distribución que satisface la relación (2.8), es decir $\Phi_2(f(x), G) \geq_p G$.

Aplicar cualquiera de estos operadores a F_n es equivalente a proyectar F_n en el conjunto de funciones de distribución simétricas (bajo Φ_1) o sobre el conjunto de funciones de distribución más picudas que G (bajo Φ_2), y no son conmutativos, es decir, aplicar Φ_1 y después Φ_2 o viceversa, no lleva al mismo resultado. En la siguiente subsección, se aplicará esta técnica de proyección para obtener dos nuevos estimadores.

2.4.1 Definición del estimador

Sean $F \geq_p G$ funciones de distribución simétricas. Sin pérdida de generalidad podemos suponer los puntos de simetría en cero (o equivalentemente conocidos), pues en caso de no serlos, se podrían estimar según Schuster (1973a) y las propiedades de $\Phi_1(F_n)$ se conservan. Entonces sea X_1, X_2, \dots, X_n una muestra proveniente de la distribución F , y siendo F_n la función de distribución empírica correspondiente, aplicando las dos transformaciones anteriores a F_n , se obtienen las funciones de distribución

$$F_{n,1} = \Phi_1(\Phi_2(F_n, G))$$

y

$$F_{n,2} = \Phi_2(\Phi_1(F_n), G),$$

como estimadores de F que poseen las siguientes propiedades:

Lema 24 Sean $F \geq_p G$ funciones de distribución simétricas con respecto al 0 y F_n la función de distribución empírica asociada a F . Entonces los estimadores para F , dados por $F_{n,1}$ y $F_{n,2}$ son simétricos alrededor del 0 y satisfacen (2.8).

Demostración. Para $F_{n,1}$, por definición del operador Φ_1 , ya se tiene la simetría, ahora solo falta probar que satisface (2.8). Sea $x \geq 0$, entonces

$$\begin{aligned} F_{n,1}(x) &= \frac{1}{2} [\Phi_2(F_n(x), G) + 1 - \Phi_2(F_n(-x^-), G)] \\ &\geq \frac{1}{2} [G(x) + 1 - G(-x^-)] \\ &\geq G(x) \end{aligned}$$

donde la última desigualdad se sigue de la condición de que G es simétrica alrededor del 0. Análogamente se prueba que para $x < 0$, $F_{n,1}(x) < G(x)$. Y por tanto, $F_{n,1}$ satisface (2.8).

Consideremos ahora $F_{n,2}$. Sea $x \geq 0$. Entonces,

$$\begin{aligned}
F_{n,2}(x) &= \max \{ \Phi_1(F_n(x)), G(x) \} \\
&= \max \{ 1 - \Phi_1(F_n(-x^-)), 1 - G(-x^-) \} \\
&= 1 - \min \{ \Phi_1(F_n(-x^-)), G(-x^-) \} \\
&= 1 - F_{n,2}(-x^-).
\end{aligned}$$

Análogamente se prueba que $F_{n,2}(x) + F_{n,2}(-x^-) = 1$, para $x < 0$, y por tanto $F_{n,2}$ es simétrico alrededor del 0. Es inmediato de la definición del operador Φ_2 que la función $F_{n,2}$ satisface la condición (2.8). ■

Denotemos con \mathcal{F}_G al conjunto de funciones de distribución que satisfacen (2.8) pero que no son necesariamente simétricas. Entonces las funciones de distribución F , $\Phi_2(F_n, G)$, $F_{n,1}$ y $F_{n,2}$ son elementos de \mathcal{F}_G . Resulta que bajo la norma supremo $\|\cdot\|_\infty$, $\Phi_2(F_n, G)$ es la proyección de F_n sobre el conjunto \mathcal{F}_G , como se establece en el siguiente Lema:

Lema 25 Sean $F \geq_p G$ funciones de distribución simétricas con respecto al 0, y F_n la función de distribución empírica asociada a F . Entonces $\|\Phi_2(F_n, G) - F_n\|_\infty = \inf_{H \in \mathcal{F}_G} \|H - F_n\|_\infty$.

Demostración. Si $\|\Phi_2(F_n, G) - F_n\|_\infty = 0$, entonces $F_n \in \mathcal{F}_G$ y el resultado es inmediato. Consideremos ahora el caso cuando $\|\Phi_2(F_n, G) - F_n\|_\infty > 0$. Procediendo por contradicción, supongamos que $\|\Phi_2(F_n, G) - F_n\|_\infty > \inf_{H \in \mathcal{F}_G} \|H - F_n\|_\infty$. Entonces existe una función de distribución $H \in \mathcal{F}_G$ tal que $\|\Phi_2(F_n, G) - F_n\|_\infty > \|H - F_n\|_\infty$. Sea x_i tal que $\|\Phi_2(F_n, G) - F_n\|_\infty = |\Phi_2(F_n(x_i), G) - F_n(x_i)| > 0$.

Si $x_i \geq 0$, entonces $|\Phi_2(F_n(x_i), G) - F_n(x_i)| = \Phi_2(F_n(x_i), G) - F_n(x_i) > |H(x_i) - F_n(x_i)|$. Ahora, $G(x_i) - F_n(x_i) \leq H(x_i) - F_n(x_i) < \Phi_2(F_n(x_i), G) - F_n(x_i)$. De la desigualdad anterior, se tiene $G(x_i) < \Phi_2(F_n(x_i), G)$, es decir, $\Phi_2(F_n(x_i)) = F_n(x_i)$ lo cual es una contradicción, ya que $|\Phi_2(F_n(x_i), G) - F_n(x_i)| > 0$.

Si $x_i < 0$, entonces $F_n(x_i) - G(x_i) \leq F_n(x_i) - H(x_i) \leq |F_n(x_i) - H(x_i)| < |\Phi_2(F_n(x_i), G) - F_n(x_i)| = F_n(x_i) - \Phi_2(F_n(x_i), G)$ y por tanto $G(x_i) > \Phi_2(F_n(x_i), G)$. Esto último implica que $F_n(x_i) - \Phi_2(F_n(x_i), G) = 0$, lo cual es una contradicción. ■

Un resultado aun más fuerte que el anterior es el siguiente:

Teorema 26 Sean $F \geq_p G$ funciones de distribución simétricas con respecto al 0, y F_n la función de distribución empírica asociada a F . Entonces

$$|\Phi_2(F_n(x), G) - F_n(x)| \leq |H(x) - F_n(x)|$$

para todo x y todo $H \in \mathcal{F}_G$.

Demostración. Sea $H \in \mathcal{F}_G$. Para $x > 0$, si $F_n(x) \geq G(x)$, entonces $|\Phi_2(F_n(x), G) - F_n(x)| = 0$; pero si $F_n(x) < G(x)$, entonces $|\Phi_2(F_n(x), G) - F_n(x)| = G(x) - F_n(x) \leq H(x) - F_n(x)$, pues $G(x) \leq H(x)$.

Para $x < 0$, $|\Phi_2(F_n(x), G) - F_n(x)| = 0$ si $F_n(x) \leq G(x)$ y $|\Phi_2(F_n(x), G) - F_n(x)| = F_n(x) - G(x) \leq F_n(x) - H(x) \leq |H(x) - F_n(x)|$ para $F_n(x) > G(x)$. ■

De lo anterior, podemos afirmar que

$$|\Phi_2(F_n(x), G) - F_n(x)| \leq |F(x) - F_n(x)| \quad \text{para todo } x \quad (2.9)$$

Un resultado similar a (2.9) que nos proporcionará las herramientas para demostrar la convergencia uniforme de nuestros estimadores es como sigue:

Lema 27 Sea f una función real no negativa. Entonces, $|\Phi_2(f(x), G) - F(x)| \leq |F(x) - f(x)|$ para todo x

Demostración. Para $x \geq 0$, si $f(x) < G(x) \leq F(x)$, entonces $|\Phi_2(f(x), G) - F(x)| = F(x) - G(x) \leq F(x) - f(x)$. Si $f(x) \geq G(x)$, entonces $|\Phi_2(f(x), G) - F(x)| = |f(x) - F(x)|$.

Para $x < 0$, si $f(x) \leq G(x)$, entonces $|\Phi_2(f(x), G) - F(x)| = |f(x) - F(x)|$. Y si $f(x) > G(x) \geq F(x)$, entonces $|\Phi_2(f(x), G) - F(x)| = G(x) - F(x) \leq f(x) - F(x) = |f(x) - F(x)|$. ■

También es posible establecer resultados similares para $\Phi_1(F_n)$. Sea \mathcal{S} el conjunto de las funciones de distribución simétricas alrededor del cero. Entonces $\Phi_1(F_n)$ es la proyección de F_n sobre \mathcal{S} , con respecto a la norma supremo:

Teorema 28 Sea F una función de distribución simétrica con respecto al 0, entonces

$$\|\Phi_1(F_n) - F_n\|_\infty = \inf_{H \in \mathcal{S}} \|H - F_n\|_\infty$$

Demostración. Ver prueba en Schuster (1987), Teorema 1. ■

De acuerdo a los dos últimos resultados, existen posibilidades de que alguno de los dos estimadores también sea una especie de proyección de F_n sobre el espacio $\mathcal{F}_G \cap \mathcal{S}$, de las funciones simétricas que satisfacen (2.8). Un resultado interesante que surgió de manera sorpresiva es que $F_{n,1}(x) \geq_p F_{n,2}(x)$, como se probará en el siguiente Lema.

Lema 29 $F_{n,1}(x) \leq F_{n,2}(x)$ para $x \geq 0$, y $F_{n,1}(x) \geq F_{n,2}(x)$ para $x < 0$.

Demostración. Sea $x \geq 0$. Dado que $F_n(x) \leq \Phi_2(F_n(x), G)$ y $\Phi_2(F_n(-x^-), G) \leq F_n(-x^-)$, entonces $F_n(x) + 1 - F_n(-x^-) \leq \Phi_2(F_n(x), G) + 1 - \Phi_2(F_n(-x^-), G)$, es decir, $\Phi_1(F_n(x)) \leq F_{n,1}(x)$. De la última desigualdad, concluimos que $F_{n,2}(x) \leq F_{n,1}(x)$.

Ahora consideremos $x < 0$. En este caso, se satisfacen las desigualdades $F_n(x) \geq \Phi_2(F_n(x), G)$ y $\Phi_2(F_n(-x^-), G) \geq F_n(-x^-)$. Esto implica que $\Phi_2(F_n(x), G) + 1 - \Phi_2(F_n(-x^-), G) \leq F_n(x) + 1 - F_n(-x^-)$, es decir, $F_{n,1}(x) \leq \Phi_1(F_n(x))$. De la última desigualdad, se concluye el resultado deseado. ■

El siguiente corolario, expresa las condiciones para igualdad entre los dos estimadores y como se puede ver, los estimadores no son iguales.

Corolario 30 a) Para $x \geq 0$, si $G(x) \leq \min \{F_n(x), 1 - F_n(-x^-)\}$ o

$$G(x) \geq \max \{F_n(x), 1 - F_n(-x^-)\}, \text{ entonces } F_{n,1}(x) = F_{n,2}(x),$$

b) Para $x < 0$, si $G(x) \leq \max \{F_n(x), 1 - F_n(-x^-)\}$ o $G(x) \geq \min \{F_n(x), 1 - F_n(-x^-)\}$, entonces $F_{n,1}(x) = F_{n,2}(x)$.

Demostración.

a) Sea $x \geq 0$. Si $G(x) \leq \min \{F_n(x), 1 - F_n(-x^-)\}$, entonces $G(x) \leq \Phi_1(F_n(x))$, y por tanto, $\Phi_1(F_n(x)) \leq F_{n,1}(x) \leq F_{n,2}(x) = \Phi_1(F_n(x))$, que es la igualdad pedida.

Ahora supongamos que $G(x) \geq \max\{F_n(x), 1 - F_n(-x^-)\}$. Entonces $G(x) \geq \Phi_1(F_n(x))$ y $\Phi_2(F_n(x), G) = G(x)$, y por tanto, $F_{n,1}(x) = \Phi_1(G(x)) = G(x) = F_{n,2}(x)$.

b) Sea $x < 0$. Si $G(x) \leq \max\{F_n(x), 1 - F_n(-x^-)\}$, entonces $G(x) \leq \Phi_1(F_n(x))$, y $\Phi_2(F_n(x), G) = G(x)$, y por lo tanto, $F_{n,1}(x) = G(x) = F_{n,2}(x)$. Consideremos ahora el caso $G(x) \geq \min\{F_n(x), 1 - F_n(-x^-)\}$. Entonces $\Phi_2(F_n(x), G) = F_n(x)$, y $G(x) \geq \Phi_1(F_n(x))$. De lo anterior podemos concluir que $F_{n,1}(x) = \Phi_1(F_n(x)) = F_{n,2}(x)$.

■

En lo que resta del capítulo, se utilizara la notación $[x]$ para el entero más grande menor o igual a x , y $[x]^*$ para el entero mas chico mayor o igual que x .

Algunas propiedades que se espera posean los estimadores es la insesgadez, asi como la consistencia y caracterización de su distribución asintótica. Desafortunadamente los estimadores $\Phi_2(F_n, G)$, $F_{n,1}$ y $F_{n,2}$ no son insesgados, aunque su sesgo no es difícil de calcular, como se muestra en el siguiente teorema:

Teorema 31 Sean $F \geq_p G$ funciones de distribución simétricas con respecto al 0, y F_n la función de distribución empírica asociada a F . Entonces:

1. El estimador $\Phi_1(F_n)$ es insesgado con varianza menor al de la distribución empírica:

$$V(\Phi_1(F_n(x))) = F(-|x|)[1 - 2F(-|x|)]/2n \leq V(F_n).$$

2. El sesgo del estimador $\Phi_2(F_n, G)$ esta dado por

$$\text{Sesgo}(\Phi_2(F_n(x), G)) = \begin{cases} -\frac{1}{n} \sum_{j=l}^n (1 - B(j; n, p)) + (G(x) - \frac{l}{n}) (1 - B(l; n, p)) & x < 0 \\ -\frac{1}{n} \sum_{j=0}^{l-1} (1 - B(j; n, p)) + (\frac{l}{n} - G(x)) (1 - B(l; n, p)) & x \geq 0 \end{cases}$$

con $l = [nG(x)]$.

3. El estimador $F_{n,1}$ es sesgado.

4. El estimador $F_{n,2}$ es sesgado.

Demostración.

1. Ver demostración en Schuster(1973b).
2. Tomando las ideas dadas en Rojo y Ma (1996), sean $b(k; n, p)$, $B(k; n, p)$ y $\overline{B}(k; n, p) = 1 - B(k; n, p)$, las funciones de densidad, distribución y supervivencia de una binomial con parámetros n y p . Entonces, $P[F_n(x) = t] = B([nt]; n, p)$, con $p = F(x)$.

Para $x < 0$, sea $0 \leq t \leq G(x)$. Entonces $\Phi_2(F_n(x), G) = t$ si y solo si, $t < F_n(x) \leq F_n(0)$ y $G(x) = t$; o $F_n(x) = t$ y $0 \leq t \leq G(x)$, y por tanto

$$P[\Phi_2(F_n(x), G) = t] = \sum_{k=0}^l b(k; n, p)I[nt = k] + \overline{B}([nt]; n, p)I[G(x) = t]$$

con $l = [nG(x)]$

Si $x \geq 0$, y $G(x) \leq t \leq 1$. Entonces $\Phi_2(F_n(x), G) = t$ si y solo si, $F_n(x) < t$ y $G(x) = t$; o $F_n(x) = t$ y $G(x) \leq t$. De lo anterior, obtenemos

$$P[\Phi_2(F_n(x), G) = t] = B([nt]; n, p)I[G(x) = t] + \sum_{k=[nG(x)]^*}^{[nt]} b(k; n, p)I[nt = k]$$

con $[x]$ igual al máximo entero $\leq x$, y $[x]^*$ igual al mínimo entero $\geq x$.

Entonces la función de distribución de $\Phi_2(F_n(x), G)$ está dada por:

Si $x < 0$,

$$P[\Phi_2(F_n(x), G) \leq t] = \begin{cases} B([nt]; n, p) & \text{si } 0 \leq t < G(x) \\ 1 & \text{si } t \geq G(x), \end{cases}$$

y si $x \geq 0$,

$$P[\Phi_2(F_n(x), G) \leq t] = \begin{cases} 0 & \text{si } t < G(x) \\ B([nt]; n, p) & \text{si } t \geq G(x) \end{cases}.$$

Utilizando la igualdad $E(X) = \int_0^\infty P[X > t] dt$, para $X \geq 0$, para $x < 0$, obtenemos,

$$\begin{aligned} E(\Phi_2(F_n(x), G)) &= \frac{1}{n} \sum_{j=0}^{l-1} \bar{B}(j; n, p) + \left(G(x) - \frac{l}{n}\right) \bar{B}(l; n, p) \\ &= F(x) - \frac{1}{n} \sum_{j=l}^n \bar{B}(j; n, p) + \left(G(x) - \frac{l}{n}\right) \bar{B}(l; n, p). \end{aligned}$$

análogamente, para $x \geq 0$,

$$\begin{aligned} E(\Phi_2(F_n(x), G)) &= \frac{1}{n} \sum_{j=l+1}^n \bar{B}(j; n, p) + \left(\frac{l+1}{n} - G(x)\right) \bar{B}(l; n, p) + G(x) \\ &= F(x) + G(x) - \frac{1}{n} \sum_{j=0}^l \bar{B}(j; n, p) + \left(\frac{l+1}{n} - G(x)\right) \bar{B}(l; n, p) \\ &= F(x) + G(x) - \frac{1}{n} \sum_{j=0}^{l-1} \bar{B}(j; n, p) + \left(\frac{l}{n} - G(x)\right) \bar{B}(l; n, p) \end{aligned}$$

de lo cual se concluye el resultado deseado.

3. Dado que $F_{n,1}(x) = (\Phi_2(F_n(x), G) + 1 - \Phi_2(F_n(-x^-), G))/2$, entonces, para $x < 0$,

$$\begin{aligned} E[F_{n,1}(x)] &= \frac{1}{2} \{E[\Phi_2(F_n(x), G)] + 1 - E[\Phi_2(F_n(-x^-), G)]\} \\ &= \frac{1}{2} \left\{ F(x) - \frac{1}{n} \sum_{j=l}^n \bar{B}(j; n, p) + \left(G(x) - \frac{l}{n}\right) \bar{B}(l; n, p) \right. \\ &\quad \left. + 1 - F(-x) - G(-x) + \frac{1}{n} \sum_{j=0}^{[nG(-x)]-1} \bar{B}(j; n, p) \right. \\ &\quad \left. - \bar{B}([nG(-x)]; n, p) \left(\frac{[nG(-x)]}{n} - G(-x)\right) \right\} \\ &= F(x) + \frac{1}{2n} \left\{ - \sum_{j=l}^n \bar{B}(j; n, p) + \sum_{j=0}^{[nG(-x)]-1} \bar{B}(j; n, p) \right\} \\ &\quad + \frac{1}{2} \left(G(x) - \frac{l}{n}\right) \bar{B}(l; n, p) - \frac{G(-x)}{2} \\ &\quad - \frac{1}{2} \left(\frac{[nG(-x)]}{n} - G(-x)\right) \bar{B}([nG(-x)]; n, p) \end{aligned}$$

de la simetría de G , tenemos las igualdades $[nG(x)] + [nG(-x)] = n - I[nG(x) \text{ no es entero}]$,

y $\bar{B}([nG(-x)]; n, p) = B([nG(x)]; n, p)$ y por tanto,

$$\begin{aligned} E[F_{n,1}(x)] &= F(x) + G(x) - \frac{1}{2} \left(1 + \frac{l}{n}\right) \\ &\quad - \frac{1}{2n} I[nG(x) \text{ no es entero}] B(l; n, p) \\ &\quad + \frac{1}{2n} \left\{ \sum_{j=0}^{l-1} \bar{B}(j; n, p) - \sum_{j=[nG(-x)]}^n \bar{B}(j; n, p) \right\} \end{aligned}$$

y para $x \geq 0$,

$$\begin{aligned} E[F_{n,1}(x)] &= \frac{1}{2} \left\{ F(x) - \frac{1}{n} \sum_{j=0}^{l-1} \bar{B}(j; n, p) + \left(\frac{l}{n} - G(x)\right) \bar{B}(l; n, p) \right\} \\ &\quad + \frac{1}{2} \left[G(x) + 1 - F(-x) + \frac{1}{n} \sum_{j=[nG(-x)]}^n \bar{B}(j; n, p) \right. \\ &\quad \left. - \left(G(-x) - \frac{[nG(-x)]}{n}\right) \bar{B}([nG(-x)]; n, p) \right] \\ &= F(x) + \frac{1}{2} \left\{ \frac{1}{n} \sum_{j=[nG(-x)]}^n \bar{B}(j; n, p) - \frac{1}{n} \sum_{j=0}^{l-1} \bar{B}(j; n, p) \right\} \\ &\quad + \frac{1}{2} \left[\left(\frac{l}{n} - G(x)\right) \bar{B}(l; n, p) + G(x) \right] \\ &\quad - \frac{1}{2} \left[\left(G(-x) - \frac{[nG(-x)]}{n}\right) \bar{B}([nG(-x)]; n, p) \right] \\ &= F(x) + \frac{1}{2} \left\{ \frac{1}{n} \sum_{j=[nG(-x)]}^n \bar{B}(j; n, p) - \frac{1}{n} \sum_{j=0}^{l-1} \bar{B}(j; n, p) \right\} \\ &\quad + \frac{1}{2} \left[\left(\frac{l}{n} - G(x)\right) (1 - 2B(l; n, p)) \right. \\ &\quad \left. + \left(G(x) - \frac{B(l; n, p)}{n} I[nG(x) \text{ no es entero}]\right) \right] \end{aligned}$$

4. Primero calculemos la función de distribución. Sea $x < 0$. Dado que $F_{n,2}(x) = t$ si y solo

si $\Phi_1(F_n(x)) \geq t$ y $G(x) = t$ ó $\Phi_1(F_n(x)) = t$ y $t < G(x)$, entonces,

$$\begin{aligned} P[F_{n,2}(x) = t] &= P[\Phi_1(F_n(x)) \geq t] I[G(x) = t] + P[\Phi_1(F_n(x)) = t] I[t < G(x)] \\ &= B([n(1-2t)]; n, F(-x) - F(x)) I[G(x) = t] \\ &\quad + \sum_{k=0}^{[nG(x)]} b(k; n, F(-x) - F(x)) I[nk = t] \end{aligned}$$

y por tanto,

$$P[F_{n,2}(x) \leq t] = \begin{cases} B([n(1-2t)]; n, F(-x) - F(x)) & \text{si } t < G(x) \\ 1 & \text{si } t \geq G(x) \end{cases}$$

Ahora, consideremos el caso $x \geq 0$. Como $F_{n,2}(x) = t$ si y solo si $\Phi_1(F_n(x)) \leq t$ y $G(x) = t$ ó $\Phi_1(F_n(x)) = t$ y $t > G(x)$, entonces,

$$\begin{aligned} P[F_{n,2}(x) = t] &= P[\Phi_1(F_n(x)) \leq t] I[G(x) = t] + P[\Phi_1(F_n(x)) = t] I[t > G(x)] \\ &= B([n(2t-1)]; n, F(x) - F(-x)) I[G(x) = t] \\ &\quad + \sum_{k=[nG(x)]}^n b(k; n, F(x) - F(-x)) I[nk = t] \end{aligned}$$

y la función de distribución de $F_{n,2}(x)$ es

$$P[F_{n,2}(x) \leq t] = B([n(2t-1)]; n, F(x) - F(-x)) I[G(x) \geq t].$$

Entonces, utilizando las funciones de distribución previamente calculadas, obtenemos para $x < 0$,

$$\begin{aligned}
E(F_{n,2}(x)) &= \frac{1}{2n} \sum_{j=0}^{[2nG(x)]-1} \overline{B}(n-j; n, F(-x) - F(x)) \\
&\quad + \left(G(x) - \frac{[2nG(x)]}{2n} \right) \overline{B}(n - [2nG(x)]; n, F(-x) - F(x)) \\
&= \frac{F(-x) - F(x)}{2} - \frac{1}{2n} \sum_{j=0}^{n-[2nG(x)]} \overline{B}(j; n, F(-x) - F(x)) \\
&\quad + \left(G(x) - \frac{[2nG(x)]}{2n} \right) \overline{B}(n - [2nG(x)]; n, F(-x) - F(x)) \\
&= \frac{1}{2} - F(x) - \frac{1}{2n} \sum_{j=0}^{n-[2nG(x)]} \overline{B}(j; n, F(-x) - F(x)) \\
&\quad + \left(G(x) - \frac{[2nG(x)]}{2n} \right) \overline{B}(n - [2nG(x)]; n, F(-x) - F(x)),
\end{aligned}$$

y para $x \geq 0$,

$$\begin{aligned}
E(F_{n,2}(x)) &= G(x) + \frac{1}{2n} \sum_{k=[2nG(x)]+1}^n \overline{B}(k-n; n, F(x) - F(-x)) \\
&\quad + \left(\frac{[2nG(x)]+1}{2n} - G(x) \right) \overline{B}([2nG(x)] - n; n, F(x) - F(-x)) \\
&= G(x) + \frac{1}{2n} \sum_{j=0}^{[2nG(x)]+1-n} \overline{B}(j; n, F(x) - F(-x)) \\
&\quad + \left(\frac{[2nG(x)]+1}{2n} - G(x) \right) \overline{B}([2nG(x)] - n; n, F(x) - F(-x)) \\
&= F(x) + G(x) - \frac{1}{2} - \frac{1}{2n} \sum_{j=[2nG(x)]+2-n}^n \overline{B}(j; n, F(x) - F(-x)) \\
&\quad + \left(\frac{[2nG(x)]+1}{2n} - G(x) \right) \overline{B}([2nG(x)] - n; n, F(x) - F(-x))
\end{aligned}$$

■

2.4.2 Propiedades asintóticas de los estimadores $F_{n,1}$ y $F_{n,2}$

Ahora se presentan los resultados correspondientes a las propiedades asintóticas de los estimadores, es decir, la consistencia uniforme y las distribuciones asintóticas respectivas.

Teorema 32 Sean F y G funciones de distribución simétricas con respecto al 0 y que satisfacen (2.8). Entonces, con probabilidad 1

a)

$$\|F_{n,1} - F\|_\infty \rightarrow 0 \text{ cuando } n \rightarrow \infty,$$

b)

$$\|F_{n,2} - F\|_\infty \rightarrow 0 \text{ cuando } n \rightarrow \infty.$$

Demostración.

a) Note que para todo x y f ,

$$\begin{aligned} \Phi_1(f(x)) - F(x) &= \frac{1}{2} [(f(x) + 1 - f(-x^-)) - F(x) - (1 - F(-x))] \\ &= \frac{1}{2} [(f(x) - F(x)) + (F(-x) - f(-x^-))] \end{aligned}$$

por lo que

$$|F_{n,1}(x) - F(x)| \leq \|\Phi_2(F_n, G) - F\|_\infty \leq \|F_n - F\|_\infty$$

donde la última desigualdad se sigue del Lema (27). La consistencia uniforme se sigue de inmediato de la desigualdad anterior.

b) Del Lema (27) y de la nota al principio de a), se sigue que

$$\|F_{n,2}(x) - F\|_\infty \leq \|\Phi_1(F_n) - F\|_\infty \leq \|F_n - F\|_\infty$$

y por tanto se tiene el resultado pedido.

■

Note que se ha probado lo siguiente:

$$\|F_{n,2}(x) - F\|_\infty \leq \|\Phi_1(F_n) - F\|_\infty \leq \|F_n - F\|_\infty$$

y

$$\|F_{n,1}(x) - F\|_\infty \leq \|\Phi_2(F_n, G) - F\|_\infty \leq \|F_n - F\|_\infty$$

es decir, ambos estimadores están mucho mas cercanos a la verdadera distribución que la empírica, en términos de la métrica uniforme.

Ahora se presentan las distribuciones asintóticas de estos estimadores.

Teorema 33 Sean F y G funciones de distribución continuas y simétricas, que satisfacen $F(x) > G(x)$ para $x > 0$, y $F(x) < G(x)$ para $x < 0$. Entonces para $i = 1, 2$

$$\sqrt{n}(F_{n,i}(x) - F(x)) \xrightarrow{D} N\left(0, \frac{F(-|x|)}{2} (2F(|x|) - 1)\right)$$

Demostración. Probaremos el resultado para $F_{n,1}$. Sea $x > 0$, con $F(x) > G(x)$, y $F(-x) > G(-x)$. Dado que

$$\begin{aligned} \sqrt{n}(F_{n,1}(x) - F(x)) &= \frac{\sqrt{n}}{2} \max\{F_n(x) - F(x), G(x) - F(x)\} \\ &\quad - \frac{\sqrt{n}}{2} \min\{F_n(-x^-) - F(-x), G(-x) - F(-x)\} \end{aligned}$$

y $G(x) - F(x) < 0$, $G(-x) - F(-x) > 0$, entonces, eventualmente y con probabilidad 1

$$\sqrt{n}(F_{n,1}(x) - F(x)) = \frac{\sqrt{n}}{2} (F_n(x) - F(x)) - \frac{\sqrt{n}}{2} (F_n(-x^-) - F(-x)).$$

Dado que el proceso empírico $\{\sqrt{n}(F_n(x) - F(x)) : -\infty < x < \infty\}$ converge debilmente al proceso Gaussiano con media cero y función de covarianza $F(s)(1 - F(t))$, $s \leq t$, entonces, obtenemos $\sqrt{n}(F_{n,1}(x) - F(x)) \xrightarrow{D} N\left(0, \frac{F(-|x|)}{2} (2F(|x|) - 1)\right)$.

Analogamente se prueba para el estimador $F_{n,2}$. ■

Mas aun, un resultado más fuerte es la convergencia debil del proceso respectivo, bajo la condición de $F(x) > G(x)$, $x > 0$.

Teorema 34 Sean F, G funciones de distribución continuas simétricas, que satisfacen $F(x) > G(x)$ para $x > 0$. Entonces, para $i = 1, 2$, el proceso $\{\sqrt{n}(F_{n,i}(x) - F(x)) : -\infty < x < \infty\}$ converge debilmente al proceso Gaussiano $\{\tilde{B}(x) : -\infty < x < \infty\}$ con media cero y función de covarianza

$$\text{Cov}\left(\tilde{B}(x), \tilde{B}(y)\right) = \begin{cases} \frac{1}{2}\bar{F}(x)(F(x) - F(-x)), & |y| > |x| \\ \frac{1}{2}F(x)(F(-y) - F(y)) & |y| < |x| \end{cases}$$

Demostración. En Rojo y Ma (1996) se prueba que $\max(\sqrt{n}(F_n(x) - F(x)), G(x) - F(x))$ y $\min(\sqrt{n}(F_n(x) - F(x)), G(x) - F(x))$ convergen a un proceso Gaussiano con media cero. Dado que los estimadores $F_{n,1}$ y $F_{n,2}$ son funciones continuas de estos procesos con respecto a la norma $\|\cdot\|_\infty$, entonces el resultado se sigue de forma inmediata. ■

2.4.3 El caso dos muestras.

Si la situación es tal que no se conoce G , pero se tiene una muestra de G de tal forma que tenemos un estimador para G dado por la distribución empírica, G_m , la teoría dada anteriormente es aplicable, a unque no directamente con la función de distribución empírica G_m , pues esta no es necesariamente simétrica con respecto al 0. Así primeramente simetrizaremos G_m , es decir, sea $\tilde{G}_m = \Phi_1(G_m)$ la versión simétrica de G_m , aplicando el proceso dado en la sección anterior se obtienen

$$F_{nm}^1 = \Phi_1\left(\Phi_2(F_n, \tilde{G}_m)\right) \text{ y } F_{nm}^2 = \Phi_2\left(\Phi_1(F_n), \tilde{G}_m\right)$$

que satisfacen la relación $F_{nm}^1 \geq_p \tilde{G}_m$, y $F_{nm}^2 \geq_p \tilde{G}_m$ y son simétricos alrededor del 0.

Como en el caso de una muestra, las propiedades asintóticas de estos estimadores son de gran importancia, y resultados similares de consistencia y distribución asintótica son fáciles de demostrar:

Teorema 35 Sean $F \geq_p G$ funciones de distribución simétricas con respecto al 0, X_1, \dots, X_n y Y_1, \dots, Y_m muestras de F y G respectivamente, entonces los estimadores F_{nm}^1 y F_{nm}^2 son uniforme y fuertemente consistentes.

Demostración. Probaremos el resultado para F_{nm}^1 , pues la prueba es similar para F_{nm}^2 . Note que

$$\begin{aligned} |F_{nm}^1(x) - F(x)| &= \frac{1}{2} \left| \max\{F_n(x), \tilde{G}_m(x)\} + 1 - \min\{F_n(-x^-), \tilde{G}_m(-x^-)\} \right. \\ &\quad \left. - \max\{F(x), G(x)\} - 1 + \min\{F(-x), G(-x)\} \right| \\ &\leq \frac{1}{2} \max\{|F_n(x) - F(x)|, |\tilde{G}_m(x) - G(x)|\} \\ &\quad + \frac{1}{2} \max\{|F_n(-x^-) - F(-x)|, |\tilde{G}_m(-x^-) - G(-x)|\} \end{aligned}$$

y por tanto,

$$\|F_{nm}^1 - F\|_\infty \leq \max \left\{ \|F_n - F\|_\infty, \|\tilde{G}_m - G\|_\infty \right\}$$

y el resultado se sigue de la consistencia fuerte uniforme de F_n y \tilde{G}_m . ■

La distribución asintótica de los estimadores tambien se conoce, y se presenta en el siguiente Teorema.

Teorema 36 Sean $F \geq_p G$ funciones de distribución simétricas con respecto al 0, X_1, \dots, X_n y Y_1, \dots, Y_m muestras de F y G respectivamente, con $F(x) > G(x)$ (para $x > 0$) o $F(x) < G(x)$ (para $x < 0$), entonces, para $i = 1, 2$,

$$\sqrt{n}(F_{nm}^i(x) - F(x)) \xrightarrow{D} N \left(0, \frac{F(-|x|)}{2} (2F(|x|) - 1) \right)$$

Demostración. Dado que

$$\begin{aligned} \sqrt{n}(F_{nm}^1(x) - F(x)) &= \frac{\sqrt{n}}{2} \left\{ \max \left\{ F_n(x) - F(x), \tilde{G}_m(x) - F(x) \right\} \right. \\ &\quad \left. - \min \left\{ F_n(-x^-) - F(-x), \tilde{G}_m(-x^-) - F(-x) \right\} \right\}, \end{aligned}$$

entonces para $F(x) > G(x)$ y $x > 0$, eventualmente y con probabilidad 1, cuando $n, m \rightarrow \infty$,

$$\sqrt{n}(F_{nm}^1(x) - F(x)) = \frac{\sqrt{n}}{2} (F_n(x) - F(x)) + \frac{\sqrt{n}}{2} (F_n(-x^-) - F(-x^-)),$$

y el resultado se obtiene.

De forma análoga se prueba para $x < 0$ y $F(x) < G(x)$. La prueba para F_{nm}^2 es similar y se omite. ■

2.4.4 El caso de una muestra con censura

Un último caso a considerar es la estimación de F con una muestra censurada y G conocida. En este caso se procede de forma natural utilizando el estimador Kaplan-Meier \hat{F}_n en lugar de la función de distribución empírica, obteniéndose los estimadores para F ,

$$\hat{F}_{nm}^1 = \Phi_1 \left(\Phi_2(\hat{F}_n, G) \right) \text{ y } \hat{F}_{nm}^2 = \Phi_2 \left(\Phi_1(\hat{F}_n), G \right).$$

Las propiedades asintóticas del estimador Kaplan-Meier, mencionados en Sección 2.3, permiten afirmar que estos dos nuevos estimadores satisfacen las mismas propiedades asintóticas que su versión con la función de distribución empírica.

2.5 Simulaciones

A continuación se presenta un estudio de simulación para cuando las funciones de distribución son simétricas. Dado que los estimadores propuestos son sesgados, una forma de comparar su desempeño con respecto a la función de distribución empírica es mediante el Error Cuadrático Medio (MSE). Se hicieron varias simulaciones bajo distintos escenarios, con tamaños de muestra $n = 10, 20$ y 30 . La siguiente tabla muestra las distribuciones utilizadas.

F		G	
$N(0, 1)$	$N(0, 1.1)$	$N(0, 1.5)$	$N(0, 2)$
$Logistica(0, 1)$	$Logistica(0, 1.1)$	$Logistica(0, 1.5)$	$Logistica(0, 2)$
$Cauchy(0, 1)$	$Cauchy(0, 1.1)$	$Cauchy(0, 1.5)$	$Cauchy(0, 2)$

Las gráficas muestran el cociente del MSE (eje vertical) de un estimador nuevo con la función de distribución empírica, calculados en el punto x , tal que $F(x) = 0.05, 0.1, 0.15, \dots, 0.95$, que son los valores que aparecen en el eje horizontal. La estimación de los MSE se basan en 10,000 corridas con muestras de tamaño n en cada una de ellas, es decir,

$$MSE(\text{estimador nuevo}, F(x)) = \frac{1}{10000} \sum (\text{estimador nuevo} - F(x))^2, \quad (2.10)$$

con estimador nuevo refiriéndose a alguno de los estimadores propuestos, la función de distribución empírica o el estimador Kaplan-Meier.

Como puede notarse de las gráficas, cualquiera de los dos nuevos estimadores es uniformemente con MSE menor que la función de distribución empírica.

También se hicieron corridas para comparar los respectivos MSE, en el caso de que se tuviera una muestra censurada de la distribución F . La función de distribución que censura en todos los casos es la exponencial estándar, es decir con media 1. De la figura (2-4) se aprecia que los

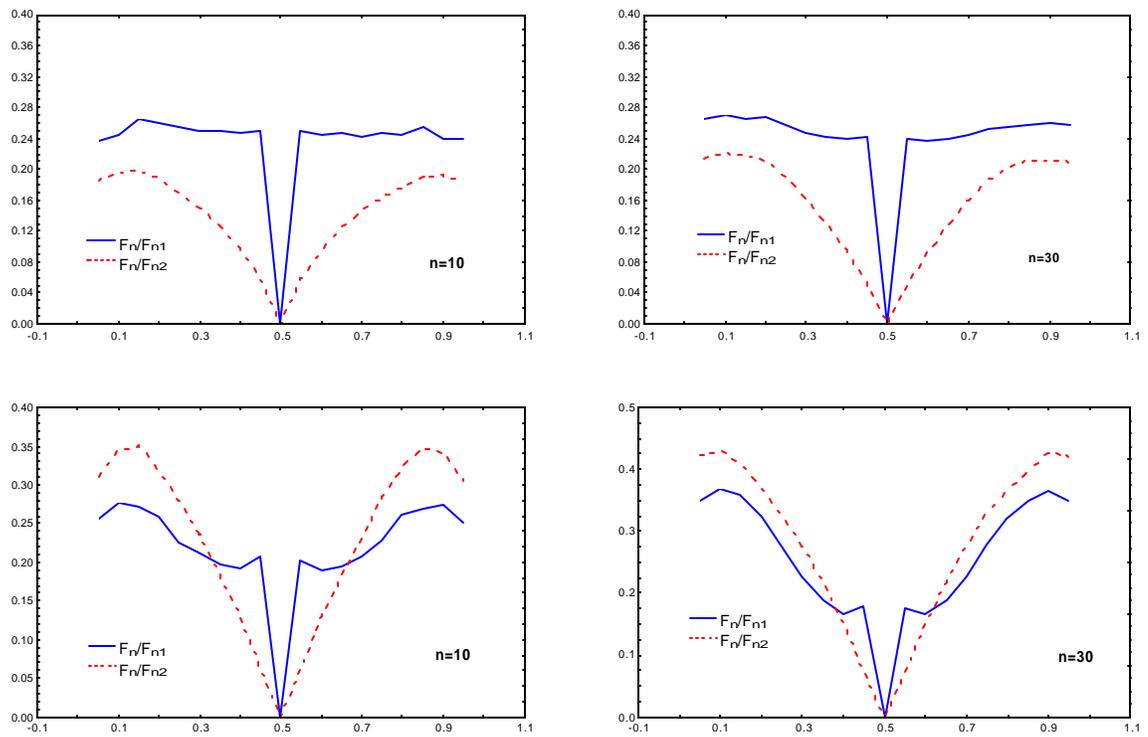


Figura 2-1: Caso una muestra. Distribución Cauchy. Las gráficas en la fila superior corresponden a $\sigma = 1.1$, y la fila inferior a $\sigma = 2$.

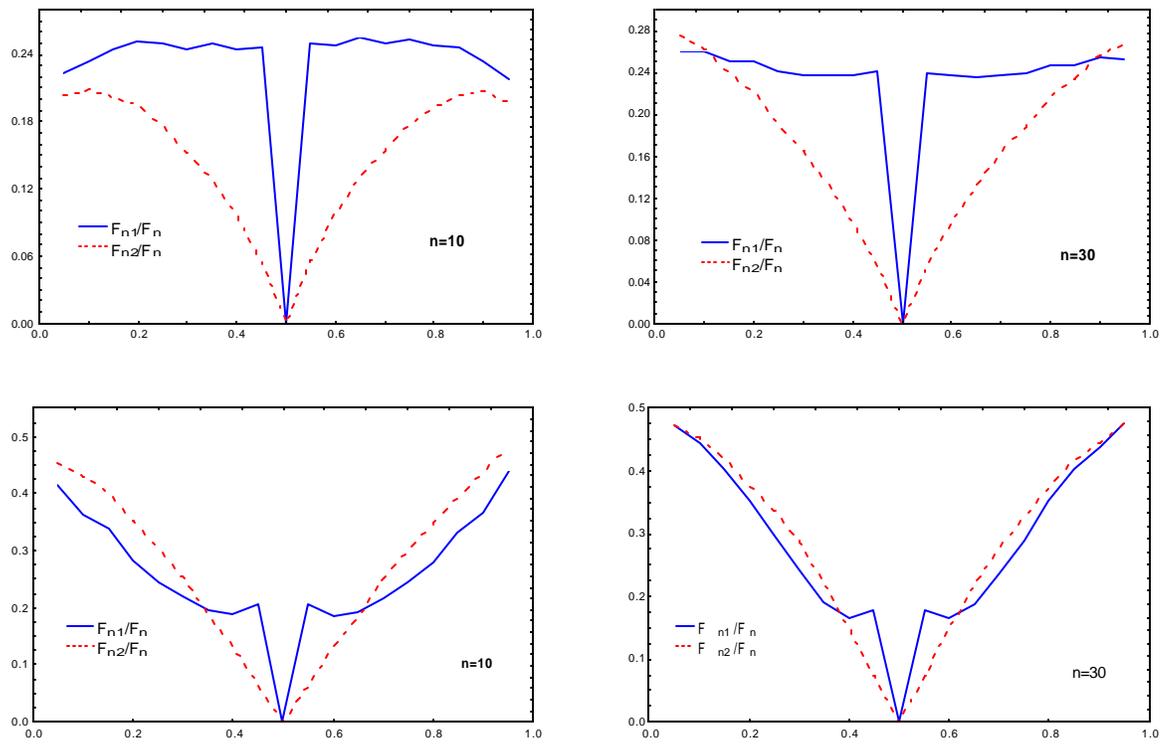


Figura 2-2: Caso una Muestra, Distribución Normal. La fila superior corresponde a $\sigma = 1.1$, la fila inferior a $\sigma = 2$.

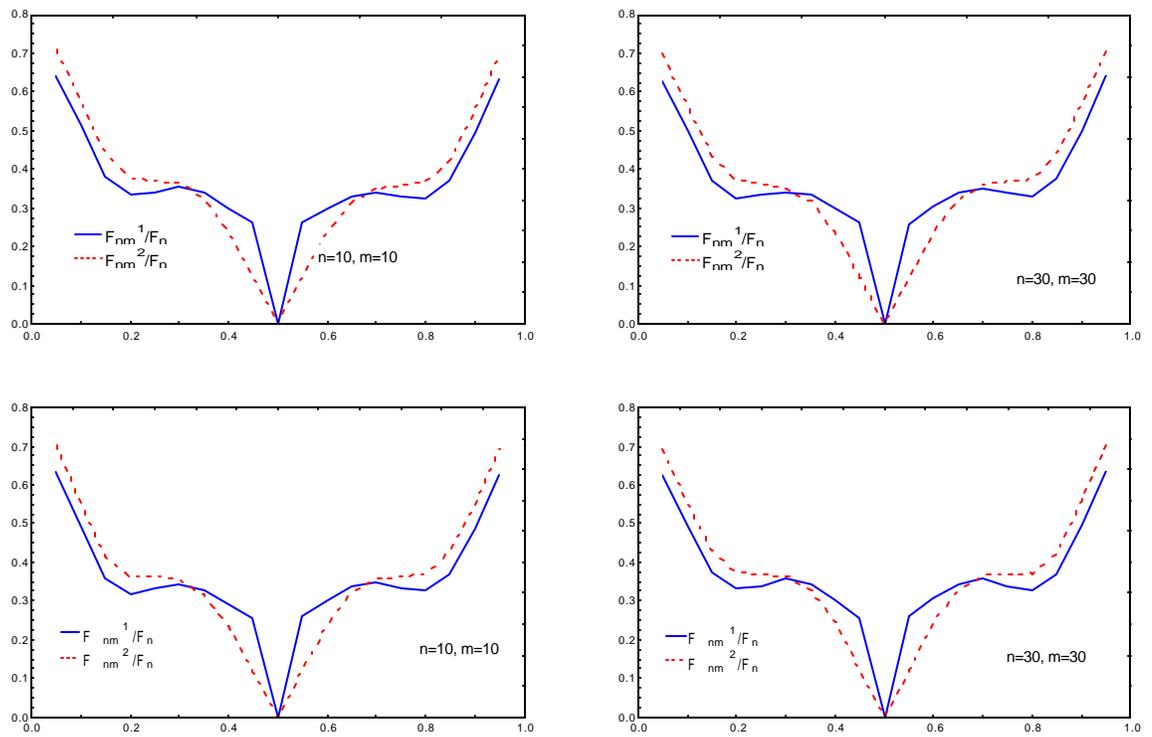


Figura 2-3: Caso dos muestras, Distribución Logística. La fila superior corresponde a $\sigma = 1.1$, y la fila inferior a $\sigma = 2$.

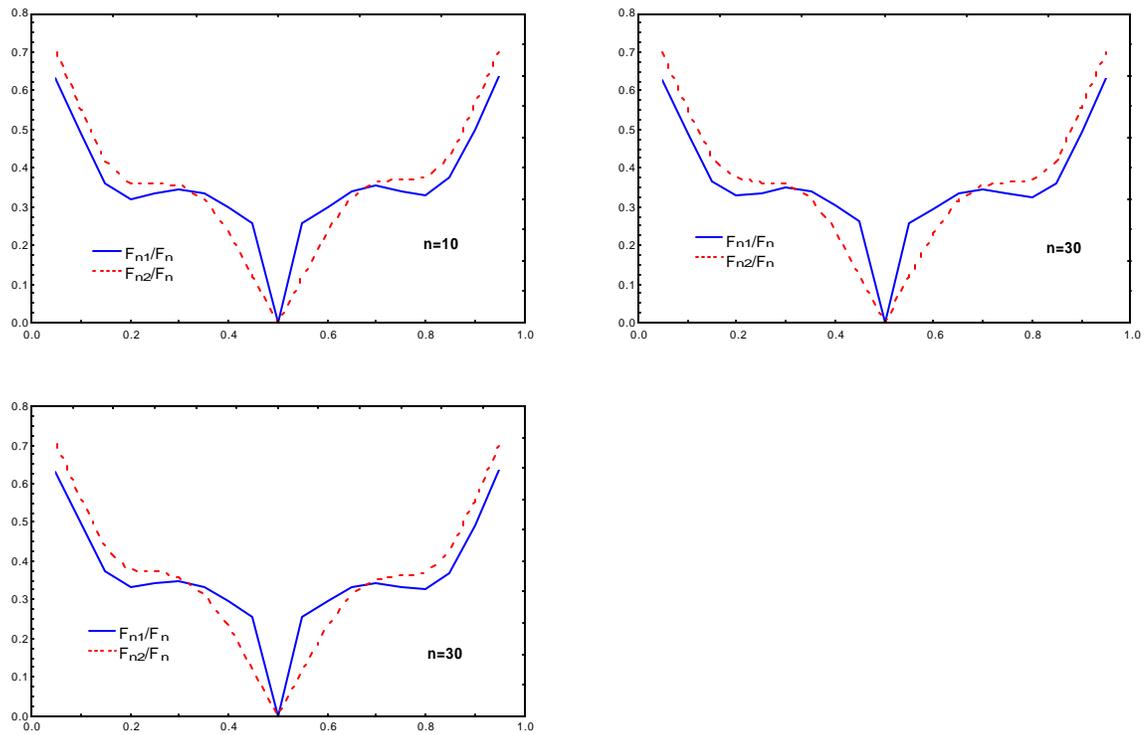


Figura 2-4: Caso con censura: Distribución Normal. Leyendo las gráficas de izquierda a derecha y de arriba a abajo, se presentan los casos para $\sigma = 1.1, 1.5$ y 2 .

nuevos estimadores también superan al estimador Kaplan-Meier, en cuanto al MSE, aunque la ganancia no es tan grande como en el caso de una muestra sin censura.

Para la situación en que tampoco se conoce G , de nuevo nuestros estimadores propuestos son mejores que la función de distribución empírica aunque la ganancia en MSE es menor que cuando se conoce G .

Se hicieron más simulaciones, pero aquí sólo se presentan algunas de las gráficas, pues el comportamiento que siguen las otras, ya sea al cambiar de distribución F , el tamaño de muestra u otro ingrediente, es similar a las gráficas presentadas.

Capítulo 3

Estimación de la Función de Supervivencia Bivariada

En estudios de Análisis de Supervivencia y Confiabilidad es posible encontrar situaciones con datos (tiempos de vida) pareados y censurados por la derecha. Algunos ejemplos son tiempos de vida de parejas casadas, el tiempo desde que se inicia un tratamiento hasta la primera respuesta en dos etapas sucesivas de un tratamiento en el mismo paciente. En estos y otros casos, es de interés estimar la función de supervivencia S (o equivalente la función de distribución F). Fijando ideas, supongamos que tenemos datos de la forma $\{(Y_{1i}, Y_{2i}, \delta_{1i}, \delta_{2i}) : i = 1, 2, \dots, n\}$, donde $Y_{ji} = \min(T_{ji}, Z_{ji})$, $\delta_{ji} = I[T_{ji} = Y_{ji}]$, $j = 1, 2$, $i = 1, 2, \dots, n$; con tiempos de falla $\mathbf{T}_i = (T_{1i}, T_{2i}) \sim F = 1 - S$, y tiempos de censura $\mathbf{Z}_i = (Z_{1i}, Z_{2i}) \sim G$. Problemas de identificabilidad surgen en el momento de trabajar con datos censurados, si no se imponen ciertas restricciones entre la variable de censura y la variable de interés, entonces los datos observados podrían ser originados por diferentes funciones de distribución. En este trabajo y en muchos otros se ha superado el problema de identificabilidad asumiendo que F y G son independientes. En Pruitt (1993) se aborda este tema y proporciona algunas condiciones más relajadas, es decir, aun cuando la independencia de F y G asegura identificabilidad, se puede debilitar esta condición sin perder la identificabilidad. Nosotros asumiremos que F y G son independientes.

Sin pretender dar una reseña completa de todos los intentos hasta ahora conocidos, procederemos a presentar en forma resumida, algunos de los más importantes. En Langberg y Shaked (1982) se ataca el problema descomponiendo la función de supervivencia S en la forma $S(s, t) = P [T_1 > s, T_2 > t] = P [T_1 > s | T_2 > t] P [T_2 > t]$. Así el problema se transforma en estimación univarida de dos funciones de supervivencia. Las propiedades de este estimador fueron desarrolladas por Campbell y Földes (1982), Campbell (1982), Horváth (1983), Burke (1984), Lo y Wang (1986), Horváth y Yandell (1986) entre otros. Este estimador tiene ciertos defectos serios, no siempre es una función de supervivencia debido a que no siempre es monótona decreciente en sus coordenadas, y puede asignar masa negativa a algún rectángulo en el plano. Aunado a esto, este estimador en caso de datos sin censura no se reduce a la función de distribución empírica usual, y depende del orden en el condicionamiento, es decir, no se obtiene el mismo estimador si ahora condicionamos en $T_1 > s$.

Por supuesto que el estimador de máxima verosimilitud no paramétrico no podía faltar, y su definición, así como sus propiedades son considerados en Campbell (1981), Hanley y Parnes (1983) y Muñoz (1980). El estimador resultante no tiene una forma cerrada, además de no ser único, y no hay garantía de que una versión en particular tenga buenas propiedades asintóticas. La razón por la que no es único es que para algún conjunto de datos, en un punto con una coordenada siendo falla y la otra siendo observación censurada, el método de máxima verosimilitud no sabe como asignar masa a la semirecta que empieza en la componente censurada. Van der Laan (1996) retoma el estimador de máxima verosimilitud no paramétrica, tratando de repararlo, mostrando como asignar la masa en los puntos problemáticos, sin embargo, de nuevo se tiene un estimador sin una forma cerrada y bastante intratable.

Otras versiones involucran representaciones de la función de supervivencia bivariada en términos de ecuaciones integrales, cuya solución se da en terminos de Series de Peano y por tanto el estimador se vuelve intratable. Ver Pons (1986).

Muchos autores (Ver Oakes, 2001) coinciden en que las mejores soluciones son hasta ahora los estimadores dados en Dabrowska (1988) y en Prentice y Cai (1992), que se describirán en las siguientes secciones. Estos estimadores son uniformemente consistentes y en el caso del estimador de Dabrowska, se tiene convergencia débil. Las propuestas, es decir, nuestros estimadores que se presentan en esta sección, son estimadores que no tienen los defectos de los estimadores

Prentice-Cai y Dabrowska y son mejores en cuanto al error cuadrático medio.

A continuación describiremos dos de los estimadores presentados en la literatura, que tratan de dar solución a este problema.

3.1 El estimador de Dabrowska

Sin duda uno de los estimadores mas importantes y conocidos de la función de supervivencia bivariada bajo censura aleatoria es el dado en Dabrowska (1988). Este estimador está basado en la representación (1.3) de S como integral producto en el plano de la función L . Esta función L está definida en la Proposición 7 y es una función que incorpora la dependencia que hay entre los tiempos de vida, conjuntando las funciones de riesgo acumuladas tanto univariadas como la bivariada.

A continuación se presenta la función L en su versión empírica, como en Dabrowska(1988). Mientras no se especifique lo contrario, el rango de las sumas es de 1, hasta n . Sean

$$\begin{aligned}\widehat{H}(s, t) &= n^{-1} \sum I [Y_{1i} > s, Y_{2i} > t], \\ \widehat{K}_1(s, t) &= n^{-1} \sum I [Y_i > s, Y_{2i} > t, \delta_{1i} = 1, \delta_{2i} = 1], \\ \widehat{K}_2(s, t) &= n^{-1} \sum I [Y_{1i} > s, Y_{2i} > t, \delta_{1i} = 1] \text{ y} \\ \widehat{K}_3(s, t) &= n^{-1} \sum I [Y_{1i} > s, Y_{2i} > t, \delta_{2i} = 1].\end{aligned}$$

Entonces, un estimador de la función bivariada de riesgo acumulado está dado por:

$$\begin{aligned}\widehat{\Lambda}_{11}(s, t) &= \int_0^s \int_0^t \frac{\widehat{K}_1(du, dv)}{\widehat{H}(u-, v-)}, \\ \widehat{\Lambda}_{10}(s, t) &= - \int_0^s \frac{\widehat{K}_2(du, t)}{\widehat{H}(u-, t)}, \\ \widehat{\Lambda}_{01}(s, t) &= - \int_0^t \frac{\widehat{K}_3(s, dv)}{\widehat{H}(s, v-)}.\end{aligned}\tag{3.1}$$

y un estimador de la función L es:

$$\widehat{L}(\Delta t_{1i}, \Delta t_{2j}) = \frac{\widehat{\Lambda}_{10}(\Delta t_{1i}, t_{2j-}) \widehat{\Lambda}_{01}(t_{1i-}, \Delta t_{2j}) - \widehat{\Lambda}_{11}(\Delta t_{1i}, \Delta t_{2j})}{\left[1 - \widehat{\Lambda}_{10}(\Delta t_{1i}, t_{2j-})\right] \left[1 - \widehat{\Lambda}_{01}(t_{1i-}, \Delta t_{2j})\right]}\tag{3.2}$$

con t_{1i}, t_{2j} los tiempos de falla observados ordenados, y $\widehat{L}(\Delta t, \Delta s) = 0$ en otro lugar. Es decir la función L sólo puede tener saltos en puntos de la forma (u, v) , con ambas coordenadas tiempos de falla observados.

Utilizando la representación (1.3) se obtiene un estimador para la función de supervivencia:

$$\widehat{S}_D(s, t) = \widehat{S}(s, 0)\widehat{S}(0, t) \prod_{t_{1i} \leq s, t_{2j} \leq t} \left[1 - \widehat{L}(\Delta t_{1i}, \Delta t_{2j}) \right]. \quad (3.3)$$

Donde $\widehat{S}(s, 0)$ y $\widehat{S}(0, t)$ son los estimadores Kaplan-Meier univariados respectivos.

Este estimador es de forma escalonada, y asigna masa en los puntos (t_{1i}, t_{2j}) , y en caso de que la muestra no tenga datos censurados, se reduce a la función de supervivencia empírica. Sus marginales son los estimadores Kaplan-Meier univariados respectivos, y por esto es que a este estimador se le conoce también como el estimador Kaplan-Meier bivariado. Bajo condiciones muy generales, se tiene consistencia uniforme, según la siguiente proposición:

Proposición 37 (*Proposición 4.1 en Dabrowska, 1988*). Sea G la función de distribución del vector censura, y $H = SG$. Entonces, bajo censura aleatoria, y para todo (τ_1, τ_2) , con $H(\tau_1, \tau_2) > 0$,

$$\sup_{[0, \tau_1] \times [0, \tau_2]} \left| \widehat{S}_D - S \right| \rightarrow 0$$

cuando $n \rightarrow \infty$, con probabilidad 1.

Demostración. Ver prueba en Dabrowska(1988). ■

Más aún, en Dabrowska (1989) se prueba la convergencia débil del estimador, así como una ley del logaritmo iterado.

Dadas estas características de este estimador, procederemos a desarrollar fórmulas que faciliten su cálculo. En principio, desarrollando las expresiones (3.1) y sustituyendo en (3.2), se obtiene la siguiente fórmula útil en el cálculo de la función \widehat{L} . Esta fórmula se utilizará en los estudios de simulación presentados más adelante.

Proposición 38 Sean $n(s, t) = n\widehat{H}(s, t)$, $d(s, t) = \sum I[Y_{1i} = s, Y_{2i} = t, \delta_{1i} = \delta_{2i} = 1]$, $d_1(s, t) =$

$\sum I [Y_{1i} = s, Y_{2i} \geq t, \delta_{1i} = 1]$ y $d_2(s, t) = I [Y_{1i} \geq s, Y_{2i} = t, \delta_{2i} = 1]$. Entonces,

$$\widehat{L}(\Delta s, \Delta t) = \begin{cases} \frac{d_1(s, t)d_2(s, t) - n(s, t)d(s, t)}{(n(s, t) - d_1(s, t))(n(s, t) - d_2(s, t))} & \text{si } (s, t) = (t_{1i}, t_{2j}) \text{ para alg\u00fan } i, j \\ & \text{y } n(s, t) \neq d_1(s, t), n(s, t) \neq d_2(s, t) \\ 0 & \text{de otra manera} \end{cases} \quad (3.4)$$

Demostraci\u00f3n. La funci\u00f3n $\widehat{K}_2(\cdot, t)$ es de tipo escalonada con saltos en los puntos de muerte observados t_{1j} , de tama\u00f1o $-n^{-1} \sum_{j=1}^n I [Y_{1j} = t_{1i}, Y_{2j} > t, \delta_{1j} = 1]$, la funci\u00f3n $\widehat{K}_3(s, \cdot)$ tiene saltos en t_{2j} , de magnitud $-n^{-1} \sum_{i=1}^n I [Y_{1i} > s, Y_{2i} = t_{2j}, \delta_{2i} = 1]$ y la funci\u00f3n $\widehat{K}_1(\cdot, \cdot)$ en el punto (t_{1i}, t_{2j}) asigna masa $n^{-1} \sum_{k=1}^n I [Y_{1k} = t_{1i}, Y_{2k} = t_{2j}, \delta_{1k} = \delta_{2k} = 1]$. Entonces,

$$\begin{aligned} \widehat{\Lambda}_{10}(s, t) &= - \int_0^s \frac{\widehat{K}_2(du, t)}{\widehat{H}(u-, t)} \\ &= \sum_{t_{1i} \leq s} \left(\sum_{a=1}^n I [Y_{1a} \geq t_{1i}, Y_{2a} > t] \right)^{-1} \left(\sum_{a=1}^n I [Y_{1a} = t_{1i}, Y_{2a} > t, \delta_{1a} = 1] \right) \end{aligned}$$

y por tanto,

$$\widehat{\Lambda}_{10}(\Delta s, t-) = \begin{cases} d_1(t_{1i}, t) (\sum_{a=1}^n I [Y_{1a} = t_{1i}, Y_{2a} \geq t])^{-1} & \text{si } s = t_{1i} \\ 0 & \text{en otro caso.} \end{cases}$$

An\u00e1logamente,

$$\widehat{\Lambda}_{10}(s-, \Delta t) = \begin{cases} d_2(s, t_{2j}) (\sum_{a=1}^n I [Y_{1a} \geq s, Y_{2a} = t_{2j}])^{-1} & \text{si } s = t_{2j} \\ 0 & \text{en otro caso,} \end{cases}$$

y

$$\widehat{\Lambda}_{11}(\Delta s, \Delta t) = \begin{cases} d(t_{1i}, t_{2j}) (\sum_{a=1}^n I [Y_{1a} \geq s, Y_{2a} \geq t_{2j}])^{-1} & \text{si } s = t_{1i} \text{ y } t = t_{2j} \\ 0 & \text{en otro caso.} \end{cases}$$

Entonces, sustituyendo estas tres \u00faltimas expresiones en (3.2), se obtiene el resultado deseado.

■

La funci\u00f3n \widehat{L} satisface las siguientes desigualdades:

Corolario 39 Sea \widehat{L} la función definida por (3.2), entonces:

- i) $1 - \widehat{L}(\Delta s, \Delta t) \geq 0$,
- ii) Si $d(s, t) = 0$, entonces $1 - \widehat{L}(\Delta s, \Delta t) \leq 1$,
- iii) Si $d(s, t) > 0$, entonces $1 - \widehat{L}(\Delta s, \Delta t) \geq 1$.

Demostración.

i) Se sigue de la igualdad

$$1 - \widehat{L}(\Delta s, \Delta t) = \frac{n(s, t) [n(s, t) - d_1(s, t) - d_2(s, t) + d(s, t)]}{(n(s, t) - d_1(s, t)) (n(s, t) - d_2(s, t))}$$

y la desigualdad $n(s, t) \geq d_1(s, t) + d_2(s, t) - d(s, t)$.

ii) Desarrollando la expresión en el denominador de (3.4), se obtiene

$$1 - \widehat{L}(\Delta s, \Delta t) = \frac{n(s, t) [n(s, t) - d_1(s, t) - d_2(s, t)]}{n(s, t) [n(s, t) - d_1(s, t) - d_2(s, t)] + d_1(s, t)d_2(s, t)} \leq 1$$

iii) De la igualdad $n(s, t) = d_1(s, t) + d_2(s, t) - d(s, t) + n\widehat{H}(s, t)$, obtenemos

$$\begin{aligned} n(s, t)d(s, t) - d_1(s, t)d_2(s, t) &= d(s, t) \left[d_1(s, t) + d_2(s, t) + n\widehat{H}(s, t) \right] \\ &\quad - d^2(s, t) - d_1(s, t)d_2(s, t) \\ &= (d(s, t) - d_1(s, t)) (d(s, t) - d_2(s, t)) + nd(s, t)\widehat{H}(s, t) \\ &\geq 0, \end{aligned}$$

es decir, $n(s, t)d(s, t) \geq d_1(s, t)d_2(s, t)$. Aplicando esta última desigualdad, obtenemos

$$1 - \widehat{L}(\Delta s, \Delta t) = \frac{n(s, t) [n(s, t) - d_1(s, t) - d_2(s, t)] + n(s, t)d(s, t)}{n(s, t) [n(s, t) - d_1(s, t) - d_2(s, t)] + d_1(s, t)d_2(s, t)} \geq 1$$

■

La expresión anterior es útil pues el cálculo de la función \widehat{L} se ha reducido a conteos de muertes y observaciones en riesgo, como en el caso de una variable y el estimador Kaplan-Meier.

Un ejemplo dado en Dabrowska (1988), página 1485, muestra como con una muestra bivariada de 4 datos, resulta un estimador de la función de supervivencia que no es propia, es decir, asigna masa negativa y no es monótona no creciente. Este es único inconveniente de este estimador y no existe una explicación clara y lógica del porque sucede esto. Dabrowska sugiere que es el resultado de la combinación de información (para el cálculo de la función \widehat{L}) proveniente de estimadores univariados y bivariados de las funciones de riesgo acumulados. Pareciera que hay una contradicción en el desempeño de este estimador, pues mientras el estimador es uniforme fuertemente consistente, el número de puntos donde asigna masa negativa no decrece al aumentar el tamaño de muestra, según Pruitt (1991), quien caracteriza todos los posibles casos donde se puede presentar un punto con masa positiva y demuestra que esta cantidad de puntos no decrece cuando el tamaño de muestra crece a ∞ .

Con esto en mente nos dimos a la tarea de encontrar modificaciones al estimador (3.3) que conserven sus propiedades de consistencia, pero que sean funciones de supervivencia propias. Más adelante se presentan estos nuevos estimadores, ahora introduzcamos otro de los estimadores importantes.

3.2 El estimador de Prentice y Cai

La representación dada en Prentice-Cai (1992) de la función S mediante (1.4) y (1.5) permite construir un estimador de la función de supervivencia como sigue: sean $\widehat{\Lambda}_i$, $i = 1, 2$, los estimadores usuales de la función de riesgo acumulada para T_1 y T_2 , es decir, $\widehat{\Lambda}_i(t) = \sum_{t_{ij} \leq t} d_{ij} n_{ij}^{-1}$, con d_{ij} denotando al número de fallas en t_{ij} y n_{ij} el número de sujetos en riesgo justo antes de t_{ij} . La fórmula recursiva para el estimador de Prentice-Cai esta dada por

$$\begin{aligned} \widehat{S}_{pc}(s, t) &= 1 - \sum_{t_{1i} \leq s} \sum_{t_{2j} \leq t} \widehat{S}_{pc}(t_{1i}^-, t_{2j}^-) \left\{ \frac{d_{1i}}{n_{1i}} \frac{d_{2j}}{n_{2j}} - \widehat{a}(t_{1i}, t_{2j}) \right\} \\ &\quad - \sum_{t_{1i} \leq s} \widehat{S}_{pc}(t_{1i}^-, t) \frac{d_{1i}}{n_{1j}} - \sum_{t_{2j} \leq t} \widehat{S}_{pc}(s, t_{2j}^-) \frac{d_{2j}}{n_{2j}} \end{aligned} \quad (3.5)$$

con $\widehat{a}(s, t) = r(s, t)^{-1} \sum_{l \in R(s, t)} \left(I [T_{1l} \leq s] - \widehat{\Lambda}_1(s \wedge T_{1l}) \right) \left(I [T_{2l} \leq t] - \widehat{\Lambda}_2(t \wedge T_{2l}) \right)$, $R(s, t)$ es el conjunto de observaciones en riesgo justo antes de (s, t) , y $r(s, t) = \#R(s, t)$.

Este estimador, al igual que el de Dabrowska, presenta saltos en los puntos de doble falla observados, y en ausencia de censura se reduce a la función de supervivencia empírica. Es uniformemente consistente, pero al igual que el esimador 3.3, no siempre es una función de supervivencia propia, ya que en ocasiones asigna masa negativa o no es monótona no creciente. Los autores sugieren una modificación al estimador (3.5), como sigue:

$$\widehat{S}_{pc}^*(s, t) = \begin{cases} \min \left\{ \widehat{S}_{pc}^*(s-, t), \widehat{S}_{pc}^*(s, t-) \right\} & \text{si } \widehat{S}_{pc}(s, t) > \min \left\{ \widehat{S}_{pc}^*(s-, t), \widehat{S}_{pc}^*(s, t-) \right\} \\ \widehat{S}_{pc}(s, t) & \text{en otro caso} \\ \widehat{S}_{pc}^*(s-t) + \widehat{S}_{pc}^*(s, t-) - \widehat{S}_{pc}^*(s-, t-) & \text{si } \widehat{S}_{pc}(s, t) < \widehat{S}_{pc}^*(s-t) + \widehat{S}_{pc}^*(s, t-) \\ & - \widehat{S}_{pc}^*(s-, t-). \end{cases} \quad (3.6)$$

Los autores prueban la consistencia del estimador (3.5), pero no prueban la consistencia del estimador modificado (3.6), quedando éste como un problema abierto.

Dadas las descripciones de los más importantes estimadores de la función de supervivencia S que se pueden encontrar en la literatura, procederemos a presentar nuestros estimadores.

3.3 Nuevos estimadores de la función de supervivencia S

Como se especifica en el ejemplo dado en Dabrowska (1988), el estimador (3.3) no siempre es una función de supervivencia, ya que en ocasiones, asigna masa negativa a algunos rectángulos en el plano o no es monótona decreciente en algún componente. Pruitt (1991) bajo el supuesto de que la máxima observación en cada eje coordenado es muerte, identifica y caracteriza los puntos donde el estimador (3.3) asigna masa negativa. Veamos otra forma de caracterizar los puntos donde un estimador de la forma (3.3), es decir, un estimador con saltos en los puntos de doble muerte y que tenga la forma

$$\widehat{P}(s, t) = \widehat{S}(s, 0)\widehat{S}(0, t) \prod_{t_{1i} \leq s, t_{2j} \leq t} \left[1 - \widehat{M}(\Delta t_{1i}, \Delta t_{2j}) \right] \quad (3.7)$$

con $\widehat{S}(s, 0)$ y $\widehat{S}(0, t)$, las marginales Kaplan-Meier univariadas respectivas, asigna masa negativa o es no decreciente.

Lema 40 Sean \widehat{M} y \widehat{P} , funciones definidas sobre los puntos de doble muerte, mediante la

ecuación (3.7). Supongamos que $\widehat{P} \geq 0$. Entonces para (s, t) punto de salto de la función \widehat{P} ,

1. \widehat{P} asigna masa positiva en (s, t) , si y sólo si,

$$\widehat{M}(\Delta s, \Delta t) \leq \left(1 - \frac{\widehat{P}(s-, t-)}{\widehat{P}(s, t-)}\right) \left(1 - \frac{\widehat{P}(s-, t-)}{\widehat{P}(s-, t)}\right) \quad (3.8)$$

2. \widehat{P} es monótona no creciente en cada componente en el punto (s, t) si y sólo si,

$$\max \left\{ 1 - \frac{\widehat{P}(s-, t-)}{\widehat{P}(s, t-)}, 1 - \frac{\widehat{P}(s-, t-)}{\widehat{P}(s-, t)} \right\} \leq \widehat{M}(\Delta s, \Delta t) \quad (3.9)$$

Demostración. De (3.7), se obtiene la ecuación

$$\widehat{P}(s, t) = \left[1 - \widehat{M}(\Delta s, \Delta t)\right] \frac{\widehat{P}(s, t-)\widehat{P}(s-, t)}{\widehat{P}(s-, t-)} \quad (3.10)$$

1. La masa en el punto (s, t) es igual a $\widehat{P}(\Delta s, \Delta t) = \widehat{P}(s, t) - \widehat{P}(s-, t) - \widehat{P}(s, t-) + \widehat{P}(s-, t-)$, y es no negativa, si y sólo si,

$$1 - \widehat{M}(\Delta s, \Delta t) \geq \frac{\widehat{P}(s-, t-)}{\widehat{P}(s, t-)\widehat{P}(s-, t)} \left(\widehat{P}(s-, t) + \widehat{P}(s, t-) - \widehat{P}(s-, t-)\right)$$

es decir,

$$\begin{aligned} \widehat{M}(\Delta s, \Delta t) &\leq 1 - \frac{\widehat{P}(s-, t-)}{\widehat{P}(s, t-)\widehat{P}(s-, t)} \left(\widehat{P}(s-, t) + \widehat{P}(s, t-) - \widehat{P}(s-, t-)\right) \\ \widehat{M}(\Delta s, \Delta t) &\leq \left(1 - \frac{\widehat{P}(s-, t-)}{\widehat{P}(s, t-)}\right) \left(1 - \frac{\widehat{P}(s-, t-)}{\widehat{P}(s-, t)}\right). \end{aligned}$$

2. Es fácil ver de (3.10) que $\widehat{P}(s, t) \leq \widehat{P}(s, t-)$ si y sólo si, $1 - \widehat{M}(\Delta s, \Delta t) \leq \frac{\widehat{P}(s-, t-)}{\widehat{P}(s-, t)}$, es decir, $\widehat{M}(\Delta s, \Delta t) \geq 1 - \frac{\widehat{P}(s-, t-)}{\widehat{P}(s-, t)}$. Análogamente, se prueba que $\widehat{P}(s, t) \leq \widehat{P}(s-, t)$ si y sólo si, $\widehat{M}(\Delta s, \Delta t) \geq 1 - \frac{\widehat{P}(s-, t-)}{\widehat{P}(s, t-)}$. Conjuntando las dos desigualdades, se obtiene el resultado deseado.

■

Teniendo en cuenta las condiciones necesarias y suficientes dadas en el Lema anterior, podemos construir un estimador de la función de supervivencia S , con marginales Kaplan-Meier univariados, modificando el estimador \widehat{S}_D mediante la siguiente ecuación:

$$\widehat{S}_B(s, t) = \widehat{S}(s, 0)\widehat{S}(0, t) \prod_{t_{1i} \leq s, t_{2j} \leq t} \left[1 - \widehat{J}(\Delta t_{1i}, \Delta t_{2j}) \right] \quad (3.11)$$

con

$$\widehat{J}(\Delta s, \Delta t) = \begin{cases} \max \left\{ 1 - \frac{S_B(s-, t-)}{S_B(s, t-)}, 1 - \frac{S_B(s-, t-)}{S_B(s-, t)} \right\} & \text{si } \widehat{L} \text{ y } \widehat{S}_B \text{ no satisfacen 3.9 en } (s, t), \\ \widehat{L}(\Delta s, \Delta t) & \text{si } \widehat{L} \text{ y } \widehat{S}_B \text{ satisfacen 3.9 y 3.8 en } (s, t), \\ \left(1 - \frac{S_B(s-, t-)}{S_B(s, t-)} \right) \left(1 - \frac{S_B(s-, t-)}{S_B(s-, t)} \right) & \text{si } \widehat{L} \text{ y } \widehat{S}_B \text{ no satisfacen 3.8 en } (s, t). \end{cases}$$

Note de la expresión anterior que si $\widehat{J}(\Delta s, \Delta t) = \max \left\{ 1 - \frac{S_B(s-, t-)}{S_B(s, t-)}, 1 - \frac{S_B(s-, t-)}{S_B(s-, t)} \right\}$ entonces $\widehat{S}_B(s, t) = \min \left\{ \widehat{S}_B(s-, t), \widehat{S}_B(s, t-) \right\}$, y $\widehat{S}_B(s, t) = \widehat{S}_B(s-, t) + \widehat{S}_B(s, t-) - \widehat{S}_B(s-, t-)$ si $\widehat{J}(\Delta s, \Delta t) = \left(1 - \frac{S_B(s-, t-)}{S_B(s, t-)} \right) \left(1 - \frac{S_B(s-, t-)}{S_B(s-, t)} \right)$. O sea, esta modificación es similar a la utilizada en Prentice y Cai (1992).

Otra forma de modificar el estimador de Dabrowska, de tal forma que asigne masa positiva y sea monótona no decreciente en cada componente, se presenta a continuación. Se define el estimador \widehat{S}_R igual a las marginales Kaplan-Meier, es decir, $\widehat{S}_R(s, 0) = \widehat{S}(s, 0)$ y $\widehat{S}_R(0, t) = \widehat{S}(0, t)$, y a partir de éstos se calcula de forma recursiva en cada punto de doble muerte mediante la fórmula,

$$\widehat{S}_R(s, t) = \begin{cases} \widehat{S}_R(s-, t) + \widehat{S}_R(s, t-) - \widehat{S}_R(s-, t-) & \text{si } \widehat{S}_D(s, t) < \widehat{S}_R(s-, t) + \widehat{S}_R(s, t-) \\ & - \widehat{S}_R(s-, t-) \\ \widehat{S}_D(s, t) & \text{en otro caso} \\ \min \left\{ \widehat{S}_R(s-, t), \widehat{S}_R(s, t-) \right\} & \text{si } \widehat{S}_D(s, t) > \min \left\{ \widehat{S}_R(s-, t), \widehat{S}_R(s, t-) \right\}. \end{cases} \quad (3.12)$$

Así, dado que $\widehat{S}_R(s-, t) + \widehat{S}_R(s, t-) - \widehat{S}_R(s-, t-) \leq \min \left\{ \widehat{S}_R(s-, t), \widehat{S}_R(s, t-) \right\}$, entonces se satisface la desigualdad

$$\widehat{S}_R(s-, t) + \widehat{S}_R(s, t-) - \widehat{S}_R(s-, t-) \leq \widehat{S}_R(s, t) \leq \min \left\{ \widehat{S}_R(s-, t), \widehat{S}_R(s, t-) \right\},$$

que garantiza las condiciones de monotonía propias de una función de supervivencia. Esta idea, aun cuando es sencilla, es suficiente para obtener un nuevo estimador, que es función de supervivencia propia. Como se observa en la definición, solo modifica al estimador \widehat{S}_D en los puntos donde existe una violación a la definición de función de supervivencia.

Los nuevos estimadores \widehat{S}_B y \widehat{S}_R no son iguales, ya que las modificaciones de las cuales provienen son muy diferentes. Por último, para comparar los estimadores (3.3), (3.5), (3.11) y (3.12), se procedió a calcular el error cuadrático medio de estos y compararlos, según se presenta en la siguiente sección.

3.4 Estudios de Simulación comparando el MSE de los nuevos estimadores con los existentes.

A fin de comparar el desempeño de estos dos estimadores \widehat{S}_R y \widehat{S}_B con los estimadores de Dabrowska y Prentice-Cai, se hicieron corridas de un programa simulando muestras bivariadas de distintos modelos, y censurando estas muestras de forma independiente mediante la distribución exponencial estándar, es decir se generaron n muestras (T_1, T_2) de alguna distribución bivariada conocida, y se censuran con un vector de dos exponenciales estándar independientes. Las simulaciones se hicieron teniendo como distribución para (T_1, T_2) a alguna de las siguientes:

1. Distribución Gumbel. $F(s, t) = 1 - e^{-s} - e^{-t} + e^{-(s+t+\theta st)}$, $s, t > 0$, $0 \leq \theta \leq 1$.

Funcion de supervivencia: $S(s, t) = \exp(-s - t - \theta st)$

Marginales: exponencial estándar.

2. Distribución Marshall-Olkin. $F(s, t) = \exp(-\lambda_1 s - \lambda_2 t - \lambda_{12} \max(s, t)) - \exp(-(\lambda_1 + \lambda_{12}) s) - \exp(-(\lambda_2 + \lambda_{12}) t) + 1$.

Función de supervivencia: $S(s, t) = \exp(-\lambda_1 s - \lambda_2 t - \lambda_{12} \max(s, t))$

Marginales: $Exp((\lambda_1 + \lambda_{12})^{-1})$ y $Exp((\lambda_2 + \lambda_{12})^{-1})$.

Correlación entre las variables= $\lambda_{12}/(\lambda_1 + \lambda_2 + \lambda_{12})$

3. Distribución Morgenstern: $F(s, t) = F_1(s)F_2(t) [1 + \alpha(1 - F_1(s))(1 - F_2(s))]$, $|\alpha| \leq 1$.

Función de supervivencia: $S(s, t) = F_1(s)F_2(t) [1 + \alpha(1 - F_1(s))(1 - F_2(s))] + 1$

Marginales: F_1, F_2 .

4. Distribución Pareto. $F(s, t) = (F_1(s)^{-1/\alpha} + F_2(t)^{-1/\alpha} - 1)^{-\alpha}$, $\alpha > 0$.

Función de supervivencia: $S(s, t) = (F_1(s)^{-1/\alpha} + F_2(t)^{-1/\alpha} - 1)^{-\alpha} - F_1(s) - F_2(t) + 1$.

Marginales: F_1, F_2 .

5. Distribución Clayton. $F(s, t) = (\exp(\frac{s}{\theta}) + \exp(\frac{t}{\theta}) - 1)^{-\theta} - \exp(-s) - \exp(-t) + 1$, $\theta \geq 0$.

Función de supervivencia: $S(s, t) = (\exp(\frac{s}{\theta}) + \exp(\frac{t}{\theta}) - 1)^{-\theta}$.

Marginales: $Exp(1)$.

Cada gráfica es el resultado de 10 000 simulaciones, en cada una de las cuales se calculan todos los estimadores en puntos de la forma (s, t) tales que $F_1(s), F_2(t) \in \{0.05, 0.1, \dots, 0.95\}$ con F_1 y F_2 las marginales de la función de distribución F . Al final se obtienen los valores del error cuadrático medio según método dado en (2.10). Los resultados indican una ganancia en cuanto al MSE, al utilizar nuestros estimadores con respecto \widehat{S}_{pc} y \widehat{S}_D , como se puede ver en las gráficas, donde se observa una superficie por arriba del plano horizontal 1. Se presentan ejemplos con algunas de las distribuciones bivariadas, pero el comportamiento de las gráficas es similar para las otras distribuciones.

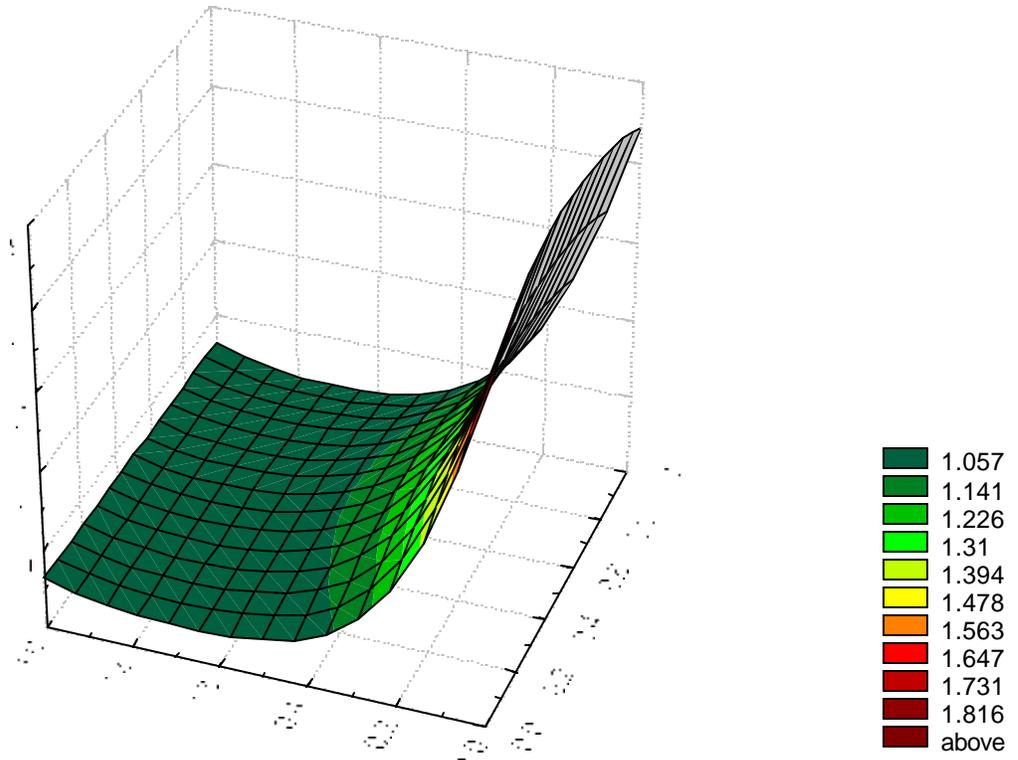


Figura 3-1: Datos generados mediante el modelo Marshall-Olkin con parámetros $\lambda_1 = \lambda_2 = 1$ y $\lambda_{12} = 2$. Se muestra el cociente $MSE(\hat{S}_D)/MSE(\hat{S}_B)$, $n=20$.

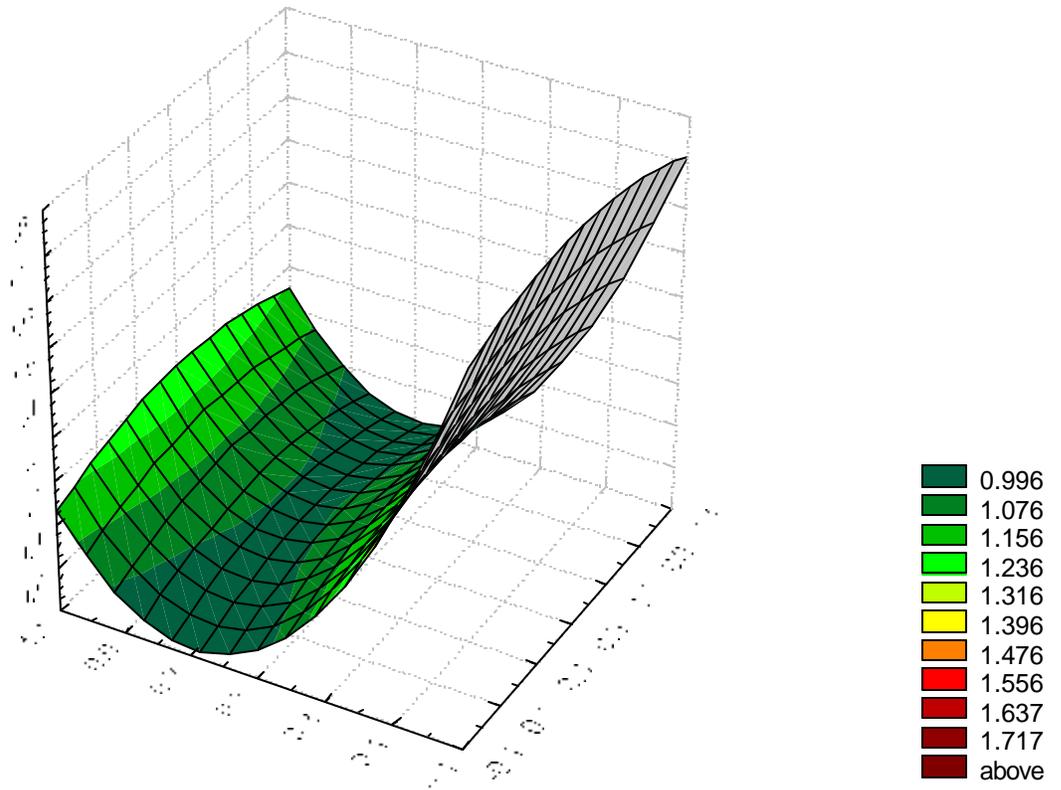


Figura 3-2: Datos generados mediante el modelo Marshall-Olkin con parámetros $\lambda_1 = \lambda_2 = 1$ y $\lambda_{12} = 2$. Se muestra el cociente $MSE(\hat{S}_{pc}) / MSE(\hat{S}_B)$, $n=20$.

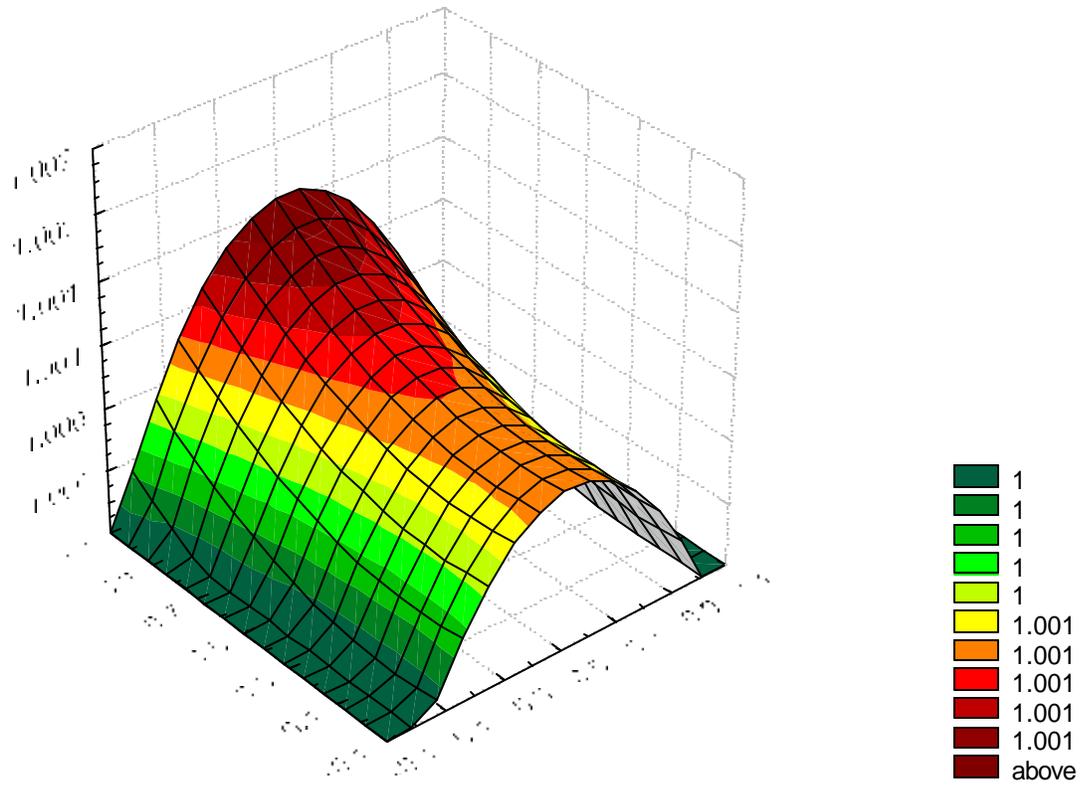


Figura 3-3: Datos generados mediante el modelo Morgersten(0.8). Se muestra el cociente $MSE(\widehat{S}_D)/MSE(\widehat{S}_R)$, $n=10$.

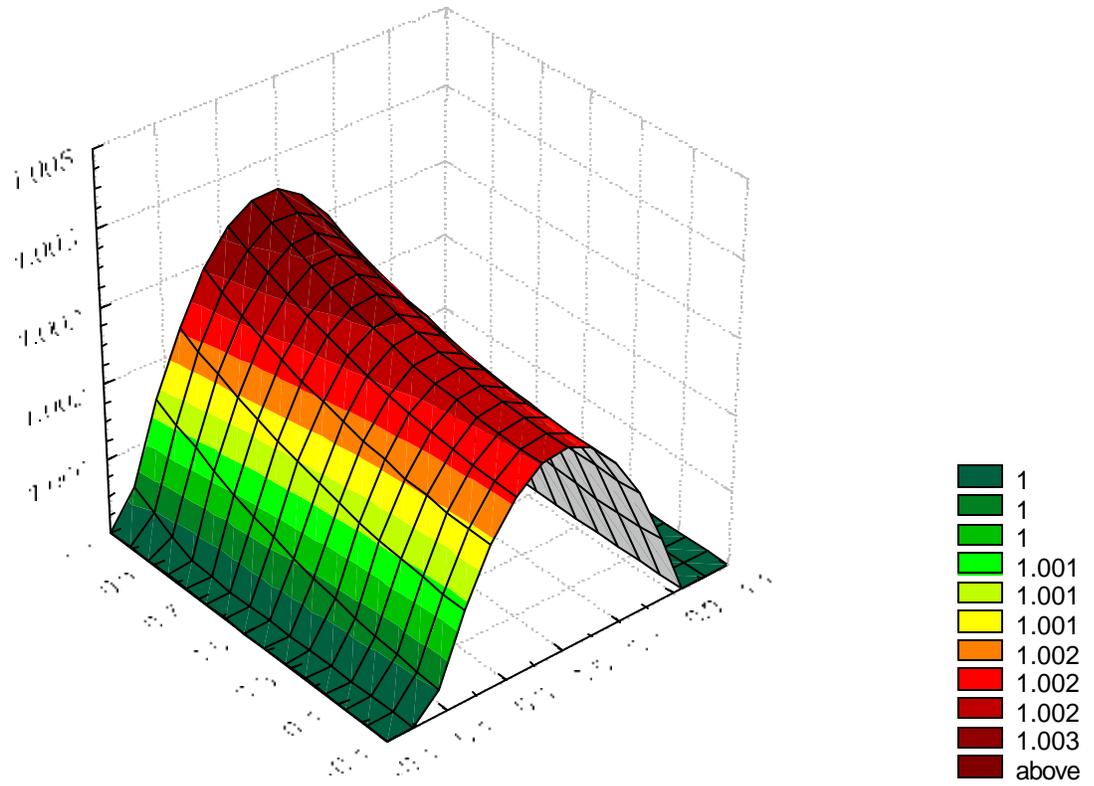


Figura 3-4: Datos generados mediante el modelo Morgersten(0.8). Se muestra el cociente $MSE(\widehat{S}_{pc})/MSE(\widehat{S}_R)$, $n=10$.

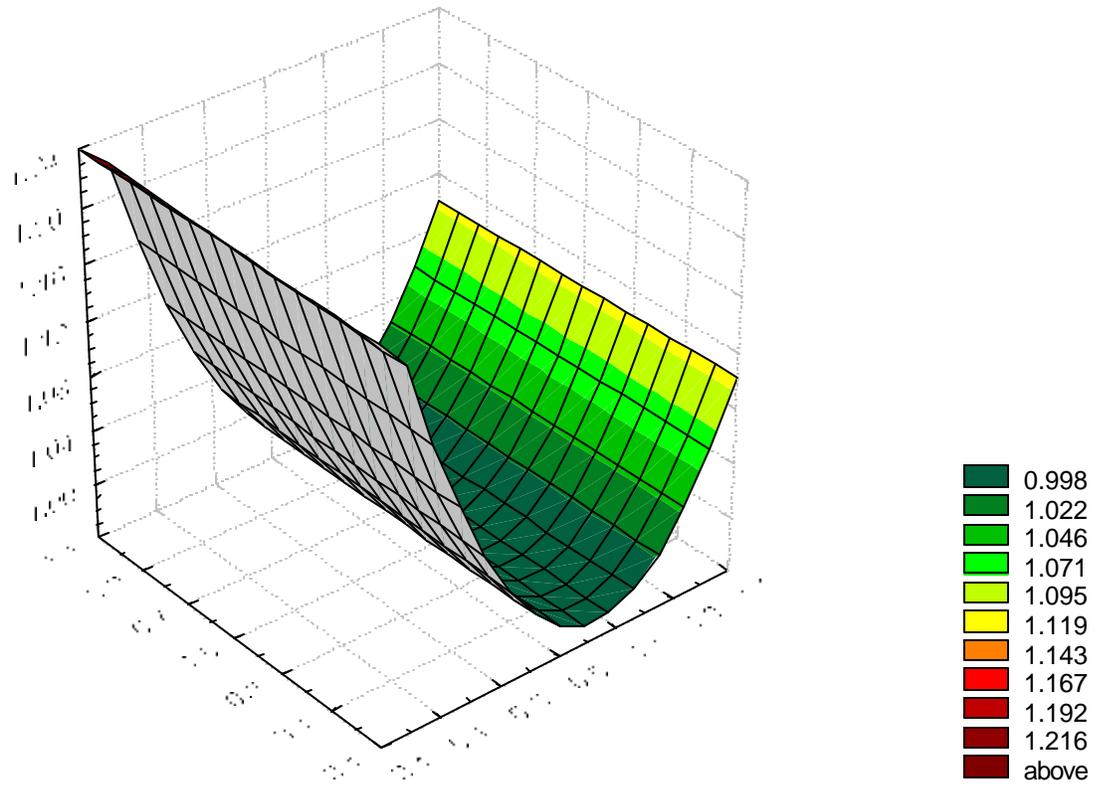


Figura 3-5: Datos generados mediante el modelo Gumbel(0.8). Se muestra el cociente $MSE(\hat{S}_D) / MSE(\hat{S}_R)$, $n=20$.

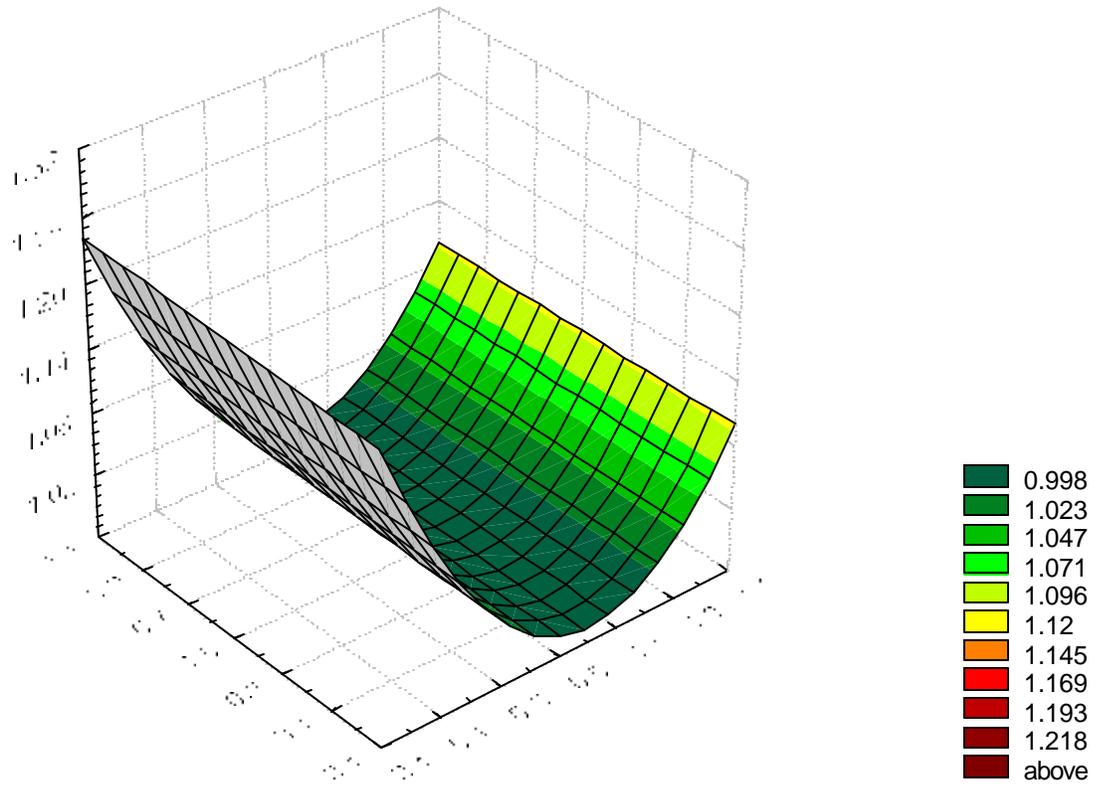


Figura 3-6: Datos generados mediante el modelo Gumbel(0.8). Se muestra el cociente $MSE(\hat{S}_{pc}) / MSE(\hat{S}_B)$, $n=20$.

Bibliografía

- [1] Breslow, N. and Crowley, J. (1974) A large sample study of the life table and product limit estimators under random censorship. *The Annals of Statistics* 2, No 3, 437-453.
- [2] Billingsley, P. (1968) *Convergence of Probability Measures*, Wiley, New York.
- [3] Birnbaum, Z. W. (1948). On random variables with comparable peakedness. *The Annals of Mathematical Statistics* 34, 1593-1601.
- [4] Bickel, P.J., & Lehmann, E. L. (1976). Descriptive statistics for nonparametric models. III: Dispersion. *The Annals of Statistics* 4,1139-1158.
- [5] Brurke, M. D. (1984) An almost-sure approximation of a multivariate product-limit estimator under random censorship. Unpublished.
- [6] Campbell, G. and Földes, A. (1982) Large-sample properties of nonparametric bivariate estimators with censored data. In *Nonparametric Statistical Inference* (B. V. Gnedenko, M. L. Puri and I. Vincze, eds.) 103-122, North-Holland, Amsterdam.
- [7] Campbell, G. (1982) Asymptotic properties of several nonparametric multivariate distribution function estimators under random censorship. In *Survival Analysis* (J. Crowley and R. A. Johnson, eds.) 243-256. IMS, Hayward, Calif.
- [8] Chung, K. L. (1974) *A course in Probability Theory*, Academic Press, New York.
- [9] Doksum, K. A. (1969) Starshaped transformations and the power of rank tests. *The Annals of Mathematical Statistics* 40, 1167-1176.

- [10] Dabrowska, D. M. (1988). Kaplan-Meier Estimate on the Plane. *The Annals of Statistics* 16, 1475-1489.
- [11] Dabrowska, D. M. (1989). Kaplan-Meier Estimate on the plane: Weak Convergence, LIL, and the Bootstrap. *Journal of Multivariate Analysis*, 29, 308-325.
- [12] Dykstra (1982) Maximum Likelihood Estimation of the Survival Functions of Stochastically Ordered Random Variables. *Journal of the American Statistical Association*, 77, No. 379, 621-628.
- [13] Dykstra, R. L., Feltz, C. J. (1989) Nonparametric Maximum Likelihood Estimation of Survival Functions with a General Stochastic Ordering and its Dual. *Biometrika*, 76, 2, 331-341.
- [14] El Barmi, H., Rojo, J. (1997) Likelihood Ratio Test for Peakedness in Multinomial Populations, *Journal of Nonparametric Statistics*, 7, 221-237.
- [15] Fraser, D. A. S. (1957) *Nonparametric Methods in Statistics*. Wiley, New York.
- [16] Gill, R. D., and Johansen, S. (1987). Product integrals and counting processes. Technical Report MS-R8707, Mathematisch Centrum. Amsterdam.
- [17] Gill, R. D. (1992a) Multivariate Survival Analysis. *Theory of Probability and Applications*, 37, No 1, 18-31.
- [18] Gill, R. D. (1992b) Multivariate Survival Analysis, Part 2, Technique. *Theory of Probability and Applications*, 37, No 2, 284-301.
- [19] Horváth, L. (1983) The rate of strong uniform consistency for the product-limit estimator. *Journal of Multivariate Analysis*, 13, 202-209.
- [20] Horváth, L. and Yandell, B. S. (1986) Bootstrapped multidimensional product limit process. Technical Report, Dept. Statistics, Univ. Wisconsin, Madison.
- [21] Johnson, M. (1987). *Multivariate Statistical Simulation*. New York: John Wiley & Sons.
- [22] Langberg, N. A., Shaked, M. (1982) On the identifiability of multivariate life distribution functions. *Annals of Probability* 10, 773-779.

- [23] Lehmann, E. L., Rojo, J. (1992) Invariant Directional Orderings. *The Annals of Mathematical Statistics*, 20, No. 4, 2100-2110.
- [24] Lo, S. H. and Wang, J. L. (1986) Iid representations for the bivariate product limit estimators and the bootstrap versions. Technical Report, Div. Statistics, Univ. California, Davis.
- [25] Muller, A; Stoyan, Dietrich (2002). *Comparison Methods for Stochastic Models and Risks*. West Sussex: John Wiley & Sons.
- [26] Oja, H. (1981) On Location, Scale, Skewness and Kurtosis of Univariate Distributions. *Scandinavian Journal of Statistics* 8, 154-168.
- [27] Oakes, D. (2001) Biometrika Centenary: Survival Analysis. *Biometrika*, 88, 1, 99-142.
- [28] Pons, O. (1986) A tests of independence between two censored survival times. *Scandinavian Journal of Statistics* 13, 173-185.
- [29] Pruitt, R. C. (1991) On Negative Mass Assigned y the Bivariate Kaplan-Meier Estimator. *The Annals of Statistics*, 19, No 1, 443-453.
- [30] Pruit, R. C. (1993) Identifiability of Bivariate Survival Curves from Censored Data. *Journal of the American Statistical Association*, Vol. 88, No. 422, 573-579.
- [31] Rojo, J. (1995) On the Weak Convergence of certain Estimators of Stochastically ordered Survival Functions. *Nonparametric Statistics* 4, 349-363.
- [32] Rojo, J; He, G. Z. (1991). On the Estimation of Stochastically Ordered Survival Functions, *Journal of Statistical Computation and Simulation*, Vol. 55, 1-51.
- [33] Rojo, J. and Ma, Z. (1996). On the Estimation of Stochastically Ordered Survival Functions. *Journal of Statistical Computation and Simulation*, Vol 55, 1-21.
- [34] Rojo, J., El Barmi, H. (2003) Estimation of Distribution Functions Under Second Order Stochastic Dominance. *Statistica Sinica* 13, 3, 903-926.
- [35] Rojo, J., Batun-Cutz, J., Durazo-Arvizu, R. (2005) . Inference Under Peakedness Restrictions. En revisión.

- [36] Schuster, E., Narvarte, J. A. (1973a) A New Nonparametric Estimator of the Center of a Symmetric Distribution. *The Annals of Statistics*, 1, No 6, 1096-1104.
- [37] Schuster, E. (1973b) On the Goodness-of-Fit Problem for Continuous Symmetric Distributions. *Journal of the American Statistical Association*, Vol. 68, No. 343, 713-715.
- [38] Schuster, E. (1974) Corrigenda: On the Goodness-of-Fit Problem for Continuous Symmetric Distributions. *Journal of the American Statistical Association*, Vol. 69, No. 345, 288.
- [39] Schuster, E. (1975). Estimating the Distribution Function of a Symmetric Distribution. *Biometrika*, 62, 3, 631-635.
- [40] Stute, W. and Wang, J. L. (1993) The strong law under random censorship. *The Annals of Statistics* 21, No 3, 1591-1607..
- [41] Van der Laan, M. J. (1996) Efficient estimator in the bivariate censoring model and repairing NPMLE. *The Annals of Statistics*, 24, 596-627.