UNIVERSIDAD DE LA LAGUNA

DEPARTAMENTO DE ESTADÍSTICA, INVESTIGACIÓN OPERATIVA Y COMPUTACIÓN

AVANCES EN LA ESTIMACIÓN DE LA SUPERVIVENCIA CON DATOS INCOMPLETOS Y DOBLEMENTE CENSURADOS

Arturo Javier Fernández Rodríguez

Memoria para optar al grado de Doctor en Ciencias Matemáticas realizada bajo la dirección del Dr. D. Miguel Ángel González Sierra

La Laguna, 1998

MIGUEL A. GONZÁLEZ SIERRA, PROFESOR TITULAR DE UNIVERSIDAD DEL DEPARTAMENTO DE ESTADÍSTICA, INVESTIGACIÓN OPERATIVA Y COMPUTACIÓN DE LA UNIVERSIDAD DE LA LAGUNA

CERTIFICO: Que la presente memoria, titulada **Avances en la Estimación de la Supervivencia con Datos Incompletos y Doblemente Censurados**, ha sido realizada bajo mi dirección por el Licenciado D. Arturo Javier Fernández Rodríguez, y constituye su Tesis para optar al grado de Doctor en Matemáticas.

Y para que conste, en cumplimiento de la legislación vigente y a los efectos que haya lugar, firmo el presente en

La Laguna, a veinte de marzo de 1998

Fdo: Miguel Ángel González Sierra

A mis padres y hermanos

por su apoyo y estímulo

Agradecimientos

El presente trabajo ha sido realizado bajo la dirección del Profesor Dr. D. Miguel Ángel González Sierra, a quien quiero expresar mi más sincero agradecimiento por todo su apoyo y dedicación.

Quiero hacer constar un especial agradecimiento a los ex compañeros y amigos José Ignacio Bravo e Íñigo de Fuentes por sus valiosas sugerencias y su inestimable colaboración.

A todas los mienbros del Departamento de Estadística, Investigación Operativa y Computación de La Universidad de La Laguna que me han ayudado, de una u otra forma, a lo largo de los años en que hemos compartido despacho, trabajo, tertulias, ... De forma expresa quiero agradecer a Natividad, Mercedes, Dionisio, Carlos y Marcos, toda su ayuda.

Arturo Javier Fernández Rodríguez

Índice

Prólogo

La necesidad de cierto rigor científico ha ocasionado una creciente aplicación de los métodos estadísticos en todas las disciplinas experimentales que, como tales, están sujetas a razonamientos de tipo inductivo que van de lo particular a lo general, es decir, que pretenden extender la conclusiones obtenidas en una parte, al todo. El único método científico adecuado para dar validez a tales extensiones es el método estadístico, pues precisamente ese es el motivo de su existencia.

Con el término *análisis de supervivencia (o fiabilidad)* se denota una amplia variedad de técnicas estadísticas orientadas al estudio de variables aleatorias no negativas, que representan el tiempo que transcurre desde un instante inicial hasta la ocurrencia de algún suceso de interés (a menudo, el fallo). Son ejemplos típicos, el tiempo de vida de un paciente con cáncer o el que tarda en fallar un sistema informático. En general, la variable en estudio se denomina génericamente *tiempo de supervivencia, de vida o de fallo.* Sin embargo, mediante estas técnicas, también se podrían analizar variables de otra clase. Por ejemplo, el número de fotocopias realizadas por una máquina o el kilometraje de un automóvil hasta su primera avería no son variables temporales sino que miden la frecuencia de uso.

Desde sus orígenes, los primeros trabajos sobre tablas de mortalidad, hasta la actualidad, con aplicaciones en campos tan diversos como la biología, ingeniería, economía, física, demografía, criminología, y un largo etcétera, el campo de actuación de estas técnicas ha ido aumentando constantemente.

La variabilidad de los resultados que se obtienen y la existencia de cierta regularidad inherente a esa incertidumbre justifican la utilización de modelos probabilísticos, que proporcionan el soporte teórico esencial y las herramientas necesarias para un análisis profundo, en los problemas de supervivencia. Una de las características distintivas de estos problemas es la presencia de complicados patrones de censura que dificultan en gran medida el estudio. Asimismo, un aspecto inherente al análisis de supervivencia es que el mecanismo del fallo/censura puede modelizarse preferentemente utilizando técnicas extraídas de la probabilidad y estadística. La necesidad de obtener métodos de análisis que incorporen la censura es quizás la razón más importante para la aparición y desarrollo de modelos y procedimientos especiales para datos de supervivencia.

La estimación de la distribución del tiempo que transcurre hasta el fallo tiene con frecuencia un interés primordial en los estudios de supervivencia. En su resolución se han utilizado una gran variedad de puntos de vista de acuerdo con el modelo supuesto y el tipo de información disponible. En bastantes ocasiones, debido a los métodos de muestreo y a ciertos factores que escapan al control experimental, puede que el *tiempo de vida* no sea directamente observable. En la mayoría de esos casos sólo se sabe que es mayor o menor que un determinado *tiempo de censura por la derecha o por la izquierda*, respectivamente. El problema de la estimación de funciones de supervivencia a partir de datos no censurados o censurados por la derecha ha sido analizado por multitud de investigadores. A pesar de que en bastantes ocasiones aparecen observaciones censuradas por la izquierda, la estimación de la supervivencia con datos doblemente censurados no ha sido tratada con la debida atención en la literatura científica. No obstante, en los últimos años han aparecido un buen número de artículos que analizan modelos de doble censura, e incluso esquemas de censura más complejos.

Por diversas razones, a menudo, cierta información relevante para el análisis de la supervivencia no se conoce. Por ejemplo, cuando ha finalizado el periodo de garantía, el vendedor desconoce el kilometraje de los automóviles que no han sufrido aún una avería. Muchos investigadores han propuesto métodos para salvar, en lo posible, esa falta de información. En esta memoria se recomienda obtener (por ejemplo, mediante el servicio postal) un apropiado porcentaje de datos censurados por la derecha (esto es, kilometrajes de automóviles sin avería en el periodo de garantía) cuando no se cuenta con sus valores. Esto es importante ya que esas observaciones censuradas contienen información relevante y pueden ayudar a conseguir la precisión requerida del estimador.

En este trabajo ahondamos en la resolución del problema de la estimación de la supervivencia en doble censura con datos incompletos. Para ello, se generalizan algunos modelos propuestos en el caso de censura a la derecha y se introducen nuevas técnicas paramétricas (en el capítulo 2) y semiparamétricas (en los capítulos

3 y 4) en este campo. Al mismo tiempo, se evalúa el efecto de la proporción de datos desconocidos sobre la eficiencia de los estimadores propuestos. Se asume que el mecanismo de censura no depende de la supervivencia y que la variable de censura a la derecha no siempre se conoce, esto es, los datos censurados a la derecha son sólo parcialmente observables. Este tipo de datos incompletos ha sido analizado con amplitud en el caso de censura a la derecha, pero no cuando existe doble censura.

El primer capítulo es meramente introductorio y, además de varias reseñas históricas, presenta algunos procedimientos de estimación y la formulación básica de los modelos probabilísticos de supervivencia. Asimismo, se define el modelo de censura aleatoria doble y el de tasas de fallo proporcionales, y se hace hincapié en la frecuente presencia de información incompleta cuando se analizan datos de supervivencia.

En el segundo capítulo se presenta un procedimiento, basado en una función de pseudo-verosimilitud, para la estimación de parámetros de distribuciones de supervivencia cuando se analizan datos incompletos y doblemente censurados. El estimador propuesto se obtiene bajo los supuestos de un modelo paramétrico conocido, seleccionado mediante estudios preliminares, y la independencia entre los *tiempos de vida* y los *de censura*. Aun cuando el mecanismo de censura es simple y conocido, la distribución teórica de los estimadores de la supervivencia suele ser compleja. Este hecho conduce, en la mayoría de casos, a la utilización de métodos asintóticos. En particular, el estimador obtenido verifica tan buenas propiedades asintóticas como las de un estimador de máxima verosimilitud; entre otras, consistencia y normalidad asintótica. Un estudio acerca de la fiabilidad de cierto sistema de gestión de bases de datos se presenta como ilustración. Además, se realiza un estudio computacional diseñado para analizar, según el nivel de censura y el tamaño muestral, la eficiencia del estimador propuesto.

Desde el punto de vista no paramétrico, la función de supervivencia muestral no es apropiada para estimar la función de supervivencia teórica cuando existe información censurada o desconocida. Una de las posibles vías de solución es la de considerar un modelo paramétrico exclusivamente para aquellos datos que no se conocen con exactitud. De este modo, en el tercer capítulo se propone un procedimiento semiparamétrico para la estimación de la función de supervivencia de la variable de interés cuando se analizan datos incompletos en un modelo de doble censura aleatoria. El estimador obtenido trata las observaciones no censuradas de forma no paramétrica y utiliza modelos paramétricos para las que están censuradas tanto a la derecha como a la izquierda. Por tanto, si el nivel de censura no es elevado, mantiene la mayoría de las propiedades de distribución libre de los estimadores no paramétricos y, además, mejora la precisión de la estimación de la supervivencia en las colas. El método se ilustra mediante un ejemplo en el que se analiza un sistema de garantía de calidad que ofrece un establecimiento que vende fotocopiadoras. Asimismo, se estudian algunas propiedades asintóticas del estimador propuesto, como la consistencia y la convergencia débil a un proceso Gaussiano, y, mediante simulación, se analiza y compara su comportamiento.

A veces existe la posiblidad de obtener alguna información previa, basada en opiniones subjetivas o en experiencias pasadas, que puede mejorar la eficacia del proceso de estimación. Además de evaluar númericamente las consecuencias de las decisiones a través de una función de pérdida, cuyo valor esperado debe minimizarse, las técnicas Bayesianas proporcionan una metodología para combinar la información a priori con la que origina la propia muestra.

En el contexto Bayesiano no paramétrico, donde la propia función de supervivencia es el "parámetro" a estimar, no es necesario someterse a restricciones impuestas por consideraciones teóricas, pero cuenta con la dificultad de definir una medida de probabilidad sobre un espacio funcional que asigne probabilidad uno al conjunto de las funciones de supervivencia. En el caso de censura doble, las enormes complicaciones que surgen han ocasionado que no se haya obtenido ningún resultado relevante hasta la actualidad. Estas dificultades han propiciado la aparición de procedimientos Bayesianos alternativos, tales como el que se propone en el capítulo 4. Este último capítulo presenta un procedimiento Bayesiano semiparamétrico para la estimación de una curva de supervivencia cuando los datos son doblemente censurados e incompletos y la función de pérdida es cuadrática. El conocimiento a priori se expresa mediante un proceso de Dirichlet sobre el vector completo de variables observables, y se considera el modelo de tasas de fallo proporcionales. Además de estimar los parámetros del modelo y comprobar su normalidad asintótica, se construye un estimador semiparamétrico de la función de supervivencia y se demuestran algunas propiedades asintóticas, como la consistencia fuerte y la convergencia débil a un proceso Gaussiano. Asimismo, se estudia y compara su eficacia mediante simulación.

Capítulo 1

Introducción al análisis de supervivencia

1.1 Introducción

El análisis de supervivencia engloba una amplia variedad de técnicas estadísticas orientadas al estudio del comportamiento de variables aleatorias no negativas, especialmente aquéllas referentes a datos recogidos en tablas de vida. El tiempo que transcurre hasta la muerte o recaída de una paciente con cáncer, el que tarda en fallar una componente eléctrica o mecánica, o en germinar una semilla, o el de duración de una huelga son algunos de los muchos ejemplos que podrían citarse como variables de interés para este análisis. Aunque todos estos casos tienen en común el estudiar una variable de tipo temporal, el *tiempo de ocurrencia* del suceso de interés, se podría analizar mediante estas técnicas variables de otro tipo, por ejemplo, el número de fotocopias realizadas por una empresa.

Los orígenes del análisis de supervivencia pueden atribuirse a los primeros trabajos sobre tablas de mortalidad de hace ya varios siglos, no obstante, su era moderna comenzó hace aproximadamente medio siglo a raíz de las aplicaciones de la teoría de la fiabilidad en ingeniería, especialmente las de tipo militar, puestas en práctica a partir de la Segunda Guerra Mundial. Hoy en día, estas técnicas no sólo se emplean con gran profusión en las ciencias biomédicas y actuariales, sino que también se han introducido en muchas otras áreas del conocimiento como son la zoología, ecología, botánica, biometría forestal, procesos industriales, fiabilidad de sistemas, física, etc.

Entre fiablidad y supervivencia existe una importante conexión: ambas tratan de modelizar, a través de conceptos probabilísticos, el *tiempo de vida* de un mecanismo, un sistema, un animal o un ser humano. Por supuesto, también existen diferencias. Un problema de gran importancia en análisis de fiabilidad es la optimización de la eficiencia de un sistema de múltiples componentes. Dicha optimización depende de la relación entre el número de componentes del sistema y su localización respectiva dentro del mismo. En análisis de supervivencia no existe ningún problema de este tipo, ya que el sistema en esta ocasión está vivo y funcionando, lo que hace que la redisposición de las componentes, que en este caso serían los órganos, no sea una alternativa viable. Por otro lado, un problema de gran relevancia en este último campo es la comparación de tiempos de supervivencia de pacientes o animales experimentales que están sometidos a dos o más tratamientos diferentes. En teoría de fiabilidad también surgen problemas de esa índole, por ejemplo, cuando se precisa comparar las distribuciones del *tiempo de vida* antes y después de ciertos cambios en el sistema; sin embargo, no son tan importantes como los problemas de comparación en análisis de supervivencia.

1.2 Tipos de datos de supervivencia

Supóngase que se somete a estudio una muestra aleatoria de N individuos sobre la que se pretende analizar el comportamiento de cierta variable de interés. Si la observación de cada unidad en estudio es continua, se obtendrá la información exacta del tiempo transcurrido, T, hasta la ocurrencia del suceso de interés. En esta ocasión, la variable observada, T, es una variable *continua*, es decir, puede suponerse que toma cualquier valor dentro de cierto intervalo de tiempo. Cuando la observación de las unidades en estudio se realiza en instantes de tiempo discretos, la información que se recoge es el número de ocurrencias del suceso de interés dentro de cada uno de los intervalos definidos por los tiempos consecutivos de observación. En este caso se consideran variables aleatorias *discretas* que representan el número de fallos en cada uno de los intervalos. Bajo esta clasificación, existen en principio dos tipos de datos: *agrupados* (caso discreto) y *no agrupados* (caso continuo).

El análisis de los datos depende del tipo de observaciones de que se disponga. A menudo, por razones de tiempo y dinero, los experimentos deben finalizar antes de que el suceso de interés haya ocurrido para todos los individuos en estudio. En tales casos, la información acerca del tiempo de ocurrencia de dicho suceso es completa (supuesto que la observación se realiza de forma continua) sólo para una parte de la muestra. De los individuos para los que no ha ocurrido aún el suceso, solamente se tiene una información parcial. Este tipo de dato suele denominarse *censurado por el tiempo.* Si todas las unidades empiezan a funcionar al mismo tiempo, o todos los individuos "nacen" a la vez, el caso es de *censura simple.* La *censura simple por el tiempo* también se denomina a menudo *censura de tipo I.*

Algunos experimentos finalizan en el instante en que se produce el r-ésimo fallo, donde r es un número entero predeterminado menor que el tamaño muestral. En estos experimentos se dice que los datos están *censurados por el número de fallos.* La *censura simple por el número de fallos* se denomina también *censura de tipo II.*

Cuando las unidades en estudio empiezan a funcionar en distintos instantes o, lo que es lo mismo, los individuos de la muestra nacen en diferentes momentos dentro de cierto intervalo $[0, t^*]$, y el experimento finaliza en el instante t^*, se dice que el caso es de *censura múltiple.*

Finalmente, debe distinguirse entre censura a la izquierda y censura a la derecha. Si algunos individuos de la muestra empezaron a funcionar, o nacieron, antes del inicio del estudio, existirá *censura a la izquierda.* El otro tipo de información censurada se origina cuando, finalizado el experimento, hay individuos en la muestra para los que no ha ocurrido todavía el suceso de interés. Esta clase de censura se denomina *censura a la derecha.* Cuando algunos datos están censurados a la derecha y otros a la izquierda, se dice que existe *doble censura*, mientras que en el caso de *censura por intervalos* existe algún dato que está censurado tanto a la derecha como a la izquierda.

1.3 Técnicas en análisis de supervivencia

Como se mencionó en el apartado anterior los datos de supervivencia pueden ser de muy diversos tipos, por lo que no es posible establecer unas reglas de procedimiento universales.

En un principio, la mayor parte de la investigación estadística, desarrollada para obtener aplicaciones en ingeniería, se centró en problemas paramétricos. En las últimas décadas, el número de ensayos clínicos sometidos a investigación médica ha experimentado un gran crecimiento, lo que ha permitido un mayor acercamiento a las técnicas no paramétricas.

Los primeros estudios realizados con datos censurados utilizaron modelos paramétricos, quizá por su simplicidad y atractivo matemático más que por su idoneidad a la hora de afrontar los problemas prácticos. Estos modelos paramétricos se basan en la suposición de que la determinación exacta de la distribución de la variable de interés, T, depende sólo y exclusivamente de un vector de parámetros desconocidos θ. Para aplicar correctamente estos métodos es imprescindible poseer gran cantidad de información con objeto de seleccionar un conjunto reducido de distribuciones, expresable paramétricamente, dentro del cual se encuentre con certeza la distribución de T.

Los procedimientos gráficos pueden utilizarse para efectuar una comprobación informal de la validez del modelo propuesto e incluso para la obtención de estimaciones aproximadas de los parámetros desconocidos. La idea básica es construir gráficas que serían aproximadamente lineales si el modelo considerado se ajustase a los datos obtenidos. De este modo, puede apreciarse a simple vista las desviaciones de la linealidad, las cuales proporcionarían evidencias de la incorrección del modelo supuesto (véase, por ejemplo, Nelson, 1982).

Si se conoce la familia paramétrica a la que pertenece la distribución de supervivencia subyacente, se prefieren, en general, los métodos paramétricos de máxima verosimilitud (véase Mann et al., 1974; Bain, 1978; Lawless, 1982; Nelson, 1982). Ya se sabe que, bajo ciertas condiciones de regularidad, el estimador máximo verosímil de θ es consistente y sigue una distribución asintóticamente normal. Entre las distribuciones utilizadas con mayor frecuencia se encuentran la exponencial (Davis, 1952; Epstein y Tsao, 1953; Epstein y Sobel, 1953, 1954; Bartholomew, 1957, 1963; Epstein, 1958; Zelen y Dannemiller, 1961), la de Weibull (Weibull, 1951; Lieblein y Zelen, 1956; Kao, 1959; Cohen, 1965; Harter y Moore, 1965), la normal y log-normal (Dixon, 1960; Feinleib, 1960; Cohen, 1961; Hill, 1963), la gamma (Gupta y Groll, 1961; Wilk et al., 1962a, 1962b; Stacy y Mihram, 1965; Harter, 1967), la Gaussiana inversa (Chhikara y Folks, 1977; Padgett, 1979), la log-logística (Kalbfleisch y Prentice, 1980, sec. 2.2.6), la de Pareto (Proschan, 1963; Davis y Feldstein, 1979), y la exponencial potencial (Smith y Bain, 1975). Debido a la frecuente dificultad para distinguir entre las distribuciones candidatas (por ejemplo, véase Bain y Englehardt, 1980), el investigador puede muy bien elegir una distribución incorrecta. Si se elige un modelo equivocado, entonces se sabe que el método de máxima verosimilitud conduce a la obtención de resultados asintóticamente sesgados y será, por tanto, inapropiado. Este problema en el sesgo de las estimaciones por un error en la especificación del modelo es quizás el más generalizado en la utilización de este método.

Cuando no se conozca con certeza el modelo paramétrico subyacente, los métodos no paramétricos se presentan como la alternativa más razonable ya que no requieren información sobre la distribución de la variable T.

El desarrollo teórico de las técnicas de estimación no paramétricas con datos censurados puede considerarse bastante reciente. En 1958, Kaplan y Meier publican un compendio de los escasos resultados que se habían obtenido hasta ese momento para observaciones censuradas a la derecha y añaden un estudio de las propiedades de un nuevo estimador de la supervivencia, el estimador producto-límite, al que se conocerá más tarde con el nombre de sus autores. En los años siguientes, Breslow y Crowley (1974), Peterson (1977) y Johansen (1978) establecen en sendos trabajos algunas de las buenas propiedades del estimador de Kaplan-Meier. Estas buenas propiedades, junto al ser de máxima verosimilitud no paramétrica, han hecho que este estimador se haya utilizado casi con exclusividad en los problemas no paramétricos con datos censurados a la derecha. En este área, Miller (1981) proporciona muchas referencias, además de un análisis conciso de los principales métodos no paramétricos de estimación.

Ambos planteamientos tienen sus pros y sus contras. Mientras que el estimador de Kaplan-Meier tiene muchas de las propiedades deseables para grandes muestras, sus propiedades para muestras pequeñas dejan mucho que desear. En particular, es sesgado para muestras finitas y la magnitud del sesgo es inversamente proporcional al tamaño muestral (véase Gross y Clark, 1975). Miller (1983) demuestra que la eficiencia asíntotica del estimador producto-límite de Kaplan-Meier es inferior a la del estimador paramétrico de máxima verosimilitud si el nivel de censura es alto o la supervivencia está proxima a cero.

Cuando los datos están doblemente censurados en los problemas de estimación no paramétrica, se utiliza con frecuencia el estimador autoconsistente de la función de supervivencia, propuesto primeramente por Turnbull (1974), el cual verifica propiedades similares al de Kaplan-Meier (véase Gu y Zhang, 1993; Mykland y Ren, 1996). En 1976, el propio Turnbull generaliza su estimador para datos arbitrariamente agrupados, censurados y truncados.

El hecho de que la información inicial que se disponga no aconseje la utilización de modelos paramétricos no significa que tenga que utilizarse sólo y exclusivamente la información muestral. Existen situaciones en las que es posible obtener información a priori y por razones prácticas debería tenerse en cuenta. La metodología Bayesiana proporciona un procedimiento para combinar las opiniones a priori con los datos observados. Además, a menudo, pueden conocerse las consecuencias de las decisiones, y es posible cuantificar númericamente la pérdida que ocasiona cada decisión en cualquiera de las situaciones mediante una función de pérdida.

En el caso Bayesiano parámetrico se considera que existe una distribución a priori sobre el espacio paramétrico Θ, al que pertenece θ, y una función de pérdida del decisor, cuyo valor esperado se intenta minimizar. En este campo, Tsokos y Shimi (1977) y Martz y Waller (1982) proporcionan un extenso compendio de técnicas Bayesianas aplicables en fiabilidad. Asimismo, ambos textos analizan algunos métodos Bayesianos empíricos.

Por otra parte, los métodos Bayesianos no paramétricos, donde la propia función de distribución de T es el "parámetro" que se desea estimar, cuentan con la dificultad de definir una distribución a priori. Este problema ha sido abordado por autores como Ferguson (1973), Doksum (1974), Antoniak (1974), Susarla y Van Ryzin (1976, 1978), Kalbfleisch (1978), Ferguson y Phadia (1979), Dykstra y Laud (1981), Tiwari y Zalkikar (1993) y Fernández et al. (1997).

También tienen especial relevancia las técnicas semiparamétricas o parcialmente paramétricas, en las cuales se supone que existe cierta información sobre las variables del modelo que permite estimaciones más eficientes de la supervivencia. Los modelos estadísticos semiparamétricos incluyen como desconocidos tanto parámetros como funciones. En general, el interés se centra en la estimación de los parámetros del modelo especificado, considerando que las funciones desconocidas constituyen un parámetro de dimensión infinita. Una motivación práctica para la utilización de modelos semiparamétricos es la de evitar restricciones sobre aspectos secundarios del problema, mientras se preserva una formulación estricta de las características de interés primordial. Mediante la imposición de restricciones paramétricas sobre los modelos no paramétricos pueden obtenerse modelos semiparamétricos (por ejemplo, los modelos de los capítulos 3 y 4). El modelo de riesgos proporcionales de Cox (1972) es el modelo semiparamétrico que más se utiliza en análisis de supervivencia.

En los modelos completamente paramétricos, sujetos a ciertas condiciones de regularidad, los estimadores de máxima verosimilitud son consistentes, y asintóticamente normales y eficientes. Para los modelos semiparamétricos no se obtienen esos resultados generales. Tradicionalmente, los procedimientos de inferencia para los modelos semiparamétricos se han obtenido ad hoc. No obstante, en las últimas décadas, han aparecido algunos artículos sobre el desarrollo de una teoría asintótica para problemas semiparamétricos suficientemente regulares que intenta imitar el desarrollo tradicional de ésta para modelos completamente paramétricos (por ejemplo, véase Gill, 1987; Bickel et al., 1993). La idea fundamental es la aproximación local de un modelo semiparamétrico por otro puramente paramétrico "menos favorable".

Entre la variedad de artículos que utilizan técnicas semiparamétricas con datos censurados se encuentran: Miller (1976), Buckley y James (1979), Finkelstein y Wolfe (1985), Finkelstein (1986), Oakes (1986), Maguluri (1993), Bravo y Esteban (1993b), Bravo et al. (1995a, 1995c) y Genest et al. (1995).

1.4 Modelo probabilístico

La variabilidad de los resultados que se obtienen al medir el tiempo transcurrido entre el instante inicial y el momento en que ocurre el suceso de interés, junto con cierta regularidad inherente a esa incertidumbre, justifican la utilización de modelos probabilísticos en los problemas de supervivencia. No obstante, hemos de decir que, en todas las aplicaciones de los métodos estadísticos basados en probabilidades, los modelos y las suposiciones en que se basan no se satisfacen, por lo general, de forma completa. Esto ocurre igualmente en el campo de la fiabilidad y supervivencia. La teoría de la probabilidad proporciona el soporte teórico esencial y las herramientas necesarias para estudiar el *tiempo de vida*, pero existen muchas y muy variadas posibilidades de que la realidad se desvie de los modelos teóricos, por lo que es esencial confrontar directamente los datos con los modelos mediante un análisis descriptivo previo.

El modelo probabilístico básico está constituido por una variable aleatoria no negativa, T, (*tiempo de vida o de fallo*) que mide el tiempo transcurrido desde un instante inicial ($t_0 = 0$) hasta que ocurre el suceso de interés en los distintos elementos de una población homogénea. Asociados a esta variable aleatoria hay una serie de conceptos imprescindibles para el desarrollo posterior que se presentan a continuación.

Se denomina *función de supervivencia o fiabilidad* de la variable T a la función S_T definida por:

$$S_T(t) = \Pr(T > t)\,, \quad t \geq 0.$$

Obviamente, $S_T(t)$ es la probabilidad de que el suceso de interés ocurra con posterioridad al instante t y se verifica que $S_T(t) = 1 - F_T(t)$, donde F_T es la función de distribución de T. Por tanto, la función de supervivencia es continua por la derecha, monótona no creciente y verifica que $S_T(0^-) = \Pr(T \geq 0) = 1$ y $\lim_{t\to\infty} S_T(t) = 0$.

Si la variable aleatoria T es absolutamente continua, y f_T es su función de densidad, entonces

$$S_T(t) = \int_t^\infty f_T(s)ds\,, \quad t \geq 0 \quad \text{y} \quad E\left[T\right] = \int_0^\infty S_T(t)dt.$$

Además, $f_T(s) = -\left(d/dt\right) S_T(s)$ en todo punto de continuidad de f_T.

Otra función utilizada con frecuencia, y de especial importancia, es la *tasa de fallo* o *función de riesgo*, cuya expresión matemática para variables absolutamente continuas es

$$h_T(t) = \lim_{\Delta\to 0+} \frac{\Pr(t < T \leq t + \Delta \mid T > t)}{\Delta}, \quad t \geq 0.$$

Intuitivamente, $h_T(t)$ representa la propensión a la ocurrencia instantánea del suceso de interés en t, sabiendo que hasta dicho instante no había ocurrido. Históricamente, en los estudios de epidemiología, se la denominó *fuerza de la mortalidad*. Es sencillo comprobar que esta función es no negativa y que

$$h_T\left(t\right) = f_T(t)/S_T(t)$$

en su campo de definición. Además, para todo $t \geq 0$, se cumple que

$$S_T(t) = \exp\left(-\int_0^t h_T(s)ds\right) \quad \text{y} \quad f_T(t) = h_T\left(t\right)\exp\left(-\int_0^t h_T(s)ds\right).$$

En ocasiones es útil definir la *tasa de fallo acumulativa* o *función de riesgo acumulativo*:

$$H_T(t) = \int_0^t h_T(s)ds = -\log S_T(t).$$

Los conceptos anteriores pueden extenderse con facilidad al caso en el cual la variable T no es absolutamente continua.

En situaciones homogéneas se considera que los tiempos de vida de todos los individuos son variables aleatorias independientes e idénticamente distribuidas. En otras ocasiones se supone que, además de los tiempos de vida de cada individuo, también se han observado sobre cada uno de ellos cierto número de covariables que, de alguna forma, afectan a la distribución del *tiempo de vida*. En el intento de incorporar este tipo de información en los modelos de supervivencia ha sido fructífero trabajar en términos de la función tasa de fallo. Por ejemplo, el modelo de riesgos proporcionales, introducido por Cox (1972), asume que el efecto de las covariables sobre la tasa de fallo es multiplicativo. Asimismo, con frecuencia, no existe una única variable de supervivencia. Campbell (1981), Korwar (1987), Oakes (1994) y otros investigadores consideran distribuciones multidimensionales.

En el campo del análisis de fiabilidad, a menudo, la variable T representa el tiempo que transcurre hasta que falla un determinado sistema (Barlow y Proschan (1965, 1975) y Gerstbakh (1989) analizan, desde la óptica estadística, la teoría de fiabilidad). Además, con frecuencia, estos sistemas constan de varios subsistemas, y cada uno de éstos de dos o más componentes que pueden reponerse y/o repararse. (véase, por ejemplo, Ascher y Feingold, 1984; Klassen y van Peppen, 1989)

1.5 Modelo de censura aleatoria doble

En muchas ocasiones sólo se conoce con exactitud el tiempo transcurrido, T, hasta la ocurrencia del suceso de interés de una parte de los N individuos de la muestra que se desea analizar, mientras que del resto solamente se tiene una información parcial. Cuando sólo se sabe que el *tiempo de vida* de un individuo, T_i, es menor que un cierto valor conocido, X_i, se dice que T_i está censurado por la izquierda en X_i. Si se conoce únicamente que $T_i > Y_i$, entonces T_i está censurado por la derecha en Y_i. Cuando existen algunos datos censurados por la derecha y otros por la izquierda, se dice que el caso es de censura doble. Este tipo de censura es el que motiva todo el trabajo de investigación que aquí se presenta.

En el modelo de doble censura aleatoria se considera que existen tres variables aleatorias definidas sobre cada individuo: T, *tiempo de ocurrencia del suceso de interés*, Y, *tiempo de censura por la derecha* y X, *tiempo de censura por la izquierda.* Además, se supone que $\Pr(X \leq Y) = 1$ y que T es independiente del intervalo aleatorio de observación $[X, Y]$.

En este caso, dada una muestra aleatoria simple (T_i, X_i, Y_i), $i = 1, ..., N$, la información disponible vendrá dada por los pares (Z_i, δ_i), $i = 1, ..., N$, donde

$$Z_i = \max\left\{\min\left(T_i, Y_i\right), X_i\right\},$$

y

$$\delta_i = \begin{cases} 0 \quad si \quad X_i \leq T_i \leq Y_i & (no\ censura), \\ 1 \quad si \qquad\quad T_i > Y_i & (censura\ por\ la\ derecha), \\ 2 \quad si \qquad\quad T_i < X_i & (censura\ por\ la\ izquierda). \end{cases}$$

Gehan (1965), Mantel (1967), Peto (1973) , Leiderman et al. (1973), Turnbull (1974), Turnbull y Weiss (1978), Morales et al. (1991), Tang et al. (1995) y otros autores proporcionan ejemplos de muestras doblemente censuradas que aparecen en situaciones prácticas.

En los estudios sobre supervivencia o fiabilidad, la estimación de la función de distribución del *tiempo de vida* es, con frecuencia, de especial interés. En esos problemas, a menudo, aparecen observaciones censuradas y, por tanto, la función de distribución empírica no es apropiada para la estimación no paramétrica de la función de distribución teórica del *tiempo de vida o de supervivencia.* En el caso de censura a la derecha se acepta generalmente el estimador producto-límite de Kaplan y Meier (1958) como sustituto de la función de distribución empírica, ya que es el estimador de máxima verosimilitud no paramétrico (EMVNP) (Cox y Oakes, 1984, pag. 48) y posee las propiedades de autoconsistencia (Efron, 1967), normalidad asintótica (Breslow y Crowley, 1974) y eficiencia asintótica (Wellner, 1982). Es natural que, en otros casos distintos a éste, se busquen estimadores que posean propiedades similares. Para una discusión detallada, véase Tsai y Crowley (1985) y Gill (1989).

En el caso más general de doble censura, Turnbull (1974) propone un algoritmo autoconsistente para hallar el EMVNP de la distribución subyacente. Chang y Yang (1987) demuestran la identificabilidad y la consistencia uniforme fuerte del estimador autoconsistente, bajo ciertas condiciones aceptables sobre la censura, para distribuciones continuas con soporte $(0, \infty)$, y Chang (1990) comprueba la normalidad asintótica en intervalos compactos supuesta una condición adicional bastante restrictiva sobre la censura. Gu y Zhang (1993) generalizan los resultados de consistencia a distribuciones no continuas con soporte arbitrario, establecen la normalidad asintótica sobre el soporte completo del *tiempo de vida* bajo condiciones más generales sobre las distribuciones de las variables de censura, y demuestran que el estimador de Turnbull es asintóticamente eficiente.

Desde otro punto de vista, Samuelsen (1989) construye estimadores no paramétricos consistentes de la función de supervivencia y de la tasa de fallo acumulativa que convergen débilmente (propiamente normalizados) a procesos Gaussianos, considerando que se conocen todos los tiempos de censura a la izquierda.

En el contexto paramétrico, Fernández et al. (1996) estudian el modelo de riesgos proporcionales de Cox (1972), a partir de datos de supervivencia doblemente censurados y covariables, mediante la modelización paramétrica de la tasa de fallo base en términos de distribuciones definidas a trozos (o con parámetros variables). Considerando que la distribución base de la supervivencia es, a trozos, exponencial, de Weibull o de valor extremo generalizada, obtienen los correspondientes sistemas de ecuaciones de verosimilitud, contrastan la influencia de las covariables en la supervivencia y proponen algunos métodos gráficos para la evaluación de los modelos propuestos.

Muchos trabajos recientes en el área de la supervivencia con datos censurados por la derecha han sido motivados por la teoría de los procesos contadores (véase Aalen, 1978; Andersen y Gill, 1982). Esta estructura ha proporcionado demostraciones elegantes y simples de muchos problemas de este campo. Fleming y Harrington (1991) y Andersen et al. (1993) presentan un compendio de las aplicaciones de las martingalas y los procesos contadores al análisis de supervivencia con datos censurados por la derecha. Sin embargo, la estructura de los procesos contadores no puede adaptarse con facilidad cuando los datos están doblemente censurados ya que es difícil definir de forma adecuada la secuencia creciente de σ-álgebras. Por este motivo, el desarrollo de una metodología apropiada para datos de supervivencia arbitrariamente censurados requiere una punto de vista esencialmente diferente, y constituye una importante área de investigación.

Desde luego, la estimación del tiempo de vida no es el único interés en el campo del análisis de supervivencia con datos doblemente censurados. Muchos artículos analizan los más diversos temas cuando existe censura doble. Por ejemplo, el contraste de homogeneidad de dos muestras (Gehan, 1965; Mantel, 1967), el contraste de bondad de ajuste (Turnbull y Weiss, 1978; Ren, 1995), la estimación de una función de supervivencia bivariada (Korwar, 1987), y el modelo lineal de regresión (Zhang y Li, 1996).

Un campo en el que aparecen con frecuencia datos doblemente censurados es en los experimentos de supervivencia/sacrificio en animales, especialmente los que se diseñan para determinar si ciertos agentes aceleran la aparición de un tumor. En tales estudios, a cada animal se le administra una dosis predeterminada de la sustancia supuestamente cancerígena y se examina, a su muerte o sacrificio, la presencia o no de un tumor. En la mayoría de las ocasiones, la aparición del tumor no es observable directamente. Tan sólo puede constatarse el tiempo de su muerte o sacrificio. En los casos donde el tumor es irreversible y no letal, los tiempos observados de muerte o sacrificio proporcionan observaciones censuradas a la izquierda (si el tumor está presente) o censuradas a la derecha (si el tumor está ausente) de los tiempos de aparición del tumor. Algunos trabajos que analizan datos sobre tumores no letales son, entre otros, Hoel y Walburg (1972), Peto et al. (1980), Dinse y Lagakos (1983) y Finkelstein (1986).

Huang (1996) estudia el estimador de máxima verosimilitud para el modelo de riesgos proporcionales de Cox (1972), considerando datos que están censurados a la derecha o a la izquierda, y aplica sus resultados a un estudio sobre tumorogénesis. Este tipo de datos también aparecen en estudios demográficos; véase, por ejemplo, Diamond et al. (1986) y Diamond and MacDonald (1991). Tang et al. (1995) proponen varios tests no paramétricos para contrastar la homogeneidad de dos muestras de esta clase y Wang y Gardiner (1996) construyen una clase de estimadores autoconsistentes de la función de supervivencia. Rabinowitz et al. (1995) tratan el problema de la estimación en modelos de regresión con datos censurados por intervalos. Además, en los estudios sobre del virus de inmunodeficiencia humana (VIH) y el síndrome de inmunodeficiencia adquirida (SIDA) aparecen esquemas de censura similares; véase, por ejemplo, Shiboski y Jewell (1992) y Jewell et al. (1994).

1.6 Modelo de tasas de fallo proporcionales

Un caso particular del modelo de censura aleatoria a la derecha, que ha sido utilizado por varios autores, es el modelo de tasas de fallo proporcionales o modelo de Koziol-Green.

En el caso de censura aleatoria a la derecha se supone que existe un número real positivo desconocido, β, tal que las funciones de supervivencia de las variables inepedientes T e Y, S_T y S_Y respectivamente, verifican la siguiente relación:

$$S_Y(t) = S_T^{\beta}(t)\,, \ \text{ para todo } t \geq 0.$$

Si T e Y son absolutamente continuas con soporte común $[0, \infty)$, entonces se verifica la siguiente relación entre las tasas de fallo de Y y T

$$h_Y(t) = \beta h_T(t)\,, \ \text{ para todo } t \geq 0,$$

expresión que justifica el nombre del modelo.

Algunos investigadores, entre otros, que estudian este modelo son: Koziol y Green (1976) proponen un test de bondad de ajuste de S_T del tipo de Cramér-von Mises; Csörgo y Horváth (1981) analizan bandas de confianza para la estimación de la supervivencia ; Ebrahimi (1985) propone un estimador de S_T y demuestra su consistencia fuerte y normalidad asintótica; Cheng y Lin (1987) examinan el estimador de máxima verosimilitud de S_T y comprueban que es asintóticamente más eficiente que el estimador de Kaplan-Meier; Herbst (1992a) construye un estadístico de Kolmogorov-Smirnov para contrastar la validez del modelo de Koziol-Green; y Herbst (1992b, 1993) estima los momentos y momentos residuales respectivamente, y demuestra su consistencia y normalidad asintótica.

Este modelo puede extenderse con facilidad al caso de doble censura: considérese que $T_1,..,T_N$, $X_1,..,X_N$ e $Y_1,..,Y_N$ son tres muestras aleatorias de las variables independientes T, X e Y respectivamente, donde T representa el tiempo de ocurrencia de un determinado suceso de interés y X e Y se corresponden con las variables de censura por la izquierda y por la derecha. Además, se supone que solamente se observan los pares (Z_i, δ_i), $i = 1, ..., N$, donde

$$Z_i = \max \left\{ \min \left(T_i, Y_i \right), X_i \right\},$$

y

$$\delta_i = \begin{cases} 0 & si & X_i \leq T_i \leq Y_i & (no\ censura) \\ 1 & si & X_i \leq Y_i < T_i & (censura\ por\ la\ derecha) \\ 2 & si & T_i < X_i \leq Y_i \ \ o \ \ Y_i < X_i & (censura\ por\ la\ izquierda) \end{cases}$$

En este caso, la extensión natural del modelo de Koziol-Green a doble censura considera que existen dos números reales positivos desconocidos, β y γ, tales que $S_X(t) = S_T^{\gamma}(t)$ y $S_Y(t) = S_T^{\beta}(t)$, para todo $t \geq 0$. Por tanto, si T, X e Y son absolutamente continuas, existe proporcionalidad entre las tasas de fallo, esto es, $h_X(t) = \gamma h_T(t)$ y $h_Y(t) = \beta h_T(t)$ en su campo de definición común.

Varios autores estudian este modelo generalizado de Koziol-Green. Por ejemplo, Bravo y Esteban (1993a) y Fernández et al. (1995a) proponen sendos estimadores de S_T que extienden a los propuestos por Ebrahimi (1985) y Herbst (1992a), respectivamente, al caso de doble censura y demuestran su consistencia uniforme fuerte y normalidad asintótica. Bravo et al. (1995b) construyen un estimador de la función de supervivencia uniformemente fuertemente consistente y asintóticamente normal cuando el *tiempo de vida* es independiente del intervalo de observación y los datos son incompletos. En el contexto Bayesiano no paramétrico, Morales et al. (1990, 1991) presentan un procedimiento para la estimación de S_T, supuesto que existe un proceso de Dirichlet a priori sobre Z y T respectivamente, mientras que Fernández et al. (1995b, 1997) consideran que la información a priori se expresa mediante un proceso de Dirichlet sobre el vector observable completo (Z, δ), y proponen un estimador de S_T uniformemente fuertemente consistente y asintóticamente normal.

1.7 Análisis de supervivencia con datos incompletos

En multitud de ocasiones, y por muy diversas razones, se desconoce cierta información relevante para el análisis de la supervivencia. Muchos investigadores han propuesto métodos para salvar, en lo posible, esa falta de información, especialmente cuando se analizan datos de supervivencia que incluyen varios tipos de fallo. Estos modelos constan de una variable que mide el tiempo que transcurre hasta que sucede el fallo y de otra que indica el tipo de fallo que se presenta, denominadas respectivamente *tiempo y tipo de fallo*. Cuando se introducen los mecanismos de censura, a menudo, una observación censurada del *tiempo de fallo* impide constatar el *tipo de fallo*. Por ejemplo, Dinse (1982) considera que el *tiempo de fallo* está censurado por la derecha y que se desconoce el *tipo de fallo* de algunos individuos, y utiliza el algoritmo EM (Dempster et al., 1977) para estimar la distribución conjunta del *tiempo y tipo de fallo*. Además, demuestra que las estimaciones sucesivas son autoconsistentes (Efron, 1967) y convergen a la estimación no paramétrica de máxima verosimilitud. Otros autores que han propuesto métodos para estimar las funciones de supervivencia en este caso son, entre otros, Miyakawa (1984), Racine-Poon y Hoel (1984) y Kodell y Chen (1987). Asimismo, Peto et al. (1980), Lagakos (1982), Dinse (1986, 1988) y Lagakos y Louis (1988) discuten el problema de la pérdida de la causa de muerte en experimentos de carcinogénesis animal. En este contexto, Goetghebeur y Ryan (1990) obtienen un test logrank modificado para comparar la supervivencia en dos poblaciones, y Dewanji (1992) sugiere una modificación de este contraste. Goetghebeur y Ryan (1995) extienden los resultados que obtuvieron en 1990 a modelos de regresión con riesgos proporcionales, y proponen un método para analizar esta clase de datos de supervivencia, que incluyen tipos de fallo desconocidos para algunos individuos, basándose en una generalización de la estructura semiparamétrica discutida por Cox y Oakes (1984, sec. 9.4) que requiere el supuesto de "pérdida al azar" de Little y Rubin (1987).

Por otra parte, Suzuki (1985a, 1985b) analiza el problema de la estimación de la función de supervivencia con la presencia de censura aleatoria a la derecha, cuando se desconoce el *tiempo de censura* de algunos individuos, en el contexto no paramétrico y paramétrico respectivamente. En el primer artículo, generaliza el estimador no paramétrico de Kaplan-Meier para esta clase de datos y comprueba que es de máxima verosimilitud, además de consistente y asintóticamente normal. En el segundo, considera que la distribución del *tiempo de vida* tiene una formulación paramétrica conocida y utiliza las distribuciones exponencial y de Weibull como ejemplos. En este caso, Suzuki obtiene un estimador consistente y asintóticamente normal de los parámetros del modelo.

En este trabajo profundizamos en el campo de la estimación de la supervivencia en doble censura con datos incompletos, extendiendo algunos modelos propuestos para el caso de censura a la derecha e introduciendo nuevas técnicas paramétricas y semiparamétricas, y analizamos el efecto de la proporción de datos desconocidos sobre la eficiencia de los estimadores considerados. De este modo, en el capítulo 2 se presenta un método para la estimación paramétrica de distribuciones de supervivencia cuando la variable de censura por la derecha es sólo parcialmente observable. El estimador propuesto, que extiende de forma natural al considerado por Suzuki (1985b) en el caso de censura por la derecha, verifica propiedades asintóticas similares a las del estimador de máxima verosimilitud, y se obtiene bajo los supuestos de un modelo paramétrico conocido y la independencia entre los tiempos de vida y los correspondientes tiempos de censura. Además, se realiza un estudio computacional diseñado para analizar la eficiencia del estimador obtenido según el nivel de censura y el tamaño muestral.

En el tercer capítulo se presenta un procedimiento semiparamétrico para estimar una curva de supervivencia a partir de datos doblemente censurados, cuando se desconoce el *tiempo de censura por la derecha* de algunos individuos, que extiende la idea dada por Klein et al. (1990) en el caso de censura por la derecha. El estimador propuesto utiliza modelos paramétricos para las observaciones censuradas y trata de forma no paramétrica a las no censuradas. También, se estudian algunas propiedades asintóticas del estimador obtenido, como la consistencia uniforme fuerte y la convergencia débil a un proceso Gaussiano, y se analiza y compara su comportamiento mediante simulación.

En el capítulo 4 se considera el modelo generalizado de Koziol-Green en el contexto Bayesiano no paramétrico, y se propone un método de estimación de la función de supervivencia S_T con datos incompletos y doblemente censurados cuando el conocimiento a priori viene dado por un proceso de Dirichlet (Ferguson, 1973) sobre el vector de variables observables (Z, δ), como alternativa al obtenido por Morales et al. (1990) cuando la información a priori es sólo sobre la variable Z y los datos son completos. Además, se analizan las propiedades asintóticas del estimador propuesto, y se estudia y compara su eficacia mediante simulación.

Capítulo 2

Estimación paramétrica de distribuciones de supervivencia

2.1 Introducción

Un fin primordial del análisis de supervivencia o fiabilidad es estimar la distribución del tiempo que transcurre hasta el fallo de un determinado sistema, o de una de sus componentes (*tiempo de supervivencia, de vida o de fallo*). Existen muchas situaciones en las que el *tiempo de fallo* no es realmente una variable temporal, sino que representa la frecuencia de uso. Por ejemplo, la vida de un automóvil se mide generalmente por su kilometraje, y la de una fotocopiadora, por el número de copias que ha realizado. Asimismo, una compañía de seguros utiliza la cantidad total que se ha pagado a un cliente con una enfermedad prolongada para medir su *tiempo de vida* y, con frecuencia, para establecer la fiabilidad de un sistema diseñado para la gestión de grandes bases de datos se utiliza la cantidad total de datos almacenados en ella. En cualquier caso, todas estas variables miden, de algún modo, la longitud del intervalo de "tiempo" que transcurre hasta que sucede el fallo.

Debido a los métodos de muestreo, o a la ocurrencia de alguna otra causa de fallo que se desea eliminar del estudio, puede que el *tiempo de fallo* no sea directamente observable. En la mayoría de esos casos sólo se conoce si el *tiempo de fallo* es mayor o menor que un determinado *tiempo de censura por la derecha o por la izquierda*, respectivamente. En particular, la censura por la derecha ha sido extensamente tratada en la literatura científica. El análisis de datos de supervivencia o fiabilidad con un esquema de censura doble se ha considerado ampliamente en los últimos años. Los artículos de Chang (1990), Bravo y Esteban (1993a, 1993b), Gu y Zhang (1993), Jewell et al. (1994), Bravo et al. (1995a), Ren (1995), Mykland y Ren (1996), Zhang y Li (1996) y Fernández et al. (1997) son sólo algunos de los muchos trabajos que han aparecido recientemente en este área.

Este capítulo presenta un procedimiento para la estimación de parámetros de distribuciones de supervivencia cuando se analizan datos incompletos y doblemente censurados. Doblemente censurados en el sentido que algunas de las observaciones están censuradas por la derecha y otras están censuradas por la izquierda. Incompletos en el sentido que la variable de censura por la derecha sólo es parcialmente observable.

El método se ilustra mediante un ejemplo basado en un estudio de la fiabilidad de un determinado sistema de base de datos diseñado para la gestión de un gran volumen de información en la sección 2.2. Además, se discute la motivación y construcción del estimador propuesto. Este estimador, el cual extiende de forma natural al considerado por Suzuki (1985b) en el caso de censura por la derecha, se obtiene bajo los supuestos de un modelo paramétrico conocido de la función de supervivencia, seleccionado mediante estudios preliminares, y la independencia entre los *tiempos de fallo* y los correspondientes *tiempos de censura*. En la sección 2.3 se demuestra que el estimador obtenido presenta tan buenas propiedades asintóticas como las de un estimador de máxima verosimilitud. Por último, la sección 2.4 está dedicada a un estudio de simulación de Monte Carlo realizado con el fin de analizar el efecto de la proporción de datos desconocidos sobre la eficiencia del estimador propuesto.

2.2 Motivación y construcción del estimador

El ejemplo que se presenta en este capítulo se basa en un estudio sobre la fiabilidad de un determinado sistema de gestión de bases de datos (SGBD). Una compañía ofrece a sus clientes un contrato de mantenimiento con una duración de un año que consiste en la actualización del producto sin costo adicional, y el soporte técnico necesario para la resolución de los problemas que les planteen. Siempre que ocurra un fallo del SGBD durante el periodo de vigencia del contrato de mantenimiento, el cliente avisa al personal técnico de la compañía para conseguir que su sistema vuelva a activarse. Esta compañía está muy interesada en demostrar que su sistema es suficientemente fiable para gestionar grandes bases de datos y, a fin de realizar un estudio de su eficacia, ha diseñado un nuevo contrato de mantenimiento, en el cual se le pregunta al cliente tanto si su sistema ha fallado anteriormente como el tamaño actual de su base de datos. Se seleccionan para este estudio los mil primeros clientes que firman este nuevo contrato de mantenimiento. Debido a las características del negocio de la compañía, se presume que el tiempo necesario para obtener esos mil clientes no es superior a dos meses. Como se plantea que el estudio tendrá una duración de nueve meses, todos los contratos de mantenimiento correspondientes a la muestra seleccionada expirarán con posterioridad a la finalización del proyecto. El análisis de fiabilidad requiere conocer el tamaño de las bases de datos de los clientes elegidos cuando ocurre el primer fallo del SGBD. Por tanto, durante los nueve meses que constituyen el periodo de estudio, siempre que un cliente seleccionado llame debido a un fallo de su sistema, el personal técnico de la compañía le pedirá el tamaño de su base de datos, si efectivamente éste es su primer fallo. Al final del periodo de estudio, el vendedor del SGBD se pone en contacto (por ejemplo, por carta) con aquellos clientes seleccionados cuyos sistemas no han fallado aún para solicitarles el tamaño de sus bases de datos, ya que esas observaciones contienen importante información acerca del *tiempo de vida* del SGBD. Este estudio constituye un ejemplo claro de análisis de datos de fiabilidad. Obsérvese que, aunque tanto el periodo de mantenimiento como el de observación se miden realmente en unidades de tiempo, la variable que representa al *tiempo de vida*, esto es, *"el tamaño de la base de datos del cliente cuando ocurre el primer fallo"*, se mide en megabytes. Nótese también que la muestra de datos que se considera está doblemente censurada, y que estos datos se originan a partir de la observación de los SGBD en las tres posibles situaciones siguientes:

- El sistema falla durante el periodo de estudio. En ese caso, el cliente avisa a la compañía suministradora del software para solicitar soporte técnico. A su vez, puede observarse el tamaño exacto de la base de datos del cliente, el cual proporciona un dato no censurado.

- El sistema ha fallado antes de la firma del contrato de mantenimiento. El tamaño de la base de datos del cliente en el momento del fallo es desconocido. En esta ocasión, la compañía sólo es capaz de saber el tamaño de la base de datos del cliente al inicio del periodo de mantenimiento, el cual se almacena en el impreso del contrato. Este caso corresponde a una observación censurada por la izquierda.

- El sistema no ha fallado al término del estudio. Esta vez, la compañía no conoce ni el tamaño de la base de datos cuando falle el SGBD por primera vez (este evento no ha ocurrido aún) ni el tamaño de la base de datos al final del estudio. Sin embargo, como se mencionó antes, estos datos contienen importante información sobre el *tiempo de vida* del SGBD. La compañía puede solicitar por carta a un cierto porcentaje de esos clientes el tamaño de sus bases de datos al final del estudio. La información así obtenida puede ayudar a incrementar la eficiencia del estimador propuesto. Esta tercera situación corresponde al caso de datos censurados por la derecha con un cierto porcentaje de información desconocida.

2.2.1 El modelo

En este capítulo se considera el modelo de doble censura aleatoria. Supóngase que T es una variable aleatoria no negativa, con función de supervivencia o fiabilidad $S_T(t) = \Pr(T > t)$, que representa a la variable *tiempo de vida* en estudio. La observación de T está sujeta al intervalo aleatorio $[X, Y]$, esto es, T es observable si y sólo si pertenece a $[X, Y]$, donde X e Y son variables aleatorias no negativas y $\Pr(X \leq Y) = 1$. Si T no pertenece a $[X, Y]$, su valor exacto no puede determinarse.

Tan sólo puede saberse si T es menor que X o mayor que Y, y solamente se observa X o Y en cada caso. Formalmente, en esta situación, la información disponible sobre T puede expresarse mediante un par de variables aleatorias Z y δ, donde

$$Z = \max\{\min(T, Y), X\} \tag{2.1}$$

y

$$\delta = \begin{cases} 0 \quad si \quad X \leq T \leq Y & (no\ censura), \\ 1 \quad si \quad \quad T > Y & (censura\ por\ la\ derecha), \\ 2 \quad si \quad \quad T < X & (censura\ por\ la\ izquierda). \end{cases} \tag{2.2}$$

Como supuesto crucial del modelo, se asume que el mecanismo de censura es independiente del *tiempo de supervivencia*, es decir, que el intervalo aleatorio $[X, Y]$ no depende de la variable T.

A partir de estudios preliminares, se supone que la función de supervivencia de T se expresa mediante un modelo paramétrico conocido $S_T(t; \theta)$, donde θ es un vector de parámetros desconocidos que toma valores en un determinado espacio paramétrico Θ, cuyo soporte, $\{t \geq 0 : 0 < S_T(t; \theta) < 1\}$, no depende de θ.

El problema que se plantea es estimar el vector de parámetros de supervivencia, θ, a partir de una muestra de N pares independientes, (Z_i, δ_i) para $i = 1, ..., N$, donde (Z_i, δ_i) está definido como en (2.1)-(2.2).

En el caso del SGBD se tendrá una relación parcial de las observaciones censuradas por la derecha. Más especifícamente, se verifica que las cantidades Z_i no se observan cuando $\delta_i = 1$, y que, después de realizar un seguimiento de esos sistemas por correo, se obtienen los valores de Y únicamente para una determinada proporción de ellos. Sin esta información, un estimador apropiado de θ sólo puede obtenerse en casos especiales.

De este modo, para $i = 1, ..., N$, se considera que $F_i = 1$ si la unidad i-ésima está sujeta a un seguimiento posterior y que $F_i = 0$ en otro caso.

El número de observaciones no censuradas, censuradas por la derecha sin seguimiento, censuradas por la derecha con seguimiento y censuradas por la izquierda se denotan por N_0, N_{10}, N_{11} y N_2 respectivamente. La probabilidad de que una unidad esté sometida a un seguimiento posterior se representa por f^*, y se supone que N_{11} y $f^* = \Pr(F = 1 \mid \delta = 1)$ son ambos positivos.

2.2.2 El estimador de máxima pseudo-verosimilitud

Para construir el estimador del párametro desconocido, θ, se considera la distribución muestral de las cantidades observadas (Z_i, δ_i), $i = 1, ..., N$, la cual viene dada por

$$L = \prod_{i=1}^{N} \left[f_T(Z_i) \left\{ S_Y(Z_i^-) - S_X(Z_i) \right\} \right]^{\mathrm{I}(\delta_i = 0)} \left[f_Y(Z_i) S_T(Z_i) \right]^{\mathrm{I}(\delta_i = 1) F_i}$$

$$\left\{ \Pr(\delta_i = 1) \right\}^{\mathrm{I}(\delta_i = 1)(1 - F_i)} \left[f_X(Z_i) \left\{ 1 - S_T(Z_i^-) \right\} \right]^{\mathrm{I}(\delta_i = 2)},$$

donde f_T, f_X y f_Y son respectivamente las funciones de densidad de probabilidad de las variables aleatorias T, X e Y, las funciones S_T, S_X y S_Y son sus correspondientes funciones de supervivencia o fiabilidad, e $\mathrm{I}(\cdot)$ representa la función indicatriz del suceso $(\cdot)$.

Si se definen las variables condicionales siguientes

$$\begin{aligned} \widetilde{Z}_i &= Z_i \mid \{\delta_i = 0\}, & i &= 1, ..., N_0, \\ \overline{Z}_j &= Z_j \mid \{\delta_j = 1, F_j = 1\}, & j &= 1, ..., N_{11}, \\ \widehat{Z}_k &= Z_k \mid \{\delta_k = 2\}, & k &= 1, ..., N_2, \end{aligned}$$

la función de verosimilitud puede expresarse como:

$$L = \prod_{i=1}^{N_0} \left[f_T(\widetilde{Z}_i) \left\{ S_Y(\widetilde{Z}_i^-) - S_X(\widetilde{Z}_i) \right\} \right] \prod_{j=1}^{N_{11}} \left[f_Y(\overline{Z}_j) S_T(\overline{Z}_j) \right] \left\{ \Pr(\delta = 1) \right\}^{N_{10}}$$

$$\prod_{k=1}^{N_2} \left[f_X(\widehat{Z}_k) \left\{ 1 - S_T(\widehat{Z}_k^-) \right\} \right].$$

Teniendo en cuenta que $\Pr(\delta = 1)$ no puede calcularse analíticamente de forma sencilla, excepto en ciertos casos especiales (por ejemplo, cuando T e Y son exponenciales), se asume que la distribución de las N_{10} observaciones censuradas por la derecha y sin seguimiento posterior coincide con la distribución de las N_{11} con seguimiento. Por este motivo, la masa de los N_{10} valores desconocidos se redistribuye equitativamente entre las N_{11} observaciones conocidas de $\overline{Z}$, y se utiliza

$$\prod_{j=1}^{N_{11}} \left[f_Y(\overline{Z}_j) S_T(\overline{Z}_j) \right]^{N_{10}/N_{11}}$$

en vez de $\left\{\Pr(\delta = 1)\right\}^{N_{10}}$. Si además se supone que la censura es no informativa, esto es, que las funciones f_X, f_Y, S_X y S_Y no incluyen ningún parámetro de interés, resulta finalmente la siguiente función de pseudo-verosimilitud

$$L^* = \prod_{i=1}^{N_0} f_T(\widetilde{Z}_i) \prod_{j=1}^{N_{11}} \left\{ S_T(\overline{Z}_j) \right\}^{1+N_{10}/N_{11}} \prod_{k=1}^{N_2} \left\{ 1 - S_T(\widehat{Z}_k^-) \right\}.$$

El estimador de θ que se propone es el vector $\theta^* \in \Theta$, en el cual la pseudo-verosimilitud L^* sea máxima.

Como $1 + N_{10}/N_{11}$ converge casi seguramente a $1/f^*$, también se podría considerar la pseudo-verosimilitud

$$L^{**} = \prod_{i=1}^{N_0} f_T(\widetilde{Z}_i) \prod_{j=1}^{N_{11}} \left\{ S_T(\overline{Z}_j) \right\}^{1/f^*} \prod_{k=1}^{N_2} \left\{ 1 - S_T(\widehat{Z}_k^-) \right\},$$

que generaliza la propuesta por Kalbfleisch y Lawless (1988) en el caso de censura a la derecha, y denotar por θ^{**} al valor del parámetro que maximice L^{**}. Sin embargo, la atención se centrará principalmente en el estimador θ^* ya que proporciona mejores resultados.

Observaciones

- Nótese que θ^* y θ^{**} son válidos incluso cuando los modelos de X e Y son no paramétricos.
- Cuando la variable de censura por la izquierda, X, también es parcialmente observable, las expresiones de L^* y L^{**} son inmediatas. No obstante, no se estudiará aquí ese caso.

2.2.3 Familias de tipo exponencial

En este apartado se supone que la función de densidad de T pertenece a una familia de tipo exponencial en su parametrización natural, esto es,

$$f_T(t) = h(t) \exp\left\{\theta^t u(t) + v(\theta)\right\}, \tag{2.3}$$

donde $h(t)$ es una función escalar de t, $u(t)$ es una vector $p \times 1$ de funciones de t linealmente independientes y $v(\theta)$ es una función escalar del parámetro θ.

Por tanto, la log pseudo-verosimilitud vendrá dada en este caso por:

$$\log L^*(\theta) = \sum_{i=1}^{N_0} \log f_T(\widetilde{Z}_i;\theta) + \left(1+\frac{N_{10}}{N_{11}}\right)\sum_{j=1}^{N_{11}} \log S_T(\overline{Z}_j;\theta) + \sum_{k=1}^{N_2} \log\left\{1 - S_T(\widehat{Z}_k;\theta)\right\}.$$

Como se verifica que

$$\frac{\partial}{\partial\theta}\log f_T(\widetilde{Z}_i;\theta) = \frac{\partial}{\partial\theta}\left\{\log h(\widetilde{Z}_i) + \theta^t u(\widetilde{Z}_i) + v(\theta)\right\} = u(\widetilde{Z}_i) + \frac{\partial}{\partial\theta}v(\theta),$$

$$\frac{\partial}{\partial\theta}\log S_T(\overline{Z}_j;\theta) = \frac{\frac{\partial}{\partial\theta}S_T(\overline{Z}_j;\theta)}{S_T(\overline{Z}_j;\theta)} = E_\theta\left[u(T_j) \mid T_j > \overline{Z}_j\right] + \frac{\partial}{\partial\theta}v(\theta),$$

$$\frac{\partial}{\partial\theta}\log\left\{1 - S_T(\widehat{Z}_k;\theta)\right\} = \frac{-\frac{\partial}{\partial\theta}S_T(\widehat{Z}_k;\theta)}{1 - S_T(\widehat{Z}_k;\theta)} = E_\theta\left[u(T_k) \mid T_k < \widehat{Z}_k\right] + \frac{\partial}{\partial\theta}v(\theta)$$

y $(\partial/\partial\theta)v(\theta) = -E_\theta\left[u(T)\right]$, donde $\partial/\partial\theta$ denota al correspondiente vector columna de derivadas parciales o gradiente respecto de θ, se obtiene que

$$\begin{aligned}\frac{\partial}{\partial\theta}\log L^*(\theta) &= \sum_{i=1}^{N_0} u(\widetilde{Z}_i) + \sum_{j=1}^{N_{11}} E_\theta\left[u(T_j) \mid T_j > \overline{Z}_j\right] \\ &+ \sum_{k=1}^{N_2} E_\theta\left[u(T_k) \mid T_k < \widehat{Z}_k\right] - N E_\theta\left[u(T)\right] \\ &= \sum_{i\in I^*} \omega_i^* \left\{E_\theta\left[u(T_i) \mid D\right] - E_\theta\left[u(T)\right]\right\},\end{aligned}$$

donde I^* es un conjunto indexado que representa a las $(N_0 + N_{11} + N_2)$ observaciones conocidas, $\omega_i^* = 1 + \mathrm{I}(\delta_i = 1,\ F_i = 1)N_{10}/N_{11}$, y D es el conjunto de los datos observados. Luego, la solución θ^* de las ecuaciones de pseudo-verosimilitud se obtiene igualando la esperanza de $u(T)$ al promedio pesado de las esperanzas de los $u(T_i)'$s condicionadas a los datos observados. Más formalmente, θ^* se obtiene como solución de

$$\frac{1}{N}\sum_{i\in I^*} \omega_i^* E_\theta\left[u(T_i) \mid D\right] = E_\theta\left[u(T)\right].$$

Análogamente, θ^{**} es solución de la ecuación

$$\frac{1}{N}\sum_{i\in I^*} \omega_i^{**} E_\theta\left[u(T_i) \mid D\right] = E_\theta\left[u(T)\right],$$

donde $\omega_i^{**} = 1 + (1/f^* - 1)\mathrm{I}(\delta_i = 1,\ F_i = 1)$, para $i \in I^*$.

2.2.4 El algoritmo EM

El algoritmo EM, propuesto por Dempster et al. (1977), es altamente eficaz cuando la log verosimilitud $\log L_0(\theta; D_0)$ de los datos D_0 que se observarían si no existiera censura tiene una expresión funcional apreciablemente más simple que la log verosimilitud $\log L(\theta; D)$ de los datos D que realmente se observan.

En el caso que se plantea, la función $Q(\theta', \theta) = E_\theta\left[\log L_0(\theta'; D_0) \mid D\right]$ puede expresarse como

$$\sum_{i=1}^{N_0} E_\theta \left[\log f_T(T_i;\theta') \mid T_i = \widetilde{Z}_i\right] + \left(1+\frac{N_{10}}{N_{11}}\right)\sum_{j=1}^{N_{11}} E_\theta \left[\log f_T(T_j;\theta') \mid T_j > \overline{Z}_j\right]$$

$$+\sum_{k=1}^{N_2} E_\theta \left[\log f_T(T_k;\theta') \mid T_k < \widehat{Z}_k\right].$$

Dada la estimación actual $\theta^*_{(j)}$ de θ, el algoritmo EM determina una nueva estimación $\theta^*_{(j+1)}$ como aquel valor de θ' que maximiza $Q(\theta',\theta^*_{(j)})$. Bajo condiciones bastante generales (véase Wu, 1983) la sucesión $\left\{\theta^*_{(j)}\right\}$ converge al valor θ^* que maximiza la log pseudo-verosimilitud $\log L^*(\theta)$.

Si la densidad de T está definida por la expresión (2.3), entonces

$$Q(\theta',\theta) = \sum_{i\in I^*} \omega_i^* \theta'^{t} E_\theta\left[u(T_i) \mid D\right] + Nv(\theta')$$

salvo una función aditiva que no depende de θ'. La maximización de Q en θ' puede obtenerse anulando $(\partial/\partial\theta')\, Q(\theta',\theta)$, lo cual proporciona la siguiente ecuación en θ', para un θ dado,

$$E_{\theta'}\left[u(T)\right] = \frac{1}{N}\sum_{i\in I^*} \omega_i^* E_\theta\left[u(T_i) \mid D\right]. \tag{2.4}$$

De este modo, si θ es un punto fijo del algoritmo EM, es decir, $\theta' = \theta$ en la expresión (2.4), entonces θ es también una solución de la ecuación de pseudo-verosimilitud $(\partial/\partial\theta)\log L^*(\theta) = 0$.

La log pseudo-verosimilitud $\log L^*(\theta)$ no decrece en ninguna iteración del algoritmo EM. Como consecuencia inmediata, el estimador de máxima pseudo-verosimilitud θ^* debe satisfacer la condición de autoconsistencia siguiente

$$Q(\theta^*,\theta^*) \geq Q(\theta',\theta^*), \text{ para todo } \theta' \in \Theta,$$

ya que, si iniciamos el proceso con $\theta = \theta^*$, la log pseudo-verosimilitud $\log L^*(\theta)$ no puede incrementarse en ninguna iteración del algoritmo EM, o de cualquier otra forma.

2.2.5 Ejemplos

Caso exponencial: Si la distribución de la variable T es exponencial de parámetro λ, entonces su función de supervivencia estará definida por $S_T(t) = \exp(-\lambda t)$ para $t > 0$. Por consiguiente,

$$L^*(\lambda) = \prod_{i=1}^{N_0} \lambda \exp(-\lambda Z_i) \prod_{j=1}^{N_{11}} \exp\left\{-\lambda\left(1+\frac{N_{10}}{N_{11}}\right) Z_j\right\} \prod_{k=1}^{N_2} \{1-\exp(-\lambda Z_k)\}.$$

De $(\partial/\partial\lambda)\log L^*(\lambda) = 0$, resulta la ecuación

$$\frac{N_0}{\lambda} - \sum_{i=1}^{N_0} Z_i - \left(1+\frac{N_{10}}{N_{11}}\right)\sum_{j=1}^{N_{11}} Z_j + \sum_{k=1}^{N_2} \frac{Z_k}{\{\exp(\lambda Z_k)-1\}} = 0.$$

La solución, λ^*, puede hallarse, por ejemplo, mediante el método de Newton-Raphson. De esta forma, si la estimación actual de λ es $\lambda^*_{(j)}$, la nueva estimación $\lambda^*_{(j+1)}$ estará definida por

$$\lambda^*_{(j+1)} = \lambda^*_{(j)} - \frac{\dfrac{\partial}{\partial\lambda}\log L^*(\lambda^*_{(j)})}{\dfrac{\partial^2}{\partial\lambda^2}\log L^*(\lambda^*_{(j)})},$$

donde

$$\frac{\partial^2}{\partial\lambda^2}\log L^*(\lambda) = -\frac{N_0}{\lambda^2} - \sum_{k=1}^{N_2}\frac{Z_k^2\exp(\lambda Z_k)}{\{\exp(\lambda Z_k)-1\}^2}.$$

También puede utilizarse el algoritmo EM. En ese caso, como $Q(\lambda', \lambda)$ es

$$\lambda'\left\{-\sum_{i=1}^{N_0} Z_i - \left(1+\frac{N_{10}}{N_{11}}\right)\sum_{j=1}^{N_{11}} Z_j + \sum_{k=1}^{N_2}\frac{Z_k}{\{\exp(\lambda Z_k)-1\}} - \frac{N_1+N_2}{\lambda}\right\} + N\log\lambda',$$

dada la estimación actual $\lambda^*_{(j)}$ de λ, la nueva estimación $\lambda^*_{(j+1)}$ estará definida por

$$\lambda^*_{(j+1)} = \frac{N}{\displaystyle\sum_{i=1}^{N_0} Z_i + \left(1+\frac{N_{10}}{N_{11}}\right)\sum_{j=1}^{N_{11}} Z_j - \sum_{k=1}^{N_2}\frac{Z_k}{\left\{\exp(\lambda^*_{(j)} Z_k)-1\right\}} + \frac{N_1+N_2}{\lambda^*_{(j)}}}.$$

Como ilustración, considérense los datos de la tabla 1 de Mantel (1967): 2^-, 3, 5^-, 6, 6^+, 7, 7, 8, 8^+, 9^+, 10, 11^-, 12, 15 y 18^+, donde los superíndices + y − indican respectivamente censura a la derecha y a la izquierda en el correspondiente dato. Si se eliminan las observaciones 6^+ y 18^+, con lo cual $N_{10} = N_{11} = 2$, y se toma como $\lambda^*_{(0)}$ a la inversa de la media muestral, resulta la siguiente tabla

j	0	1	2	3	4	5	6
$\lambda^*_{(j)}$	0.1263	0.1063	0.1018	0.1006	0.1003	0.1002	0.1002
$\log L^*\left(\lambda^*_{(j)}\right)$	-31.979	-31.685	-31.667	-31.666	-31.666	-31.666	-31.666

Mediante el método de Newton-Raphson los resultados hubieran sido

j	0	1	2	3	4
$\lambda^*_{(j)}$	0.1262	0.0934	0.0997	0.1002	0.1002
$\log L^*\left(\lambda^*_{(j)}\right)$	-31.979	-31.693	-31.666	-31.666	-31.666

Por otra parte, utilizando los datos completos de Mantel, la estimación de máxima verosimilitud de λ hubiera sido 0.094145.

Una tercera posibilidad es la aplicación de métodos de bisección ya que es posible acotar el valor de λ^*. Como $0 \leq T_i \leq X_i$ cuando $\delta_i = 2$, puede obtenerse una cota inferior, λ^*_L, y una cota superior, λ^*_U, de λ^* considerando respectivamente, en tales casos, que $T_i = X_i$ y que $T_i = 0$. Se comprueba con facilidad que

$$\lambda^*_L = \frac{N_0 + N_2}{\sum_{i=1}^{N_0} Z_i + \left(1 + \frac{N_{10}}{N_{11}}\right) \sum_{j=1}^{N_{11}} Z_j + \sum_{k=1}^{N_2} Z_k} \quad \text{y} \quad \lambda^*_U = \frac{N_0 + N_2}{\sum_{i=1}^{N_0} Z_i + \left(1 + \frac{N_{10}}{N_{11}}\right) \sum_{j=1}^{N_{11}} Z_j}.$$

El intervalo inicial de acotación $[\lambda^*_L, \lambda^*_U]$ puede reducirse repetidamente (por ejemplo, mediante el método regula falsi) hasta conseguir la precisión deseada para λ^*. En el ejemplo que se analizó anteriormente se obtiene que $\lambda^*_L = 0.091667$ y $\lambda^*_U = 0.10784$, mientras que $\lambda^* = 0.10021$.

El valor de λ^{**} puede hallarse mediante un desarrollo similar sustituyendo $1 + N_{10}/N_{11}$ por $1/f^*$ en las fórmulas precedentes.

Caso Weibull: Si T sigue la distribución de Weibull de parámetro $\theta = (\lambda, r)^t$, $S_T(t) = \exp(-\lambda t^r)$ para $t > 0$, se verifica que

$$L^*(\theta) = \prod_{i=1}^{N_0} \lambda r Z_i^{r-1} \exp(-\lambda Z_i^r) \prod_{j=1}^{N_{11}} \exp\left\{-\left(1 + \frac{N_{10}}{N_{11}}\right) \lambda Z_j^r\right\} \prod_{k=1}^{N_2} \{1 - \exp(-\lambda Z_k^r)\}.$$

A partir de $(\partial/\partial\lambda) \log L^*(\theta) = 0$ y $(\partial/\partial r) \log L^*(\theta) = 0$, resulta que

$$\frac{N_0}{\lambda} - \sum_{i=1}^{N_0} Z_i^r - \left(1 + \frac{N_{10}}{N_{11}}\right) \sum_{j=1}^{N_{11}} Z_j^r + \sum_{k=1}^{N_2} \frac{Z_k^r}{\exp(\lambda Z_k^r) - 1} = 0,$$

$$\frac{N_0}{r} + \sum_{i=1}^{N_0} (1 - \lambda Z_i^r) \log Z_i - \left(1 + \frac{N_{10}}{N_{11}}\right) \lambda \sum_{j=1}^{N_{11}} Z_j^r \log Z_j + \lambda \sum_{k=1}^{N_2} \frac{Z_k^r \log Z_k}{\exp(\lambda Z_k^r) - 1} = 0.$$

Este sistema de dos ecuaciones puede resolverse utilizando, por ejemplo, el método de Newton-Raphson, obteniéndose así su correspondiente solución $\theta^* = (\lambda^*, r^*)^t$. De este modo, si la estimación actual de θ es $\theta^*_{(j)}$, la nueva estimación $\theta^*_{(j+1)}$ estará definida por

$$\theta^*_{(j+1)} = \theta^*_{(j)} - \begin{pmatrix} \dfrac{\partial^2}{\partial\lambda^2} \log L^*(\theta^*_{(j)}) & \dfrac{\partial^2}{\partial\lambda\partial r} \log L^*(\theta^*_{(j)}) \\ \dfrac{\partial^2}{\partial r\partial\lambda} \log L^*(\theta^*_{(j)}) & \dfrac{\partial^2}{\partial r^2} \log L^*(\theta^*_{(j)}) \end{pmatrix}^{-1} \frac{\partial}{\partial\theta} \log L^*(\theta^*_{(j)}),$$

donde

$$
\begin{aligned}
\frac{\partial^2}{\partial \lambda^2} \log L^*(\theta) &= -\frac{N_0}{\lambda^2} - \sum_{k=1}^{N_2} \frac{Z_k^{2r} \exp(\lambda Z_k^r)}{\{\exp(\lambda Z_k^r) - 1\}^2}\,, \\
\frac{\partial^2}{\partial \lambda \partial r} \log L^*(\theta) &= -\sum_{i=1}^{N_0} Z_i^r \log Z_i - \left(1 + \frac{N_{10}}{N_{11}}\right) \sum_{j=1}^{N_{11}} Z_j^r \log Z_j + \sum_{k=1}^{N_2} \frac{Z_k^r \log Z_k}{\exp(\lambda Z_k^r) - 1} \\
&\quad -\lambda \sum_{k=1}^{N_2} \frac{Z_k^{2r} \exp(\lambda Z_k^r) \log Z_k}{\{\exp(\lambda Z_k^r) - 1\}^2}\,, \\
\frac{\partial^2}{\partial r \partial \lambda} \log L^*(\theta) &= \frac{\partial^2}{\partial \lambda \partial r} \log L^*(\theta), \\
\frac{\partial^2}{\partial r^2} \log L^*(\theta) &= -\frac{N_0}{r^2} - \lambda \sum_{i=1}^{N_0} Z_i^r (\log Z_i)^2 - \lambda \left(1 + \frac{N_{10}}{N_{11}}\right) \sum_{j=1}^{N_{11}} Z_j^r (\log Z_j)^2 \\
&\quad + \lambda \left\{ \sum_{k=1}^{N_2} \frac{Z_k^r (\log Z_k)^2}{\exp(\lambda Z_k^r) - 1} - \lambda \sum_{k=1}^{N_2} \frac{Z_k^{2r} (\log Z_k)^2 \exp(\lambda Z_k^r)}{\{\exp(\lambda Z_k^r) - 1\}^2} \right\}.
\end{aligned}
$$

Como en el ejemplo anterior, se consideran los datos de Mantel (1967), salvo 6^+ y 18^+, y se toma como valores iniciales $\lambda^*_{(0)} = 0.0045$ y $r^*_{(0)} = 2.465$, estimados mediante el método de los momentos sin tener en cuenta la censura de algunos datos. Entonces, la nueva estimación de $\theta^* = (\lambda^*, r^*)^{\mathrm{t}}$ sería:

$$
\begin{pmatrix} \lambda^*_{(1)} \\ r^*_{(1)} \end{pmatrix} = \begin{pmatrix} 0.0045 \\ 2.4650 \end{pmatrix} - \begin{pmatrix} -533037 & -5017.34 \\ -5017.34 & -74.7898 \end{pmatrix}^{-1} \begin{pmatrix} -555.962 \\ -7.18571 \end{pmatrix}.
$$

Esto es, $\lambda^*_{(1)} = 0.0041238$ y $r^*_{(1)} = 2.39416$. Continuando con el proceso se obtienen las estimaciones $\lambda^* = 0.0117448$ y $r^* = 1.95939$.

Como en el caso exponencial, puede acotarse el valor de λ^*. Si r^* es conocido, se verifica que $\lambda^*_L \leq \lambda^* \leq \lambda^*_U$, donde

$$
\lambda^*_L = \frac{N_0 + N_2}{\sum_{i=1}^{N_0} Z_i^{r^*} + \left(1 + \frac{N_{10}}{N_{11}}\right) \sum_{j=1}^{N_{11}} Z_j^{r^*} + \sum_{k=1}^{N_2} Z_k^{r^*}} \quad \text{y} \quad \lambda^*_U = \frac{N_0 + N_2}{\sum_{i=1}^{N_0} Z_i^{r^*} + \left(1 + \frac{N_{10}}{N_{11}}\right) \sum_{j=1}^{N_{11}} Z_j^{r^*}}.
$$

Con los mismos datos de Mantel que se utilizaron con anterioridad se obtiene que $\lambda_L^* = 0.0108145$ y $\lambda_U^* = 0.0124989$, mientras que $\lambda^* = 0.0117448$.

De forma análoga, es posible hallar el valor de $\theta^{**} = (\lambda^{**}, r^{**})^{\mathrm{t}}$ considerando en las fórmulas precedentes $1/f^*$ en vez de $1 + N_{10}/N_{11}$.

2.3 Propiedades asintóticas del estimador

Es bien sabido que, cuando todas las unidades están sujetas a seguimiento y se satisfacen las apropiadas condiciones de regularidad, el estimador de máxima verosimilitud, $\widehat{\theta}$, del parámetro θ es asintóticamente la única solución de la ecuación de verosimilitud $(\partial/\partial\theta) \log L_0(\theta) = 0$, donde $L_0(\theta)$ es la verosimilitud total que se obtiene cuando cada valor, censurado o no, es observado, es decir, cuando $f^* = 1$. Mäkeläinen et al. (1981) proporcionan condiciones suficientes para asegurar la existencia y unicidad de este estimador. A su vez, $\widehat{\theta}$ es un estimador consistente de θ con distribución asintóticamente normal de media θ y matriz de varianzas-covarianzas $J(\theta)^{-1}$, donde

$$J(\theta) = -E\left[(\partial^2/\partial\theta\partial\theta^{\mathrm{t}}) \log L_0(\theta)\right]$$

se denomina matriz de información de Fisher (véase, por ejemplo, Le Cam, 1953, 1970; Cox y Hinkley, 1974, sec 9.2; Lehmann, 1983, sec. 6.2).

En esta sección se demuestra que, bajo las mismas condiciones de regularidad que satisface $\widehat{\theta}$, el estimador de máxima pseudo-verosimilitud propuesto, θ^*, verifica idénticas propiedades asintóticas. De igual modo, se comprueba que θ^{**} presenta propiedades análogas.

Teorema 2.1

El estimador de máxima pseudo-verosimilitud θ^* es un estimador consistente de θ. Asimismo, θ^{**} también verifica esta propiedad.

Demostración

Para probar este teorema se demuestra que las soluciones de los sistemas

$R^*(\theta) = (\partial/\partial\theta) \log L^*(\theta) = 0$ y $R_0(\theta) = (\partial/\partial\theta) \log L_0(\theta) = 0$ coinciden asintóticamente.

Para el sistema $R^*(\theta) = 0$ se verifica que

$$\begin{aligned}
\frac{1}{N} R^*(\theta) &= \frac{1}{N} \sum_{i=1}^{N_0} R(\widetilde{Z}_i; \theta) + \frac{1}{N} \left(1 + \frac{N_{10}}{N_{11}}\right) \sum_{j=1}^{N_{11}} R(\overline{Z}_j; \theta) + \frac{1}{N} \sum_{k=1}^{N_2} R(\widehat{Z}_k; \theta) \\
&= \frac{N_0}{N} \frac{1}{N_0} \sum_{i=1}^{N_0} R(\widetilde{Z}_i; \theta) + \left(\frac{N_{10} + N_{11}}{N}\right) \frac{1}{N_{11}} \sum_{j=1}^{N_{11}} R(\overline{Z}_j; \theta) \\
&\quad + \frac{N_2}{N} \frac{1}{N_2} \sum_{k=1}^{N_2} R(\widehat{Z}_k; \theta) = 0,
\end{aligned}$$

donde

$$\begin{aligned}
R(\widetilde{Z}_i; \theta) &= \frac{\partial}{\partial\theta} \log f_T(\widetilde{Z}_i; \theta), & i &= 1, ..., N_0, \\
R(\overline{Z}_j; \theta) &= \frac{\partial}{\partial\theta} \log S_T(\overline{Z}_j; \theta), & j &= 1, ..., N_{11}, \\
R(\widehat{Z}_k; \theta) &= \frac{\partial}{\partial\theta} \log \left\{1 - S_T(\widehat{Z}_k; \theta)\right\}, & k &= 1, ..., N_2.
\end{aligned}$$

Tomando límite en probabilidad, resulta que

$$\frac{1}{N} R^*(\theta) \xrightarrow[N \to \infty]{\Pr} p_0 E\left[R(\widetilde{Z}; \theta)\right] + p_1 E\left[R(\overline{Z}; \theta)\right] + p_2 E\left[R(\widehat{Z}; \theta)\right],$$

donde $p_i = \Pr(\delta = i)$, para $i = 0, 1, 2$. Asimismo, se observa que el miembro de la derecha de la expresión anterior es igual al límite en probabilidad de $R_0(\theta)/N$ cuando $N \to \infty$, donde

$$R_0(\theta) = \sum_{i=1}^{N_0} R(\widetilde{Z}_i; \theta) + \sum_{j=1}^{N_1} R(\overline{Z}_j; \theta) + \sum_{k=1}^{N_2} R(\widehat{Z}_k; \theta).$$

De este modo, las soluciones de ambos sistemas, $R^*(\theta) = 0$ y $R_0(\theta) = 0$, coinciden asintóticamente. Por tanto, θ^* es un estimador consistente de θ. Idéntica propiedad verifica θ^{**} ya que también

$$\frac{1}{N}R^{**}(\theta) = \frac{\partial}{\partial\theta}\log L^{**}(\theta) \xrightarrow[N\to\infty]{\text{Pr}} p_0 E\left[R(\widetilde{Z};\theta)\right] + p_1 E\left[R(\overline{Z};\theta)\right] + p_2 E\left[R(\widehat{Z};\theta)\right],$$

donde $R^{**}(\theta) = (\partial/\partial\theta)\log L^{**}(\theta)$.

□

Teorema 2.2

La distribución asintótica de $\sqrt{N}(\theta^* - \theta)$ es normal de media cero y matriz de varianzas-covarianzas $\Phi(\theta)^{-1}\Sigma^*(\theta, f^*)\Phi(\theta)^{-1}$, donde

$$\begin{aligned}\Phi(\theta) = p_0 E\left[\frac{\partial^2}{\partial\theta\partial\theta^{\mathrm{t}}}\log f_T(\widetilde{Z};\theta)\right] + p_1 E\left[\frac{\partial^2}{\partial\theta\partial\theta^{\mathrm{t}}}\log S_T(\overline{Z};\theta)\right] \\ + p_2 E\left[\frac{\partial^2}{\partial\theta\partial\theta^{\mathrm{t}}}\log\left\{1 - S_T(\widehat{Z};\theta)\right\}\right]\end{aligned} \tag{2.5}$$

y

$$\begin{aligned}\Sigma^*(\theta, f^*) \quad = \quad & p_0 E\left[\frac{\partial}{\partial\theta}\log f_T(\widetilde{Z};\theta)\frac{\partial}{\partial\theta^{\mathrm{t}}}\log f_T(\widetilde{Z};\theta)\right] \\ & + \frac{p_1}{f^*}E\left[\frac{\partial}{\partial\theta}\log S_T(\overline{Z};\theta)\frac{\partial}{\partial\theta^{\mathrm{t}}}\log S_T(\overline{Z};\theta)\right] \\ & + p_2 E\left[\frac{\partial}{\partial\theta}\log\left\{1 - S_T(\widehat{Z};\theta)\right\}\frac{\partial}{\partial\theta^{\mathrm{t}}}\log\left\{1 - S_T(\widehat{Z};\theta)\right\}\right] \\ & - \frac{p_1(1-f^*)}{f^*}E\left[\frac{\partial}{\partial\theta}\log S_T(\overline{Z};\theta)\right]E\left[\frac{\partial}{\partial\theta^{\mathrm{t}}}\log S_T(\overline{Z};\theta)\right].\end{aligned} \tag{2.6}$$

A su vez, la distribución de $\sqrt{N}(\theta^{**} - \theta)$ es asintóticamente normal de media cero y matriz de varianzas-covarianzas $\Phi(\theta)^{-1}\Sigma^{**}(\theta, f^*)\Phi(\theta)^{-1}$, donde

$$
\begin{aligned}
\Sigma^{**}(\theta, f^*) \;\; = \;\; & p_0 E\left[\frac{\partial}{\partial\theta}\log f_T(\widetilde{Z};\theta)\frac{\partial}{\partial\theta^{\mathrm{t}}}\log f_T(\widetilde{Z};\theta)\right] \\
& +\frac{p_1}{f^*}E\left[\frac{\partial}{\partial\theta}\log S_T(\overline{Z};\theta)\frac{\partial}{\partial\theta^{\mathrm{t}}}\log S_T(\overline{Z};\theta)\right] \\
& +p_2 E\left[\frac{\partial}{\partial\theta}\log\left\{1-S_T(\widehat{Z};\theta)\right\}\frac{\partial}{\partial\theta^{\mathrm{t}}}\log\left\{1-S_T(\widehat{Z};\theta)\right\}\right]
\end{aligned} \tag{2.7}
$$

Demostración

Como $R^*(\theta^*) = (\partial/\partial\theta)\log L^*(\theta^*) = 0$, mediante un desarrollo de Taylor de $R^*(\theta^*)$ en un entorno del verdadero valor del parámetro θ, se obtiene que

$$
R^*(\theta^*) = R^*(\theta) + NM(\widetilde{\theta})(\theta^* - \theta) = 0, \tag{2.8}
$$

donde

$$
M(\widetilde{\theta}) = \frac{1}{N}\frac{\partial^2}{\partial\theta\partial\theta^{\mathrm{t}}}\log L^*(\widetilde{\theta})
$$

y $\widetilde{\theta}$ es un determinado punto del segmento que une θ^* y θ.

Puesto que $E\left[M(\theta)\right] = \Phi(\theta)$, se verifica por la ley de los grandes números que $M(\theta) \xrightarrow[N\to\infty]{\mathrm{Pr}} \Phi(\theta)$.

Por las condiciones de regularidad, se sabe que $J(\theta)$ es definida positiva. Por tanto, la matriz $\Phi(\theta)$, que coincide con $-N \cdot J(\theta)$, es definida negativa. Este hecho implica la no singularidad de $M(\theta)$ con probabilidad uno para todo N suficientemente grande. Luego, se obtiene a partir de (2.8) que

$$
\sqrt{N}(\theta^* - \theta) = -M(\widetilde{\theta})^{-1}R^*(\theta)/\sqrt{N} \;\; \text{casi seguramente } (c.\,s.) \;\; \text{cuando} \;\; N \to \infty.
$$

Como θ^* es un estimador consistente de θ por el teorema 2.1 y se cumple que $\widetilde{\theta} = \mu\theta^* + (1-\mu)\theta$ para algún $\mu \in [0,1]$, entonces $\widetilde{\theta} \xrightarrow[N\to\infty]{\mathrm{Pr}} \theta$. Por tanto, resulta que $M(\widetilde{\theta}) - M(\theta) \xrightarrow[N\to\infty]{\mathrm{Pr}} 0$, esto es, $M(\widetilde{\theta}) = M(\theta) + o_p(1)$.

De esta forma, se cumple que

$$
\sqrt{N}(\theta^* - \theta) = -M(\theta)^{-1}R^*(\theta)/\sqrt{N} + o_p(1) = -\Phi(\theta)^{-1}R^*(\theta)/\sqrt{N} + o_p(1).
$$

Además, las condiciones de regularidad aseguran que $R^*(\theta)/\sqrt{N}$ verifica el teorema central del límite multidimensional. Por tanto, converge en distribución a una variable aleatoria normal, esto es,

$$\frac{1}{\sqrt{N}} R^*(\theta) \xrightarrow[N\to\infty]{d} N(0, \Sigma^*(\theta, f^*)),$$

donde $\Sigma^*(\theta, f^*)$ está definida por la expresión (2.6), y consecuentemente

$$\sqrt{N}(\theta^* - \theta) \xrightarrow[N\to\infty]{d} N(0, \Phi(\theta)^{-1}\Sigma^*(\theta, f^*)\Phi(\theta)^{-1}).$$

Mediante un desarrollo similar, resulta que

$$\sqrt{N}(\theta^{**} - \theta) \xrightarrow[N\to\infty]{d} N(0, \Phi(\theta)^{-1}\Sigma^{**}(\theta, f^*)\Phi(\theta)^{-1})$$

ya que en esta ocasión

$$\frac{1}{\sqrt{N}} R^{**}(\theta) \xrightarrow[N\to\infty]{d} N(0, \Sigma^{**}(\theta, f^*)) \quad \text{y} \quad \sqrt{N}(\theta^{**} - \theta) + \Phi(\theta)^{-1} R^*(\theta)/\sqrt{N} \xrightarrow[N\to\infty]{\Pr} 0.$$

Para finalizar, se procede a comprobar que la esperanza de $R^*(\theta)/\sqrt{N}$ y de $R^{**}(\theta)/\sqrt{N}$ es el vector cero y que sus correspondientes matrices de varianzas-covarianzas asintóticas vienen dadas por la expresiones (2.6) y (2.7).

Por simplicidad, se supone que $\theta = (\theta_1, \theta_2)^{\mathrm{t}}$ y que

$$\frac{\partial}{\partial\theta_1} \log f_T(\widetilde{Z}_i;\theta) \equiv U_i \ ; \quad \frac{\partial}{\partial\theta_1} \log S_T(\overline{Z}_j;\theta) \equiv V_j \ ; \quad \frac{\partial}{\partial\theta_1} \log\left\{1 - S_T(\widehat{Z}_k;\theta)\right\} \equiv W_k,$$

$$\frac{\partial}{\partial\theta_2} \log f_T(\widetilde{Z}_i;\theta) \equiv U_i' \ ; \quad \frac{\partial}{\partial\theta_2} \log S_T(\overline{Z}_j;\theta) \equiv V_j' \ ; \quad \frac{\partial}{\partial\theta_2} \log\left\{1 - S_T(\widehat{Z}_k;\theta)\right\} \equiv W_k'.$$

Por tanto, resulta que

$$\frac{1}{\sqrt{N}} R^*(\theta) = \frac{1}{\sqrt{N}} \begin{pmatrix} R_1^*(\theta) \\ R_2^*(\theta) \end{pmatrix} = \frac{1}{\sqrt{N}} \begin{pmatrix} \displaystyle\sum_{i=1}^{N_0} U_i + \left(1 + \frac{N_{10}}{N_{11}}\right) \sum_{j=1}^{N_{11}} V_j + \sum_{k=1}^{N_2} W_k \\ \displaystyle\sum_{i=1}^{N_0} U_i' + \left(1 + \frac{N_{10}}{N_{11}}\right) \sum_{j=1}^{N_{11}} V_j' + \sum_{k=1}^{N_2} W_k' \end{pmatrix}.$$

Utilizando la notación anterior, se obtiene que

$$
\begin{aligned}
E\left[R_1^*(\theta)\right] &= E\left[E\left[\sum_{i=1}^{N_0} U_i + \left(1+\frac{N_{10}}{N_{11}}\right)\sum_{j=1}^{N_{11}} V_j + \sum_{k=1}^{N_2} W_k \mid N_0, N_{10}, N_{11}, N_2\right]\right] \\
&= E\left[E\left[\sum_{i=1}^{N_0} U_i \mid N_0\right]\right] + E\left[E\left[\left(1+\frac{N_{10}}{N_{11}}\right)\sum_{j=1}^{N_{11}} V_j \mid N_{10}, N_{11}\right]\right] \\
&\quad + E\left[E\left[\sum_{k=1}^{N_2} W_k \mid N_2\right]\right].
\end{aligned}
$$

Por consiguiente,

$$
\begin{aligned}
E\left[R_1^*(\theta)\right] &= E\left[N_0 E\left[U\right]\right] + E\left[N_1 E\left[V\right]\right] + E\left[N_2 E\left[W\right]\right] \\
&= N\left(p_0 E\left[U\right] + p_1 E\left[V\right] + p_2 E\left[W\right]\right).
\end{aligned}
$$

Luego, teniendo en cuenta que $p_0 E\left[U\right] + p_1 E\left[V\right] + p_2 E\left[W\right]$ vale

$$
\int_0^\infty \frac{\partial}{\partial\theta_1} \log f_T(z;\theta)\left\{S_Y(z) - S_X(z)\right\} f_T(z) dz + \int_0^\infty \frac{\partial}{\partial\theta_1} \log S_T(z;\theta) S_T(z) f_Y(z) dz
$$

$$
+ \int_0^\infty \frac{\partial}{\partial\theta_1} \log\left\{1 - S_T(z;\theta)\right\}\left\{1 - S_T(z)\right\} f_X(z) dz
$$

$$
= \int_0^\infty \frac{\partial}{\partial\theta_1} f_T(z;\theta)\left\{S_Y(z) - S_X(z)\right\} dz + \int_0^\infty \frac{\partial}{\partial\theta_1} S_T(z;\theta) f_Y(z) dz
$$

$$
+ \int_0^\infty \frac{\partial}{\partial\theta_1}\left\{1 - S_T(z;\theta)\right\} f_X(z) dz
$$

$$
= \frac{\partial}{\partial\theta_1}\int_0^\infty \Pr(X < z < Y) f_T(z) dz + \frac{\partial}{\partial\theta_1}\int_0^\infty \Pr(T > z) f_Y(z) dz
$$

$$
+ \frac{\partial}{\partial\theta_1}\int_0^\infty \Pr(z > T) f_X(z) dz
$$

$$
= \frac{\partial}{\partial\theta_1}\left\{p_0 + p_1 + p_2\right\} = 0,
$$

resulta que $E\left[R_1^*(\theta)\right] = 0$. Del mismo modo, se comprueba que $E\left[R_2^*(\theta)\right] = 0$. Por tanto, se verifica que la esperanza de $R^*(\theta)/\sqrt{N}$ es el vector cero.

Ahora, se plantea el cálculo de la matriz asintótica de varianzas-covarianzas del vector $R^*(\theta)/\sqrt{N}$.

En primer lugar, la varianza de $R_1^*(\theta)/\sqrt{N}$ es igual a

$$\frac{1}{N}Var\left(E\left[R_1^*(\theta) \mid N_0, N_{10}, N_{11}, N_2\right]\right) + \frac{1}{N}E\left[Var\left(R_1^*(\theta) \mid N_0, N_{10}, N_{11}, N_2\right)\right].$$

El primero de los términos anteriores vale

$$\begin{aligned}
&\frac{1}{N}Var\left(N_0E\left[U\right] + N_1E\left[V\right] + N_2E\left[W\right]\right) \\
&= \frac{1}{N}Var\left(N_0(E\left[U\right] - E\left[W\right]) + N_1(E\left[V\right] - E\left[W\right]) + NE\left[W\right]\right) \\
&= \frac{1}{N}(E\left[U\right] - E\left[W\right])^2 Var(N_0) + \frac{1}{N}(E\left[V\right] - E\left[W\right])^2 Var(N_1) \\
&\quad + \frac{2}{N}(E\left[U\right] - E\left[W\right])(E\left[V\right] - E\left[W\right])Cov(N_0, N_1) \\
&= p_0(1-p_0)(E\left[U\right] - E\left[W\right])^2 + p_1(1-p_1)(E\left[V\right] - E\left[W\right])^2 \\
&\quad -2p_0p_1(E\left[U\right] - E\left[W\right])(E\left[V\right] - E\left[W\right]) \\
&= p_0E\left[U\right]^2 + p_1E\left[V\right]^2 + p_2E\left[W\right]^2.
\end{aligned}$$

Nótese que en la última línea se ha utilizado que

$$p_0 + p_1 + p_2 = 1 \quad \text{y} \quad p_0E\left[U\right] + p_1E\left[V\right] + p_2E\left[W\right] = 0.$$

El segundo término es igual a

$$\begin{aligned}
&\frac{1}{N}\left\{E\left[N_0\right]Var\left(U\right) + E\left[N_{11}\left(1 + \frac{N_{10}}{N_{11}}\right)^2\right]Var\left(V\right) + E\left[N_2\right]Var\left(W\right)\right\} \\
&= \frac{1}{N}\left\{Np_0Var\left(U\right) + \frac{\left(Np_1\right)^2}{Np_1f^*}Var\left(V\right) + o_p(N) + Np_2Var\left(W\right)\right\},
\end{aligned}$$

donde $o_p(N)$ es una expresión que dividida por N converge en probabilidad a 0.

Luego,

$$
\begin{aligned}
\frac{1}{N} Var\left(R_1^*(\theta)\right) &= p_0 E\left[U\right]^2 + p_1 E\left[V\right]^2 + p_2 E\left[W\right]^2 + p_0 Var\left(U\right) + \frac{p_1}{f^*} Var\left(V\right) \\
&\quad + o_p(1) + p_2 Var\left(W\right) \\
&= p_0 E\left[U^2\right] + \frac{p_1}{f^*} E\left[V^2\right] + p_2 E\left[W^2\right] - \frac{p_1(1-f^*)}{f^*} E\left[V\right]^2 + o_p(1).
\end{aligned}
$$

En segundo lugar, se procede al cálculo de la covarianza

$$
Cov(R_1^*(\theta)/\sqrt{N}, R_2^*(\theta)/\sqrt{N}) = E\left[R_1^*(\theta) R_2^*(\theta)\right]/N.
$$

Al hallar $E\left[R_1^*(\theta) R_2^*(\theta)\right]$ aparecerán los nueve términos siguientes:

$$
\begin{aligned}
1)\ E\left[\sum_{i=1}^{N_0} U_i \sum_{i=1}^{N_0} U_i'\right] &= E\left[E\left[\sum_{i=1}^{N_0} U_i U_i' + \sum_{i \neq i'}^{N_0} U_i U_{i'}' \mid N_0\right]\right] \\
&= E\left[N_0 E[UU'] + N_0(N_0-1) E\left[U\right] E[U']\right] \\
&= E\left[N_0\right] E[UU'] + \left(E\left[N_0^2\right] - E\left[N_0\right]\right) E\left[U\right] E[U'] \\
&= N p_0 E[UU'] + N(N-1) p_0^2 E\left[U\right] E[U'],
\end{aligned}
$$

$$
\begin{aligned}
2)\ E\left[\left(1+\frac{N_{10}}{N_{11}}\right) \sum_{i=1}^{N_0} U_i \sum_{j=1}^{N_{11}} V_j'\right] &= E\left[E\left[\left(1+\frac{N_{10}}{N_{11}}\right) \sum_{i=1}^{N_0} \sum_{j=1}^{N_{11}} U_i V_j' \mid N_0, N_{10}, N_{11}\right]\right] \\
&= E\left[N_0 N_1 E\left[U\right] E[V']\right] \\
&= N(N-1) p_0 p_1 E\left[U\right] E[V'],
\end{aligned}
$$

$$
3)\ E\left[\sum_{i=1}^{N_0} U_i \sum_{k=1}^{N_2} W_k'\right] = N(N-1) p_0 p_2 E\left[U\right] E[W'],
$$

$$
4)\ E\left[\left(1+\frac{N_{10}}{N_{11}}\right) \sum_{i=1}^{N_0} U_i' \sum_{j=1}^{N_{11}} V_j\right] = N(N-1) p_0 p_1 E[U'] E\left[V\right].
$$

$$5)\ E\left[\left(1+\frac{N_{10}}{N_{11}}\right)^2\sum_{j=1}^{N_{11}}V_j\sum_{j=1}^{N_{11}}V_j'\right]$$

$$= E\left[N_{11}\left(1+\frac{N_{10}}{N_{11}}\right)^2 E[VV']\right] + E\left[N_{11}(N_{11}-1)\left(1+\frac{N_{10}}{N_{11}}\right)^2 E\left[V\right]E[V']\right]$$

$$= \left(\frac{Np_1}{f^*}+o_p(N)\right)E[VV']$$

$$+\left(Np_1(1-p_1)+N^2p_1^2-\frac{Np_1}{f^*}-o_p(N)\right)E\left[V\right]E[V'],$$

$$6)\ E\left[\left(1+\frac{N_{10}}{N_{11}}\right)\sum_{j=1}^{N_{11}}V_j\sum_{k=1}^{N_2}W_k'\right] = N(N-1)p_1p_2E\left[V\right]E[W'],$$

$$7)\ E\left[\sum_{k=1}^{N_2}W_k\sum_{i=1}^{N_0}U_i'\right] = N(N-1)p_0p_2E\left[W\right]E[U'],$$

$$8)\ E\left[\left(1+\frac{N_{10}}{N_{11}}\right)\sum_{k=1}^{N_2}W_k\sum_{j=1}^{N_{11}}V_j'\right] = N(N-1)p_1p_2E\left[W\right]E[V'],$$

$$9)\ E\left[\sum_{k=1}^{N_2}W_k\sum_{k=1}^{N_2}W_k'\right] = Np_2E[WW'] + N(N-1)p_2^2E\left[W\right]E[W'].$$

Obsérvese que la suma de los términos que incluyen N^2 se anula ya que

$$p_0E\left[U\right]+\ p_1E\left[V\right]\ +p_2E\left[W\right]=\ p_0E[U']+p_1E[V']+p_2E[W']=0.$$

Sumando los nueve términos anteriores y dividiendo por N, se obtiene que:

$$\begin{aligned} Cov(R_1^*(\theta)/\sqrt{N},R_2^*(\theta)/\sqrt{N}) \ &= \ p_0E[UU']+\frac{p_1}{f^*}E[VV']+p_2E[WW'] \\ &\quad -\frac{p_1(1-f^*)}{f^*}E\left[V\right]E[V']+o_p(1), \end{aligned}$$

con lo que finaliza la comprobación de la fórmula (2.6).

Veamos que la esperanza de $R^{**}(\theta)/\sqrt{N}$ es el vector cero y que su matriz de varianzas-covarianzas asintótica viene dada por la expresión (2.7).

En esta ocasión se verifica que

$$\frac{1}{\sqrt{N}} R^{**}(\theta) = \frac{1}{\sqrt{N}} \begin{pmatrix} R_1^{**}(\theta) \\ R_2^{**}(\theta) \end{pmatrix} = \frac{1}{\sqrt{N}} \begin{pmatrix} \sum_{i=1}^{N_0} U_i + \frac{1}{f^*} \sum_{j=1}^{N_{11}} V_j + \sum_{k=1}^{N_2} W_k \\ \\ \sum_{i=1}^{N_0} U_i' + \frac{1}{f^*} \sum_{j=1}^{N_{11}} V_j' + \sum_{k=1}^{N_2} W_k' \end{pmatrix}.$$

De ahí, se tiene que

$$\begin{aligned} E\left[R_1^{**}(\theta)\right] &= E\left[E\left[\sum_{i=1}^{N_0} U_i + \frac{1}{f^*}\sum_{j=1}^{N_{11}} V_j + \sum_{k=1}^{N_2} W_k \mid N_0, N_{10}, N_{11}, N_2\right]\right] \\ &= E\left[E\left[\sum_{i=1}^{N_0} U_i \mid N_0\right]\right] + E\left[\frac{1}{f^*} E\left[\sum_{j=1}^{N_{11}} V_j \mid N_{10}, N_{11}\right]\right] \\ &\quad + E\left[E\left[\sum_{k=1}^{N_2} W_k \mid N_2\right]\right]. \end{aligned}$$

Por consiguiente,

$$\begin{aligned} E\left[R_1^{**}(\theta)\right] &= E\left[N_0 E\left[U\right]\right] + \frac{1}{f^*} E\left[N_{11} E\left[V\right]\right] + E\left[N_2 E\left[W\right]\right] \\ &= N\left(p_0 E\left[U\right] + p_1 E\left[V\right] + p_2 E\left[W\right]\right) = 0. \end{aligned}$$

Análogamente, $E\left[R_2^{**}(\theta)\right] = 0$. Por tanto, se verifica que la esperanza de $R^{**}(\theta)/\sqrt{N}$ coincide con el vector cero.

En este punto, se plantea el cálculo de los elementos de la matriz de varianzas-covarianzas asintótica de $R^{**}(\theta)/\sqrt{N}$.

En primer lugar, la varianza de $R_1^{**}(\theta)/\sqrt{N}$ es igual a

$$\frac{1}{N}Var\left(E\left[R_1^{**}(\theta) \mid N_0, N_{10}, N_{11}, N_2\right]\right) + \frac{1}{N}E\left[Var\left(R_1^{**}(\theta) \mid N_0, N_{10}, N_{11}, N_2\right)\right].$$

El primero de los términos anteriores vale

$$\frac{1}{N}Var\left(N_0E\left[U\right] + \frac{1}{f^*}N_1E\left[V\right] + N_2E\left[W\right]\right)$$

$$= \frac{1}{N}\left(E[N_0^2]E\left[U\right]^2 + \frac{1}{f^{*2}}E[N_1^2]E\left[V\right]^2 + E[N_2^2]E\left[W\right]^2\right.$$

$$+\frac{2}{f^*}E[N_0N_{11}]E\left[U\right]E\left[V\right] + 2E[N_0N_2]E\left[U\right]E\left[W\right]$$

$$\left.+\frac{2}{f^{*2}}E[N_{11}N_2]E\left[V\right]E\left[W\right]\right)$$

$$= \left\{Np_0^2 + p_0(1-p_0)\right\}E\left[U\right]^2 + \frac{1}{f^{*2}}\left\{Np_1^2f^{*2} + p_1f^*(1-p_1f^*)\right\}E\left[V\right]^2$$

$$+\left\{Np_2^2 + p_2(1-p_2)\right\}E\left[W\right]^2 + 2(N-1)p_0p_1E\left[U\right]E\left[V\right]$$

$$+2(N-1)p_0p_2E\left[U\right]E\left[W\right] + 2(N-1)p_1p_2E\left[V\right]E\left[W\right]$$

$$= p_0E\left[U\right]^2 + \frac{p_1}{f^*}E\left[V\right]^2 + p_2E\left[W\right]^2.$$

El segundo término es igual a

$$\frac{1}{N}\left\{E\left[N_0\right]Var\left(U\right) + \frac{1}{f^{*2}}E\left[N_{11}\right]Var\left(V\right) + E\left[N_2\right]Var\left(W\right)\right\}$$

$$= p_0Var\left(U\right) + \frac{p_1}{f^*}Var\left(V\right) + p_2Var\left(W\right).$$

Luego,

$$\frac{1}{N}Var\left(R_1^{**}(\theta)\right) = p_0E[U^2] + \frac{p_1}{f^*}E\left[V^2\right] + p_2E\left[W^2\right].$$

En segundo lugar, se procede al cálculo de

$$Cov(R_1^{**}(\theta)/\sqrt{N}, R_2^{**}(\theta)/\sqrt{N}) = E\left[R_1^{**}(\theta)R_2^{**}(\theta)\right]/N.$$

En el cálculo de $E\left[R_1^{**}(\theta)R_2^{**}(\theta)\right]$ aparecerán los mismos nueve términos que se obtenían al hallar $E\left[R_1^{*}(\theta)R_2^{*}(\theta)\right]$ salvo el quinto que, en este caso, será:

$$\begin{aligned} 5)E\left[\frac{1}{f^{*2}}\sum_{j=1}^{N_{11}}V_j\sum_{j=1}^{N_{11}}V_j'\right] &= E\left[N_{11}\frac{1}{f^{*2}}E[VV'] + N_{11}(N_{11}-1)\frac{1}{f^{*2}}E\left[V\right]E[V']\right] \\ &= \frac{Np_1}{f^*}E[VV'] + N(N-1)p_1^2 + E\left[V\right]E[V']. \end{aligned}$$

Sumando esos nueve términos y dividiendo por N, resulta que:

$$Cov(R_1^{**}(\theta)/\sqrt{N}, R_2^{**}(\theta)/\sqrt{N}) = p_0E[UU'] + \frac{p_1}{f^*}E[VV'] + p_2E[WW'] + o_p(1).$$

Por consiguiente, queda demostrada la válidez de la expresión (2.7).

$\square$

Observaciones

- Nótese que

$$\sqrt{N}(\widehat{\theta}-\theta) \xrightarrow[N\to\infty]{d} N(0, -\Phi(\theta)^{-1}).$$

- Como θ^* es una estimador consistente de θ, el estimador semiparamétrico obtenido en Bravo et al. (1995a, 1995c) es consistente. Este estimador puede utilizarse como alternativa al puramente paramétrico $S_T(t;\theta^*)$. Resultados análogos se obtienen si se considera al estimador θ^{**}.

- Obsérvese que

$$E\left[h(\widetilde{Z};\theta)\right] = \frac{1}{p_0}\int_0^\infty h(z;\theta)\left\{S_Y(z)-S_X(z)\right\}f_T(z;\theta)dz,$$

$$E\left[h(\overline{Z};\theta)\right] = \frac{1}{p_1}\int_0^\infty h(z;\theta)S_T(z;\theta)f_Y(z)dz,$$

$$E\left[h(\widehat{Z};\theta)\right] = \frac{1}{p_2}\int_0^\infty h(z;\theta)\left\{1-S_T(z;\theta)\right\}f_X(z)dz.$$

- Aplicando la ley de los grandes números, es posible obtener estimaciones de las matrices Φ y Σ^* aproximando θ, p_i, $i = 0, 1, 2$, $E\left[h(\widetilde{Z};\theta)\right]$, $E\left[h(\overline{Z};\theta)\right]$ y $E\left[h(\widehat{Z};\theta)\right]$ por θ^*, N_i/N, $i = 0, 1, 2$, $\sum_{i=1}^{N_0} h(\widetilde{Z}_i;\theta^*)/N_0$, $\sum_{j=1}^{N_{11}} h(\overline{Z}_j;\theta^*)/N_{11}$ y $\sum_{i=1}^{N_2} h(\widehat{Z}_k;\theta^*)/N_2$, respectivamente. De forma similar, se pueden conseguir estimaciones de las matrices Φ y Σ^{**} intercambiando a θ^* por θ^{**}.

Corolario 2.1

Si la función $g(\cdot) = (g_1(\cdot), ...g_r(\cdot))^{\mathrm{t}}$ es diferenciable con continuidad en un entorno de θ, en el cual la matriz de derivadas parciales $D_g(\cdot) = ((\partial/\partial\theta_i)g_j(\cdot))_{kxr}$ es no singular, entonces

$$\sqrt{N}(g(\theta^*) - g(\theta)) \xrightarrow[N\to\infty]{d} N(0, D_g(\theta)^{\mathrm{t}}\Phi(\theta)^{-1}\Sigma^*(\theta, f^*)\Phi(\theta)^{-1}D_g(\theta)),$$

y

$$\sqrt{N}(g(\theta^{**}) - g(\theta)) \xrightarrow[N\to\infty]{d} N(0, D_g(\theta)^{\mathrm{t}}\Phi(\theta)^{-1}\Sigma^{**}(\theta, f^*)\Phi(\theta)^{-1}D_g(\theta)).$$

Por supuesto, $g(\theta^*)$ y $g(\theta^{**})$ son ambos estimadores consistentes de $g(\theta)$, y de máxima pseudo-verosimilitud L^* y L^{**} respectivamente.

Demostración

Mediante un desarrollo en serie de $g(\theta^*)$ en un entorno de θ, se obtiene que

$$g(\theta^*) = g(\theta) + D_g(\widetilde{\theta})^{\mathrm{t}}(\theta^* - \theta),$$

donde $\widetilde{\theta}$ está en el segmento que une θ^* y θ.

Como $\widetilde{\theta} - \theta = o_p(1)$, se cumple que $D_g(\widetilde{\theta}) = D_g(\theta) + o_p(1)$. Además, $\sqrt{N}(\theta^* - \theta)$ está acotado en probabilidad. Por consiguiente, se verifica que

$$\sqrt{N}\,(g(\theta^*) - g(\theta)) = D_g(\theta)^{\mathrm{t}}\sqrt{N}(\theta^* - \theta) + o_p(1),$$

y de ahí, como

$$\sqrt{N}(\theta^* - \theta) \xrightarrow[N\to\infty]{d} N(0, \Phi(\theta)^{-1}\Sigma^*(\theta, f^*)\Phi(\theta)^{-1}),$$

resulta que

$$\sqrt{N}(g(\theta^*) - g(\theta)) \xrightarrow[N\to\infty]{d} N(0, D_g(\theta)^{\mathrm{t}}\Phi(\theta)^{-1}\Sigma^*(\theta, f^*)\Phi(\theta)^{-1}D_g(\theta)).$$

La normalida asintótica de $g(\theta^{**})$ se demuestra de forma similar.

$\square$

Como caso particular, los estimadores de la supervivencia y de los percentiles son consistentes y asintóticamente normales, esto es, se verifica el corolario siguiente.

Corolario 2.2

Dado $\alpha \in (0,1)$, sea $t_\alpha(\cdot)$ el cuantil α de la distribución de T y supóngase que el vector columna de derivadas parciales $Q(\cdot) = (\partial/\partial\theta)t_\alpha(\cdot)$ es no nulo y continuo en un entorno de θ, entonces $t_\alpha(\theta^*)$ y $t_\alpha(\theta^{**})$ son estimadores consistentes, y de máxima pseudo-verosimilitud L^* y L^{**} respectivamente, de $t_\alpha(\theta)$. Además,

$$\sqrt{N}(t_\alpha(\theta^*) - t_\alpha(\theta)) \xrightarrow[N\to\infty]{d} N(0, Q(\theta)^{\mathrm{t}}\Phi(\theta)^{-1}\Sigma^*(\theta, f^*)\Phi(\theta)^{-1}Q(\theta)),$$

y

$$\sqrt{N}(t_\alpha(\theta^{**}) - t_\alpha(\theta)) \xrightarrow[N\to\infty]{d} N(0, Q(\theta)^{\mathrm{t}}\Phi(\theta)^{-1}\Sigma^{**}(\theta, f^*)\Phi(\theta)^{-1}Q(\theta)).$$

Si el vector $G_t(\cdot) = (\partial/\partial\theta)S_T(t;\cdot)$ es no nulo y continuo en un entorno de θ, entonces $S_T(t;\theta^*)$ y $S_T(t;\theta^{**})$ son estimadores consistentes, y de máxima pseudo-verosimilitud L^* y L^{**} respectivamente, de $S_T(t;\theta)$. Además,

$$\sqrt{N}\,(S_T(t;\theta^*) - S_T(t;\theta)) \xrightarrow[N\to\infty]{d} N(0, G_t(\theta)^{\mathrm{t}}\Phi(\theta)^{-1}\Sigma^*(\theta, f^*)\Phi(\theta)^{-1}G_t(\theta)),$$

y

$$\sqrt{N}\,(S_T(t;\theta^{**}) - S_T(t;\theta)) \xrightarrow[N\to\infty]{d} N(0, G_t(\theta)^{\mathrm{t}}\Phi(\theta)^{-1}\Sigma^{**}(\theta, f^*)\Phi(\theta)^{-1}G_t(\theta)).$$

Observación

- Dada la normalidad asintótica de

$$\sqrt{N}\,(S_T(t;\theta^*) - S_T(t;\theta)) \quad \text{y} \quad \sqrt{N}(t_\alpha(\theta^*) - t_\alpha(\theta)),$$

es posible construir intervalos de confianza asintóticos para $S_T(t;\theta^*)$ y $t_\alpha(\theta^*)$, y efectuar contrastes de hipótesis. Similares resultados se obtienen si intercambiamos a θ^* por θ^{**}. Hemos de precisar que, en general, es preferible utilizar la normalidad asintótica de $\log\left[-\log S_T(t;\theta^*)\right]$ a la de $S_T(t;\theta^*)$.

En los corolarios 2.3 y 2.4 se analizan dos casos particulares.

Corolario 2.3

Si T_i, Y_i y X_i, para $i = 1, .., N$, son variables aleatorias exponenciales con tasas de fallo λ, μ y γ respectivamente, entonces $\sqrt{N}(\lambda^* - \lambda)$ converge en distribución a una variable aleatoria normal con media 0 y varianza:

$$va(\sqrt{N}\lambda^*) = \frac{\lambda^2 \left\{ \dfrac{1}{1+\xi} + \dfrac{\xi(1-f^*)}{f^*(1+\xi)^3} - \dfrac{1}{1+\eta} + 2\eta\zeta(3, 1+\eta) \right\}}{\left\{ \dfrac{1}{1+\xi} - \dfrac{1}{1+\eta} + 2\eta\zeta(3, 1+\eta) \right\}^2}, \tag{2.9}$$

donde $\xi = \mu/\lambda$, $\eta = \gamma/\lambda$ y ζ es la función Zeta de Riemann generalizada definida por

$$\zeta(z, q) = \frac{1}{\Gamma(z)} \int_0^\infty \frac{t^{z-1} e^{-qt}}{1 - e^{-t}} dt.$$

Asimismo, la distribución de $\sqrt{N}(\lambda^{**} - \lambda)$ es asintóticamente normal con media 0 y varianza:

$$va(\sqrt{N}\lambda^{**}) = \frac{\lambda^2 \left\{ \dfrac{1}{1+\xi} + \dfrac{2\xi(1-f^*)}{f^*(1+\xi)^3} - \dfrac{1}{1+\eta} + 2\eta\zeta(3, 1+\eta) \right\}}{\left\{ \dfrac{1}{1+\xi} - \dfrac{1}{1+\eta} + 2\eta\zeta(3, 1+\eta) \right\}^2}. \tag{2.10}$$

Demostración

En primer lugar, se verifica que

$$p_1 = \Pr(T > Y) = \int_0^\infty S_T(y) f_Y(y) dy = \int_0^\infty e^{-\lambda y} \mu e^{-\mu y} dt = \frac{\mu}{\lambda + \mu},$$

$$p_2 = \Pr(T < X) = \int_0^\infty S_X(t) f_T(t) dt = \int_0^\infty e^{-\gamma t} \lambda e^{-\lambda t} dt = \frac{\lambda}{\lambda + \gamma},$$

y

$$p_0 = \Pr(X \leq T \leq Y) = 1 - p_1 - p_2.$$

Sustituyendo los valores de p_j, $j = 0, 1, 2$, y las expresiones:

$$\frac{\partial}{\partial \lambda} \log f_T(\widetilde{Z}; \lambda) = \frac{1}{\lambda} - \widetilde{Z}, \quad \frac{\partial^2}{\partial \lambda^2} \log f_T(\widetilde{Z}; \lambda) = -\frac{1}{\lambda^2}, \quad \frac{\partial}{\partial \lambda} \log S_T(\overline{Z}; \lambda) = -\overline{Z},$$

$$\frac{\partial^2}{\partial \lambda^2} \log S_T(\overline{Z}; \lambda) = 0, \quad \frac{\partial}{\partial \lambda} \log \left\{1 - S_T(\widehat{Z}; \lambda)\right\} = \frac{\widehat{Z} \exp(-\lambda \widehat{Z})}{1 - \exp(-\lambda \widehat{Z})} \quad \text{y}$$

$$\frac{\partial^2}{\partial \lambda^2} \log \left\{1 - S_T(\widehat{Z}; \lambda)\right\} = \frac{-\widehat{Z}^2 \exp(-\lambda \widehat{Z})}{\left\{1 - \exp(-\lambda \widehat{Z})\right\}^2}$$ en (2.5), (2.6) y (2.7), se obtiene (2.9) y (2.10) después de algunos cálculos que se proceden a exponer con detalle.

Como

$$\begin{aligned} p_0 E\left[\frac{\partial^2}{\partial \lambda^2} \log f_T(\widetilde{Z}; \lambda)\right] &= -\int_0^\infty \frac{1}{\lambda^2} \left(e^{-\mu z} - e^{-\gamma z}\right) \lambda e^{-\lambda z} dz \\ &= -\frac{1}{\lambda} \left(\frac{1}{\lambda + \mu} - \frac{1}{\lambda + \gamma}\right), \end{aligned}$$

$$p_1 E\left[\frac{\partial^2}{\partial \lambda^2} \log S_T(\overline{Z}; \lambda)\right] = 0,$$

$$\begin{aligned} p_2 E\left[\frac{\partial^2}{\partial \lambda^2} \log \left\{1 - S_T(\widehat{Z}; \lambda)\right\}\right] &= \int_0^\infty \frac{-z^2 e^{-\lambda z}}{\{1 - e^{-\lambda z}\}^2} \gamma e^{-\gamma z} (1 - e^{-\lambda z}) dz \\ &= -\gamma \int_0^\infty \frac{z^2 e^{-(\lambda + \gamma) z}}{1 - e^{-\lambda z}} dz \\ &= -\frac{2\gamma}{\lambda^3} \zeta(3, 1 + \gamma/\lambda), \end{aligned}$$

se obtiene, en este caso, que

$$\Phi(\lambda) = -\frac{1}{\lambda^2} \left(\frac{1}{1 + \xi} - \frac{1}{1 + \eta} + 2\eta \zeta(3, 1 + \eta)\right),$$

donde $\xi = \mu/\lambda$, $\eta = \gamma/\lambda$ y ζ es la función Zeta de Riemann generalizada.

Ahora, se procede a hallar la matriz $\Sigma^*(\lambda, f^*)$. Resulta que

$$p_0 E\left[\left\{\frac{\partial}{\partial\lambda}\log f_T(\widetilde{Z};\lambda)\right\}^2\right] = \int_0^\infty (\frac{1}{\lambda} - z)^2 \lambda e^{-\lambda z}(e^{-\mu z} - e^{-\gamma z})dz$$
$$= \frac{\lambda^2+\mu^2}{\lambda(\lambda+\mu)^3} - \frac{\lambda^2+\gamma^2}{\lambda(\lambda+\gamma)^3},$$

$$\frac{p_1}{f^*}E\left[\left\{\frac{\partial}{\partial\lambda}\log S_T(\overline{Z};\lambda)\right\}^2\right] = \frac{1}{f^*}\int_0^\infty (-z)^2 \mu e^{-\mu z}e^{-\lambda z}dz = \frac{2\mu}{f^*(\lambda+\mu)^3},$$

$$p_2 E\left[\left\{\frac{\partial}{\partial\lambda}\log\left\{1 - S_T(\widehat{Z};\lambda)\right\}\right\}^2\right] = \int_0^\infty \frac{z^2 e^{-2\lambda z}}{\left\{1-e^{-\lambda z}\right\}^2}\gamma e^{-\gamma z}(1-e^{-\lambda z})dz$$
$$= \gamma\int_0^\infty \frac{z^2 e^{-(2\lambda+\gamma)z}}{1-e^{-\lambda z}}dz$$
$$= \frac{2\gamma}{\lambda^3}\zeta(3, 2+\gamma/\lambda),$$

y, puesto que $E\left[\frac{\partial}{\partial\lambda}\log S_T(\overline{Z};\lambda)\right] = -\frac{1}{p_1}\int_0^\infty z\mu e^{-\mu z}e^{-\lambda z}dz = -\frac{1}{\lambda+\mu}$,

$$\frac{p_1(1-f^*)}{f^*}E\left[\frac{\partial}{\partial\lambda}\log S_T(\overline{Z};\lambda)\right]^2 = \frac{(1-f^*)\mu}{f^*\left(\lambda+\mu\right)^3}.$$

Considerando que $\xi = \mu/\lambda$, $\eta = \gamma/\lambda$ y $\zeta(3, 2+\eta) = \zeta(3, 1+\eta) - (1+\eta)^{-3}$, y utilizando la fórmula (2.6), se verifica que

$$\Sigma^*(\lambda, f^*) = \frac{1}{\lambda^2}\left\{\frac{1}{1+\xi} + \frac{\xi(1-f^*)}{f^*(1+\xi)^3} - \frac{1}{1+\eta} + 2\eta\zeta(3, 1+\eta)\right\}.$$

Por tanto, $\Phi(\lambda)^{-1}\Sigma^*(\lambda, f^*)\Phi(\lambda)^{-1}$ coincide con la expresión (2.9).

De forma análoga, se cumple que

$$\Sigma^{**}(\lambda, f^*) = \frac{1}{\lambda^2}\left\{\frac{1}{1+\xi} + \frac{2\xi(1-f^*)}{f^*(1+\xi)^3} - \frac{1}{1+\eta} + 2\eta\zeta(3, 1+\eta)\right\},$$

por lo cual, $\Phi(\lambda)^{-1}\Sigma^{**}(\lambda, f^*)\Phi(\lambda)^{-1}$ es idéntica a la expresión (2.10).

□

Observación

- Nótese que, si la variable T es exponencial, se verifica que:

  Para todo $\alpha \in (0,1)$,

$$\sqrt{N}(t_\alpha(\lambda^*) - t_\alpha(\lambda)) \xrightarrow[N\to\infty]{d} N(0, \log^2(1-\alpha)va(\sqrt{N}\lambda^*)/\lambda^4),$$

$$\sqrt{N}(t_\alpha(\lambda^{**}) - t_\alpha(\lambda)) \xrightarrow[N\to\infty]{d} N(0, \log^2(1-\alpha)va(\sqrt{N}\lambda^{**})/\lambda^4).$$

  Para todo $t > 0$,

$$\sqrt{N}\left(S_T(t;\lambda^*) - S_T(t;\lambda)\right) \xrightarrow[N\to\infty]{d} N(0, t^2\exp(-2\lambda t)va(\sqrt{N}\lambda^*)),$$

$$\sqrt{N}\left(S_T(t;\lambda^{**}) - S_T(t;\lambda)\right) \xrightarrow[N\to\infty]{d} N(0, t^2\exp(-2\lambda t)va(\sqrt{N}\lambda^{**})),$$

  de donde

$$\sqrt{N}\left(\log\left[-\log S_T(t;\lambda^*)\right] - \log\left[-\log S_T(t;\lambda)\right]\right) \xrightarrow[N\to\infty]{d} N(0, va(\sqrt{N}\lambda^*)/\lambda^2),$$

$$\sqrt{N}\left(\log\left[-\log S_T(t;\lambda^{**})\right] - \log\left[-\log S_T(t;\lambda)\right]\right) \xrightarrow[N\to\infty]{d} N(0, va(\sqrt{N}\lambda^{**})/\lambda^2).$$

Basado en L, la cota inferior de Cramér-Rao de $\sqrt{N}\lambda$, $CR(\sqrt{N}\lambda)$, para el caso exponencial es

$$CR(\sqrt{N}\lambda) = \frac{\lambda^2}{\dfrac{1}{1+\xi} + \dfrac{\xi(f^*-1)}{(1+\xi)^3} - \dfrac{1}{1+\eta} + 2\eta\zeta(3, 1+\eta)}.$$

Obviamente, $va(\lambda^*)$ coincide con $CR(\lambda)$ cuando $f^* = 1$ ya que el estimador de λ que maximiza la verosimilitud es el mejor estimador asintóticamente normal y su varianza asintótica alcanza la cota inferior de Cramér-Rao.

La tabla 2.1 muestra los valores de $CR(\sqrt{N}\lambda)/\lambda^2$ y $va(\sqrt{N}\lambda^*)/\lambda^2$ para $f^* = 0.1, .., 1.0$ cuando el nivel de censura es del 10%, 25% y 50%. También se presentan los porcentajes de eficiencia $ef(\lambda^*) = 100 \times CR(\lambda)/va(\lambda^*)$.

censura	f^*	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
10%	va	1.503	1.253	1.169	1.128	1.103	1.086	1.074	1.065	1.058	1.053
	CR	1.100	1.094	1.088	1.083	1.078	1.073	1.068	1.063	1.058	1.053
	ef	73.18	87.35	93.12	96.09	97.79	98.81	99.42	99.77	99.94	100.0
25%	va	2.272	1.646	1.437	1.332	1.270	1.228	1.198	1.176	1.158	1.144
	CR	1.270	1.254	1.239	1.225	1.211	1.197	1.183	1.170	1.157	1.144
	ef	55.87	76.22	86.26	91.93	95.35	97.47	98.76	99.51	99.89	100.0
50%	va	3.661	2.378	1.950	1.736	1.607	1.522	1.461	1.415	1.380	1.351
	CR	1.630	1.593	1.558	1.525	1.493	1.462	1.433	1.405	1.377	1.351
	ef	44.51	67.00	79.91	87.83	92.85	96.06	98.06	99.24	99.83	100.0

Tabla 2.1 : Valores de $va(\sqrt{N}\lambda^*)/\lambda^2$, $CR(\sqrt{N}\lambda)/\lambda^2$ y $ef(\lambda^*)$ en el caso exponencial.

Obsérvese que, como se esperaba, los porcentajes de eficiencia crecen cuando aumenta la proporción de seguimiento, f^*, y cuando disminuye el nivel de censura. Nótese también que se necesitan bajas proporciones de seguimiento para alcanzar buenos porcentajes de eficiencia. Por ejemplo, para obtener un 90% de eficiencia, una proporción de seguimiento de 0.2371, 0.3598 y 0.4378 es suficiente para un nivel de censura del 10%, 25% y 50% respectivamente.

La tabla 2.2 presenta los valores de $va(\sqrt{N}\lambda^{**})/\lambda^2$, $CR(\sqrt{N}\lambda)/\lambda^2$ y $ef(\lambda^{**})$ para $f^* = 0.1, .., 1.0$ cuando el nivel de censura es del 10%, 25% y 50%. Obsérvese que, también en esta ocasión, $va(\lambda^{**}) = CR(\lambda)$ cuando $f^* = 1$, pero que la eficiencia de λ^{**} es inferior a la de λ^* en los demás casos.

censura	f^*	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
10%	va	1.953	1.453	1.286	1.203	1.153	1.1196	1.096	1.078	1.064	1.053
	CR	1.100	1.094	1.088	1.083	1.078	1.073	1.068	1.063	1.058	1.053
	ef	56.31	75.32	84.67	90.09	93.55	95.87	97.48	98.62	99.43	100.0
25%	va	3.400	2.147	1.729	1.520	1.395	1.311	1.252	1.207	1.172	1.144
	CR	1.270	1.254	1.239	1.225	1.211	1.197	1.183	1.170	1.157	1.144
	ef	37.34	58.42	71.67	80.56	86.78	91.26	94.52	96.93	98.71	100.0
50%	va	5.972	3.405	2.549	2.121	1.865	1.693	1.571	1.479	1.408	1.351
	CR	1.630	1.593	1.558	1.525	1.493	1.462	1.433	1.405	1.377	1.351
	ef	27.29	46.80	61.13	71.89	80.07	86.35	91.19	94.93	97.81	100.0

Tabla 2.2 : Valores de $va(\sqrt{N}\lambda^{**})/\lambda^2$, $CR(\sqrt{N}\lambda)/\lambda^2$ y $ef(\lambda^{**})$ en el caso exponencial.

Corolario 2.4

Si la variable T sigue una distribución exponencial de parámetro λ y la ventana de observación $[X,Y]$ es el intervalo de tiempo que transcurre entre el r-ésimo y el s-ésimo ($r<s$) eventos en un proceso homogéneo de Poisson de parámetro μ, entonces $\sqrt{N}(\lambda^*-\lambda)$ converge en distribución a una variable aleatoria normal con media 0 y varianza $va(\sqrt{N}\lambda^*)$ dada por:

$$\frac{\lambda^2\left\{\left(\frac{\xi}{1+\xi}\right)^r-\left(\frac{\xi}{1+\xi}\right)^s+\frac{s(1-f^*)\xi^s}{f^*(1+\xi)^{s+2}}+r(r+1)\xi^r\zeta(r+2,\xi+1)\right\}}{\left\{\left(\frac{\xi}{1+\xi}\right)^r-\left(\frac{\xi}{1+\xi}\right)^s+r(r+1)\xi^r\zeta(r+2,\xi+1)\right\}^2}, \qquad (2.11)$$

donde $\xi=\mu/\lambda$ y ζ es la función Zeta de Riemann generalizada.

De forma análoga, la distribución de $\sqrt{N}(\lambda^{**}-\lambda)$ es asintóticamente normal con media 0 y varianza $va(\sqrt{N}\lambda^{**})$ dada por:

$$\frac{\lambda^2\left\{\left(\frac{\xi}{1+\xi}\right)^r-\left(\frac{\xi}{1+\xi}\right)^s+\frac{s\,(s+1)\,(1-f^*)\xi^s}{f^*(1+\xi)^{s+2}}+r(r+1)\xi^r\zeta(r+2,\xi+1)\right\}}{\left\{\left(\frac{\xi}{1+\xi}\right)^r-\left(\frac{\xi}{1+\xi}\right)^s+r(r+1)\xi^r\zeta(r+2,\xi+1)\right\}^2}. \qquad (2.12)$$

Demostración

Se supone que existen $E_1, ..., E_s$, variables aleatorias independientes e idénticamente distribuidas con función de supervivencia común $S_E(t) = \exp(-\mu t)$ para $t > 0$, tal que $X = \sum_{j=1}^{r} E_j$ e $Y = \sum_{j=1}^{s} E_j$. Por tanto, X e Y siguen sendas distribuciones gammas $G(r, \mu)$ y $G(s, \mu)$ respectivamente y su función de densidad conjunta viene dada por:

$$f_{(X,Y)}(x, y) = \begin{cases} \dfrac{\mu^r}{(r-1)!(s-r)!} \exp(-\mu x) x^{r-1}(y-x)^{s-r-1} & si\ 0 < x < y, \\ 0 & en\ otro\ caso. \end{cases}$$

Luego, se cumple que

$$p_1 = \int_0^\infty S_T(y) f_Y(y) dy = \int_0^\infty e^{-\lambda y} \frac{\mu^s}{\Gamma(s)} e^{-\mu y} y^{s-1} dt = \left(\frac{\mu}{\lambda+\mu}\right)^s,$$

$$p_2 = \int_0^\infty S_X(t) f_T(t) dt = \sum_{i=1}^{r-1} \int_0^\infty e^{-\mu t} \frac{(\mu t)^i}{i!} \lambda e^{-\lambda t} dt = 1 - \left(\frac{\mu}{\lambda+\mu}\right)^r$$

y

$$p_0 = 1 - p_1 - p_2 = \left(\frac{\mu}{\lambda+\mu}\right)^r - \left(\frac{\mu}{\lambda+\mu}\right)^s.$$

Teniendo en cuenta las fórmulas anteriores y el teorema 2.2, se obtienen las expresiones (2.11) y (2.12) a partir de los cálculos que se especifican a continuación.

Como

$$p_0 E\left[\frac{\partial^2}{\partial \lambda^2} \log f_T(\tilde{Z}; \lambda)\right] = -E\left[\frac{1}{\lambda^2} \mid X < T < Y\right] = \frac{1}{\lambda^2}\left\{\left(\frac{\mu}{\lambda+\mu}\right)^s - \left(\frac{\mu}{\lambda+\mu}\right)^r\right\},$$

$$p_1 E\left[\frac{\partial^2}{\partial \lambda^2} \log S_T(\overline{Z}; \lambda)\right] = 0,$$

y, utilizando la función Zeta de Riemann generalizada ζ,

$$p_2 E\left[\frac{\partial^2}{\partial\lambda^2}\log\left\{1 - S_T(\widehat{Z};\lambda)\right\}\right] = -\int_0^\infty \frac{z^2 e^{-\lambda z}}{\{1-e^{-\lambda z}\}^2}\frac{\mu^r}{(r-1)!}e^{-\mu z}z^{r-1}(1-e^{-\lambda z})dz$$

$$= -\frac{\mu^r}{(r-1)!}\int_0^\infty \frac{z^{r+1}e^{-(\lambda+\mu)z}}{1-e^{-\lambda z}}dz$$

$$= -\frac{r(r+1)\mu^r}{\lambda^{r+2}}\zeta(r+2, 1+\mu/\lambda),$$

se obtiene que, en esta ocasión,

$$\Phi(\lambda) = -\frac{1}{\lambda^2}\left\{\left(\frac{\xi}{1+\xi}\right)^r - \left(\frac{\xi}{1+\xi}\right)^s + r(r+1)\xi^r\zeta(r+2,\xi+1)\right\},$$

donde $\xi = \mu/\lambda$.

Ahora, se plantea el cálculo de la matriz $\Sigma^*(\lambda, f^*)$. Resulta que

$$p_0 E\left[\left\{\frac{\partial}{\partial\lambda}\log f_T(\widetilde{Z};\lambda)\right\}^2\right]$$

$$= \int_0^\infty (\frac{1}{\lambda} - z)^2\lambda e^{-\lambda z}(\sum_{i=1}^s e^{-\mu z}\frac{(\mu z)^i}{i!} - \sum_{i=1}^r e^{-\mu z}\frac{(\mu z)^i}{i!})dz$$

$$= \frac{1}{\lambda^2}\left\{\left(\frac{\mu}{\lambda+\mu}\right)^r - \left(\frac{\mu}{\lambda+\mu}\right)^s\right\} + r(r+1)\frac{\mu^r}{(\lambda+\mu)^{r+2}} - s(s+1)\frac{\mu^s}{(\lambda+\mu)^{s+2}}$$

$$= \frac{1}{\lambda^2}\left\{\left(\frac{\xi}{1+\xi}\right)^r - \left(\frac{\xi}{1+\xi}\right)^s + r(r+1)\frac{\xi^r}{(1+\xi)^{r+2}} - s(s+1)\frac{\xi^s}{(1+\xi)^{s+2}}\right\},$$

$$\frac{p_1}{f^*}E\left[\left\{\frac{\partial}{\partial\lambda}\log S_T(\overline{Z};\lambda)\right\}^2\right] = \frac{1}{f^*}\int_0^\infty (-z)^2 e^{-\lambda z}\frac{\mu^s}{(s-1)!}e^{-\mu z}z^{s-1}dz = \frac{s(s+1)\mu^s}{f^*(\lambda+\mu)^{s+2}},$$

$$
\begin{aligned}
p_2 E\left[\left\{\frac{\partial}{\partial\lambda}\log\left\{1-S_T(\widehat{Z};\lambda)\right\}\right\}^2\right] &= \int_0^\infty \frac{z^2 e^{-2\lambda z}}{1-e^{-\lambda z}}\frac{\mu^s}{(s-1)!}e^{-\mu z}z^{s-1}dz \\
&= \frac{\mu^r}{(r-1)!}\int_0^\infty \frac{z^{r+1}e^{-(2\lambda+\mu)z}}{1-e^{-\lambda z}}dz \\
&= \frac{r(r+1)\mu^r}{\lambda^{r+2}}\zeta(r+2,2+\mu/\lambda)
\end{aligned}
$$

y

$$
\frac{p_1(1-f^*)}{f^*}E\left[\frac{\partial}{\partial\lambda}\log S_T(\overline{Z};\lambda)\right]^2 = \frac{s^2(1-f^*)\mu^s}{f^*\,(\lambda+\mu)^{s+2}},
$$

ya que

$$
E\left[\frac{\partial}{\partial\lambda}\log S_T(\overline{Z};\lambda)\right] = -\frac{1}{p_1}\int_0^\infty z e^{-\lambda z}\frac{\mu^s}{(s-1)!}e^{-\mu z}z^{s-1}dz = \frac{s}{\lambda+\mu}.
$$

Teniendo en cuenta que $\zeta(r+2,2+\xi)=\zeta(r+2,1+\xi)-(1+\xi)^{-(r+2)}$ y $\xi=\mu/\lambda$, se obtiene, a partir de la fórmula (2.6), que $\Sigma^*(\lambda,f^*)$ es igual a

$$
\frac{1}{\lambda^2}\left\{\left(\frac{\xi}{1+\xi}\right)^r - \left(\frac{\xi}{1+\xi}\right)^s + \frac{s(1-f^*)\xi^s}{f^*(1+\xi)^{s+2}} + r(r+1)\xi^r\zeta(r+2,\xi+1)\right\}.
$$

Por tanto, $\Phi(\lambda)^{-1}\Sigma^*(\lambda,f^*)\Phi(\lambda)^{-1}$ coincide con la expresión (2.11).

Del mismo modo, se cumple que $\Sigma^{**}(\lambda,f^*)$ es igual a

$$
\frac{1}{\lambda^2}\left\{\left(\frac{\xi}{1+\xi}\right)^r - \left(\frac{\xi}{1+\xi}\right)^s + \frac{s(s+1)(1-f^*)\xi^s}{f^*(1+\xi)^{s+2}} + r(r+1)\xi^r\zeta(r+2,\xi+1)\right\},
$$

por lo cual, $\Phi(\lambda)^{-1}\Sigma^{**}(\lambda,f^*)\Phi(\lambda)^{-1}$ es idéntica a la expresión (2.12).

□

En este caso, la cota inferior de Cramér-Rao de $\sqrt{N}\lambda$ basada en L es

$$CR(\sqrt{N}\lambda) = \frac{\lambda^2}{\left(\frac{\xi}{1+\xi}\right)^r - \left(\frac{\xi}{1+\xi}\right)^s + \frac{s(1-f^*)\xi^s}{(1+\xi)^{s+2}} + r(r+1)\xi^r\zeta(r+2,\xi+1)}.$$

La tabla 2.3 muestra los valores de $va(\sqrt{N}\lambda^*)/\lambda^2$, $CR(\sqrt{N}\lambda)/\lambda^2$ y $ef(\lambda^*)$, para $f^* = 0.1, .., 1.0$ y varios niveles de censura, cuando $r = 1$ (*Caso I*).

p_1	p_2	f^*	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
0.051	0.050	va	1.128	1.087	1.073	1.066	1.062	1.059	1.057	1.056	1.055	1.054
		CR	1.061	1.061	1.060	1.059	1.058	1.057	1.056	1.056	1.055	1.054
		ef	94.10	97.58	98.75	99.31	99.61	99.79	99.90	99.96	99.99	100.0
0.118	0.125	va	1.478	1.287	1.224	1.192	1.173	1.161	1.152	1.145	1.140	1.135
		CR	1.171	1.167	1.163	1.159	1.155	1.151	1.147	1.143	1.139	1.135
		ef	79.22	90.61	94.98	97.17	98.41	99.14	99.59	99.84	99.96	100.0
0.237	0.250	va	2.506	1.851	1.634	1.525	1.459	1.416	1.384	1.361	1.343	1.328
		CR	1.458	1.442	1.427	1.412	1.4397	1.383	1.369	1.355	1.341	1.328
		ef	58.16	77.87	87.33	92.60	95.75	97.69	98.87	99.56	99.90	100.0

Tabla 2.3 : Valores de $va(\sqrt{N}\lambda^*)/\lambda^2$, $CR(\sqrt{N}\lambda)/\lambda^2$ y $ef(\lambda^*)$ en el *Caso I*.

Se observan conclusiones similares a la expuestas en el caso exponencial. Asimismo, se necesitan bajas proporciones de seguimiento para alcanzar buenos porcentajes de eficiencia. Por ejemplo, para obtener un 90% de eficiencia es suficiente una proporción de seguimiento de 0.0614, 0.1909 y 0.3440 cuando el nivel de censura es del 10.1%, 24.3% y 48.7% respectivamente.

p_1	p_2	f^*	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
0.051	0.050	va	5.419	2.994	2.186	1.781	1.539	1.377	1.262	1.175	1.108	1.054
		CR	1.061	1.061	1.060	1.059	1.058	1.057	1.056	1.056	1.055	1.054
		ef	19.58	35.42	48.48	59.44	68.75	76.76	83.72	89.82	95.21	100.0
0.118	0.125	va	6.956	3.722	2.644	2.105	1.782	1.566	1.412	1.297	1.207	1.135
		CR	1.171	1.167	1.163	1.159	1.155	1.151	1.147	1.143	1.139	1.135
		ef	16.83	31.34	43.96	55.03	64.79	73.46	81.19	88.12	94.36	100.0
0.237	0.250	va	8.394	4.469	3.160	2.506	2.113	1.852	1.665	1.525	1.416	1.328
		CR	1.458	1.442	1.427	1.412	1.4397	1.383	1.369	1.355	1.341	1.328
		ef	17.37	32.27	45.15	56.34	66.11	74.68	82.22	88.88	94.77	100.0

Tabla 2.4 : Valores de $va(\sqrt{N}\lambda^{**})/\lambda^2$, $CR(\sqrt{N}\lambda)/\lambda^2$ y $ef(\lambda^{**})$ en el *Caso I*.

Los valores de $va(\sqrt{N}\lambda^{**})/\lambda^2$, $CR(\sqrt{N}\lambda)/\lambda^2$ y $ef(\lambda^{**}) = 100\times CR(\lambda)/va(\lambda^{**})$ para $f^* = 0.1, .., 1.0$ y varios niveles de censura se presentan en la tabla 2.4. Al igual que ocurría anteriormente, $va(\lambda^*)$ y $va(\lambda^{**})$ coinciden con $CR(\lambda)$ sólo cuando $f^* = 1$; en los demás casos, la eficiencia de λ^{**} es inferior a la de λ^*.

2.4 Experiencia computacional

En esta sección se presentan los resultados de un estudio computacional diseñado para investigar la eficacia del estimador de θ de máxima pseudo-verosimilitud, θ^*, y de los percentiles estimados de la distribución de T, $t^*_{0.10}$, $t^*_{0.50}$ y $t^*_{0.90}$. Se analiza el efecto de la proporción de seguimiento, f^*, sobre el comportamiento de los estimadores precedentes, según el nivel de censura y el tamaño muestral, en dos casos particulares: *Caso I* y *Caso II*.

El *sesgo*, $E\left[\theta^*\right]-\theta$, y el error cuadrático medio (ecm), $E\left[(\theta^* - \theta)^2\right]$, se estiman mediante los estadísticos

$$\frac{1}{k}\sum_{i=1}^{k}\theta^*_{(i)} - \theta \quad \text{y} \quad \frac{1}{k}\sum_{i=1}^{k}\left(\theta^*_{(i)} - \theta\right)^2$$

respectivamente, donde k es el número de muestras que se han generado y $\theta^*_{(i)}$ es el valor de θ^* para la i-ésima muestra de la simulación.

El estudio consiste en la simulación de 5000 muestras de tamaños $25, 50$ y 100 para diferentes niveles de censura (aprox. 10%, 25% y 50%), mientras que f^* varía de 0.1 a 1.0 con incrementos de 0.3.

En primer lugar, se considera el *Caso I*, esto es, la variable T sigue una distribución exponencial de parámetro λ y la ventana de observación es el intervalo de tiempo que transcurre entre el primer y el s-ésimo eventos en un proceso homogéneo de Poisson de parámetro μ. De este modo, $T \sim Exp(\lambda)$, $X \sim G(1,\mu)$ e $Y \sim G(s,\mu)$, donde por simplicidad λ vale 1, el parámetro μ se obtiene resolviendo la ecuación

$$p_2 = \Pr(T < X) = \int_0^\infty S_X(t) f_T(t)dt = \int_0^\infty e^{-\mu t}\lambda e^{-\lambda t}dt = \frac{\lambda}{\lambda+\mu} = \frac{1-p_0}{2},$$

y el valor del parámetro entero s se halla de forma que

$$p_1 = \Pr(T > Y) = \int_0^\infty S_T(y) f_Y(y) dy = \int_0^\infty e^{-\lambda y} \frac{\mu^s}{\Gamma(s)} e^{-\mu y} y^{s-1} dt = \left(\frac{\mu}{\lambda + \mu}\right)^s$$

sea aproximadamente $(1 - p_0)/2$. Resulta que $\mu = \lambda(1 + p_0)/(1 - p_0)$, mientras que s se obtiene redondeando $\log\{(1 - p_0)/2\} / \log\{(1 + p_0)/2\}$ al entero más próximo. Consecuentemente, para una censura aproximada del 10%, 25% y 50%, se verifica respectivamente que $\mu = 19, 7$ y 3, y que $s = 58, 16$ y 5. Aproximadamente $p_1 = 0.051$, 0.118 y 0.237 en las tres situaciones precedentes, con lo cual los correspondientes niveles de censura son en realidad del 10.1%, 24.3% y 48.7%.

La tabla 2.5 muestra el efecto de la proporción de seguimiento, f^*, sobre el *sesgo* $(\times 10^4)$ de λ^* y el error cuadrático medio (*ecm*) de $\sqrt{N}\lambda^*$. Obsérvese que, aun para muestras pequeñas, los valores del *ecm* están próximos a la varianza asintótica teórica. Como era de esperar, los resultados del *sesgo* y *ecm* decrecen cuando aumenta la proporción de seguimiento o el tamaño muestral para cada nivel de censura. Por otra parte, si aumenta el nivel de censura, los valores del *sesgo* y *ecm* crecen.

			$N = 25$		$N = 50$		$N = 100$		*varianza*
p_1	p_2	f^*	*sesgo*	*ecm*	*sesgo*	*ecm*	*sesgo*	*ecm*	*asintótica*
0.051	0.050	0.1	470	1.749	209	1.514	87	1.152	1.128
		0.4	276	1.305	166	1.176	74	1.072	1.066
		0.7	275	1.302	164	1.174	71	1.070	1.057
		1.0	274	1.301	163	1.172	69	1.068	1.068
0.118	0.125	0.1	559	1.869	241	1.577	100	1.536	1.478
		0.4	326	1.371	191	1.280	89	1.248	1.192
		0.7	320	1.351	180	1.258	77	1.205	1.152
		1.0	319	1.344	177	1.246	74	1.184	1.135
0.237	0.250	0.1	654	2.921	301	2.728	203	2.564	2.506
		0.4	456	1.730	203	1.607	99	1.576	1.525
		0.7	393	1.602	197	1.475	84	1.439	1.384
		1.0	381	1.558	178	1.433	79	1.402	1.328

Tabla 2.5 : Resultados del *sesgo* $\times 10^4$ de λ^*, *ecm* de $\sqrt{N}\lambda^*$ y *va* $\left(\sqrt{N}\lambda^*\right)$ en el *Caso I* para $\lambda = 1$.

Resultados similares a los expuestos anteriormente se obtienen para el estimador de la mediana poblacional, $t^*_{0.50}$, en la tabla 2.6. Nótese que, como $t^*_{0.50}$ es igual a $\log 2/\lambda^*$, se verifica que la varianza asintótica de $\sqrt{N}t^*_{0.50}$ viene dada por

$$va\left(\sqrt{N}t^*_{0.50}\right) = (\log 2)^2\, va\left(\sqrt{N}\lambda^*\right)/\lambda^4.$$

Asimismo, del análisis de los percentiles estimados $t^*_{0.10}$ y $t^*_{0.90}$ se deducen consecuencias análogas.

			$N=25$		$N=50$		$N=100$		*varianza*
p_1	p_2	f^*	*sesgo*	*ecm*	*sesgo*	*ecm*	*sesgo*	*ecm*	*asintótica*
0.051	0.050	0.1	-98	0.667	-44	0.594	-23	0.554	0.542
		0.4	48	0.556	35	0.542	20	0.522	0.512
		0.7	46	0.553	33	0.541	16	0.520	0.508
		1.0	45	0.551	32	0.539	14	0.519	0.506
0.118	0.125	0.1	105	0.770	45	0.748	24	0.731	0.710
		0.4	101	0.666	44	0.633	23	0.612	0.573
		0.7	97	0.643	43	0.601	22	0.589	0.553
		1.0	95	0.633	43	0.585	22	0.578	0.545
0.237	0.250	0.1	123	1.343	73	1.291	35	1.261	1.204
		0.4	116	0.816	71	0.786	33	0.761	0.733
		0.7	111	0.756	68	0.724	31	0.701	0.665
		1.0	108	0.736	67	0.706	30	0.680	0.638

Tabla 2.6 : Resultados del *sesgo* $\times 10^4$ de $t^*_{0.50}$, *ecm* de $\sqrt{N}t^*_{0.50}$ y $va\left(\sqrt{N}t^*_{0.50}\right)$ en el *Caso I* para $\lambda = 1$.

En segundo término, como *Caso II*, se considera que la amplitud aleatoria de la ventana de observación, $Y - X$, y las variables aleatorias T y X siguen sendas distribuciones de Weibull. De forma explícita, $Y - X \sim W(\mu, r)$, $T \sim W(\lambda, r)$ y $X \sim W(\sigma, r)$, donde $\lambda = 1$, r puede ser 0.5, 1.0 ó 2.0, el parámetro σ se obtiene resolviendo la ecuación

$$p_2 = \Pr(T < X) = \int_0^\infty S_X(t) f_T(t)dt = \int_0^\infty e^{-\sigma t^r}\lambda r t^{r-1} e^{-\lambda t^r} dt = \frac{\lambda}{\lambda+\sigma} = \frac{1-p_0}{2},$$

y el valor de μ debe verificar que

$$p_1 = \Pr(T > Y) = \int_0^\infty S_T(y) f_Y(y)dy = (1-p_0)/2.$$

De esto resulta que $\sigma = \lambda\,(1+p_0)\,/\,(1-p_0)$, mientras que el valor de μ sólo puede hallarse analíticamente cuando $r=1$, en cuyo caso $\mu = \lambda\,(1-p_0)\,/\,(2p_0)$ ya que

$$p_1 = \int_0^\infty e^{-\lambda y}\frac{\sigma\mu}{\sigma-\mu}\left(e^{-\mu y}-e^{-\sigma y}\right)dy = \frac{\sigma\mu}{(\lambda-\sigma)\,(\lambda-\mu)}.$$

Por tanto, para $r=1$ y un nivel de censura del 10%, 25% y 50% se verifica respectivamente que $\sigma = 19, 7$ y 3, mientras que $\mu = 1/18$, $1/6$ y $1/2$.

La tabla 2.7 presenta los valores del *sesgo* $(\times 10^4)$, *ecm* $(\times 10^4)$ y porcentajes de eficiencia relativa (ef) de las estimaciones λ^* y r^* para un 10% y 25% de censura y $N = 100$. La eficiencia relativa se obtiene multiplicando por 100 la razón entre (*ecm* cuando $f^*=1$) y (*ecm* cuando $f^*=i$) para $i = 0.1$, 0.4, 0.7 y 1.0.

		10% *censura*				25% *censura*			
r	f^*	0.1	0.4	0.7	1.0	0.1	0.4	0.7	1.0
0.5	$sesgo(\lambda^*)$	91	78	75	70	111	92	88	78
	$ecm(\lambda^*)$	130	122	119	118	174	131	127	125
	$ef(\lambda^*)$	90.7	96.7	99.3	100	71.8	95.0	98.3	100
	$sesgo(r^*)$	84	80	79	77	182	171	169	168
	$ecm(r^*)$	21	19	18	18	30	24	22	21
	$ef(r^*)$	86.9	95.6	99.2	100	70.8	89.7	95.6	100
1.0	$sesgo(\lambda^*)$	126	116	106	104	154	138	125	116
	$ecm(\lambda^*)$	143	134	130	129	188	148	137	135
	$ef(\lambda^*)$	90.0	96.1	98.8	100	72.1	90.9	98.5	100
	$sesgo(r^*)$	181	170	163	161	435	393	374	369
	$ecm(r^*)$	79	73	70	69	137	105	97	93
	$ef(r^*)$	86.4	95.2	98.5	100	68.1	88.7	96.7	100
2.0	$sesgo(\lambda^*)$	81	73	69	68	99	86	81	75
	$ecm(\lambda^*)$	134	122	120	119	177	137	128	126
	$ef(\lambda^*)$	88.5	97.4	99.0	100	71.2	91.9	98.2	100
	$sesgo(r^*)$	272	256	247	250	859	753	709	687
	$ecm(r^*)$	337	304	295	292	572	421	381	359
	$ef(r^*)$	86.7	96.2	98.9	100	62.9	85.4	94.4	100

Tabla 2.7 : Resultados del *sesgo* $\times 10^4$, *ecm* $\times 10^4$ y *eficiencia* de λ^* y r^* en el *Caso II* para $N = 100$.

El efecto de la proporción de seguimiento sobre el *sesgo* y el *ecm* es similar al que se obtuvo en el *Caso I*. Obsérvese que un 40% de seguimiento es suficiente para alcanzar un alto nivel de eficiencia. Cuando aumenta el nivel de censura también lo hacen los valores del *sesgo* y *ecm* de los estimadores λ^* y r^*, sobre todo para el segundo de éstos. Además, el *sesgo* y *ecm* de r^* crecen rápidamente si aumenta el valor del parámetro r.

Cuando se estiman los percentiles de la distribución de supervivencia se obtienen resultados parecidos. La tabla 2.8 muestra los porcentajes de eficiencia relativa de los percentiles estimados $t^*_{0.10}$, $t^*_{0.50}$ y $t^*_{0.90}$. En este caso, la eficiencia de $t^*_{0.10}$ y $t^*_{0.50}$ es bastante similar, mientras que la de $t^*_{0.90}$ es apreciablemente inferior, sobre todo cuando aumenta el nivel de censura y se reduce la proporción de seguimiento.

		10% *censura*				25% *censura*			
r	f^*	0.1	0.4	0.7	1.0	0.1	0.4	0.7	1.0
0.5	$ef(t^*_{0.10})$	94.5	98.4	99.8	100	88.3	95.4	97.9	100
	$ef(t^*_{0.50})$	94.0	98.2	99.6	100	85.1	96.5	99.0	100
	$ef(t^*_{0.90})$	67.2	91.2	98.6	100	37.9	83.1	94.7	100
1.0	$ef(t^*_{0.10})$	97.0	98.9	99.8	100	88.8	96.4	98.2	100
	$ef(t^*_{0.50})$	94.2	98.4	99.7	100	85.3	97.4	98.8	100
	$ef(t^*_{0.90})$	74.5	91.5	96.6	100	43.0	80.7	94.5	100
2.0	$ef(t^*_{0.10})$	96.0	98.8	99.6	100	87.7	96.5	98.0	100
	$ef(t^*_{0.50})$	94.2	98.8	99.5	100	86.9	97.6	99.1	100
	$ef(t^*_{0.90})$	71.3	92.9	97.5	100	40.7	76.1	93.6	100

Tabla 2.8 : Resultados de la *eficiencia* de $t^*_{0.10}$, $t^*_{0.50}$ y $t^*_{0.90}$ en el *Caso II* para $N = 100$.

Capítulo 3

Estimación semiparamétrica de la función de supervivencia

3.1 Introducción

El propósito de este capítulo es presentar un método semiparamétrico para la estimación de la función de supervivencia de la variable de interés cuando se analizan datos incompletos en un modelo de doble censura aleatoria. Al igual que en el capítulo anterior, los datos que se disponen son incompletos en un doble sentido. En el modelo de doble censura aleatorio, el tiempo de supervivencia de algunos individuos es desconocido; en su lugar, se observa solamente el tiempo de censura a la derecha o a la izquierda. Además, se considera que la variable de censura a la derecha no siempre se conoce, esto es, los datos censurados a la derecha son sólo parcialmente observables. Este tipo de datos incompletos han sido analizados con amplitud en el caso de censura a la derecha por Suzuki (1985a, 1985b), en la estimación de la funcion de supervivencia desde el punto de vista tanto no paramétrico como paramétrico, y por Kalbfleisch y Lawless (1988), en la estimación de los coeficientes de regresión en modelos paramétricos de fiabilidad. Al igual que argumentaba Suzuki, se recomienda obtener algún porcentaje de los datos censurados a la derecha cuando no se cuenta con sus valores. Esto es fundamental ya que esas observaciones censuradas contienen información relevante acerca del tiempo de supervivencia, y pueden ayudar a conseguir la precisión requerida del estimador.

Por otra parte, Klein et al. (1990) proponen un estimador semiparamétrico de la función de supervivencia para datos censurados por la derecha cuando el tiempo de vida/censura es completamente observable.

En la sección 3.2 se ilustra el modelo mediante un ejemplo en el que se analiza un sistema de garantía de calidad que cada vez es más popular. Asimismo, bajo la suposición de que el mecanismo de censura no está relacionado con el tiempo de supervivencia de los individuos, se propone un estimador semiparamétrico de la función de supervivencia que extiende la idea dada por Klein et al. (1990) al caso de datos incompletos y doblemente censurados. El estimador propuesto trata las observaciones no censuradas de forma no paramétrica y utiliza modelos paramétricos para las observaciones censuradas tanto a la derecha como a la izquierda. Por tanto, mantiene la mayoría de las propiedades de distribución libre de los estimadores no paramétricos autoconsistentes si el nivel de censura no es elevado, y permite estimar la función de supervivencia en las colas con una precisión razonable. En la sección 3.3 se examinan algunas propiedades asintóticas del estimador obtenido, concretamente su consistencia y normalidad asintótica, todo ello supuesto una elección correcta del modelo paramétrico teórico. Por último, la sección 3.4 presenta un estudio computacional diseñado para comparar el comportamiento del estimador semiparamétrico de la supervivencia frente al paramétrico propuesto en el capítulo anterior y al no paramétrico autoconsistente (Turnbull, 1974; Mykland y Ren, 1996). En los mismos dos casos estudiados en el capítulo 2, se analiza el efecto de la proporción de seguimiento, f^*, sobre la eficiencia de los estimadores precedentes, según el nivel de censura y el tamaño muestral.

3.2 Construcción del estimador

3.2.1 Una aplicación práctica

Muchos establecimientos que venden fotocopiadoras ofrecen un sistema de garantía que cubre cualquier desperfecto que tenga cada nueva máquina durante un cierto periodo de tiempo y/o hasta un determinado número de fotocopias. Esta clase de sistemas de garantía de calidad se ha hecho muy popular en los últimos años. Son útiles tanto para los compradores, que quieren tener la certeza de que la máquina que han comprado funcionará correctamente, como para los empresarios, que desean conocer la fiabilidad de los productos que venden.

Veamos qué tipos de datos aparecerán en esta situación. Si una fotocopiadora falla durante el periodo de garantía, su propietario solicitará al establecimiento que la repare. Cuando la máquina llegue al servicio de reparación, el establecimiento podrá obtener una información exacta acerca del fallo, por ejemplo, el número de fotocopias que ha hecho. Este tipo de datos se denominan completos o no censurados. Nótese que en este caso la variable de interés T, *"número de fotocopias realizadas antes del primer fallo"*, no es una variable temporal sino que mide la frecuencia de uso. Si la fotocopiadora no falla durante el periodo de garantía, no puede determinarse el valor exacto de T. Se dice que tales datos están censurados por la derecha. En esta aplicación, incluso la empresa no puede conocer el valor de Y, esto es, *"el número de fotocopias realizadas durante el periodo de garantía"*. No obstante, puede obtenerse información de la variable Y mediante el seguimiento de un cierto porcentaje de esas unidades censuradas por la derecha (por ejemplo, utilizando el servicio postal). Obsérvese, sin embargo, que incluso algunos de los propietarios no responderán a lo solicitado. Por tanto, el establecimiento tendrá una relación parcial de esta clase de datos. Los datos se denominarán incompletos en este sentido de observación parcial: la variable de censura por la derecha, Y, no siempre se conoce. En esta clase de estudios a posteriori será primordial estimar el número apropiado de máquinas que deben someterse a un seguimiento posterior para alcanzar un equilibrio entre el coste y la precisión del estimador.

Con objeto de incrementar las ventas, algunos empresarios admiten un periodo adicional de prueba previo al de garantía (por ejemplo, quince días), en el cual se contrasta la fiabilidad del producto, y se consideran fallos de naturaleza distinta al que ocasiona el uso normal (piezas de repuesto defectuosas, fallos de fábrica, etc.). Si una fotocopiadora falla durante el periodo de prueba, dicha máquina será automáticamente enviada al fabricante y reemplazada por otra nueva. Esto es importante no sólo para los establecimientos, que pueden así incrementar las ventas y conocer la calidad de los productos que venden, sino también para los compradores, quienes, si ocurre un fallo, pueden cambiar la fotocopiadora por otra nueva (la mayoría de la gente piensa que una máquina reparada no es comparable a una nueva) o recuperar su dinero. En tal caso, se supone que el vendedor no puede obtener el valor de T ya que la fotocopiadora no pasará por el servicio de reparación del establecimiento. Por consiguiente, no conoce el valor de X, *"número de fotocopias hechas en el periodo de prueba"*. En esta aplicación práctica se considera que X toma un valor por defecto para todas las unidades censuradas por la izquierda, por ejemplo, "el número medio de fotocopias realizadas hasta el inicio del periodo de garantía por todas las máquinas que sobreviven al periodo de prueba". Luego, es necesario que esta información esté disponible. Sin embargo, los resultados presentados en este capítulo se obtienen en el caso general, donde X se observa para todas las unidades que fallan en el periodo de prueba.

Nótese que, tanto la variable de interés, T, como las variables de censura, X e Y, se miden en las mismas unidades (número de fotocopias), mientras que los periodos de prueba y garantía se miden en unidades de tiempo.

Supuesto el modelo de doble censura aleatoria con información incompleta que se presentó en el capítulo anterior e idéntica notación, las fotocopiadoras del estudio se clasifican en los cuatro grupos siguientes: las máquinas que fallan en el periodo de garantia, aquéllas que lo hacen en el de prueba, las que no fallan antes de la expiración de la garantía y su propietario no informó del valor de Y (fotocopiadoras sin seguimiento), y aquéllas de las sí informó (fotocopiadoras con seguimiento). El número de unidades en cada grupo se denota por N_0, N_2, N_{10} y N_{11} respectivamente, donde

$$N_0 = \sum_{i=1}^{N} \mathrm{I}(\delta_i = 0)\,, \qquad N_2 = \sum_{i=1}^{N} \mathrm{I}(\delta_i = 2)\,,$$
$$N_{10} = \sum_{i=1}^{N} \mathrm{I}(\delta_i = 1)(1 - F_i)\,, \qquad N_{11} = \sum_{i=1}^{N} \mathrm{I}(\delta_i = 1)F_i\,,$$

$F_i = 1$ si la fotocopiadora i-ésima está sujeta a seguimiento ($F_i = 0$ en otro caso), y $N = N_0 + N_2 + N_{10} + N_{11}$ es el número total de fotocopiadoras en estudio. Al igual que en el capítulo 2, se supone que tanto N_{11} como la probabilidad de que una fotocopiadora tenga un seguimiento posterior, f^*, son positivos.

3.2.2 El estimador semiparamétrico

Se desea estimar la función de supervivencia de T a partir de una muestra de N pares independientes, (Z_i, δ_i), $i = 1, ..., N$, donde (Z_i, δ_i) está definido como en (2.1)-(2.2), pero, cuando $\delta_i = 1$, los valores de Z_i sólo se observan si $F_i = 1$. A su vez, a partir de estudios preliminares o mediante una análisis gráfico de los datos, se selecciona un modelo paramétrico, $S(t;\theta)$, para la función de supervivencia de T, donde θ es un vector de parámetros desconocidos que pertence a determinado espacio paramétrico Θ. Se supone que el soporte del modelo, $\{t \geq 0 : 0 < S(t;\theta) < 1\}$, no depende de θ y que $\widehat{\theta}$ es un estimador consistente de θ basado en la muestra observada o en el análisis preliminar de otra muestra.

Si se hubiera observado el valor de T_i, para todo $i = 1, ..., N$, se tendría un problema con datos completos, y S_T podría estimarse mediante la correspondiente función de supervivencia empírica. Por analogía al caso anterior, y después de prestar atención especial a aquellos datos que no son directamente observables y a las observaciones censuradas a la derecha que son desconocidas, se propone una especie de función de supervivencia semi-empírica como estimador de S_T.

Obsérvese que si $Z_i > t$ y $\delta_i = 0$, o bien $\delta_i = 1$ y $F_i = 1$, se sabe con certeza que $T_i > t$. Análogamente, cuando $Z_i \leq t$ y $\delta_i = 0$ ó 2, se cumple con seguridad que $T_i \leq t$. Sin embargo, no es posible determinar si T_i es mayor que t o no cuando $Z_i \leq t$, $\delta_i = 1$ y $F_i = 1$, o cuando $Z_i > t$ y $\delta_i = 2$. Lo mismo sucede para aquellas unidades con $\delta_i = 1$ y $F_i = 0$, en las cuales no se ha observado el valor de Z_i. Se asumirá que la distribución de las N_{10} unidades censuradas a la derecha sin seguimiento coincide con la de las N_{11} unidades con seguimiento. Esto es, las pérdidas de información son aleatorias, o bien, las unidades con seguimiento se eligen al azar. Por tanto, la masa de las unidades sin seguimiento se redistribuirá equitativamente entre las que sí lo tienen. Consecuentemente, cada una de las N_{11} observaciones censuradas a la derecha que se conocen tendrá una masa adicional de N_{10}/N_{11}.

En todos los casos indeterminados mencionados anteriormente, se utilizará la información disponible y el modelo paramétrico seleccionado, $S(t;\theta)$, para estimar

$\Pr(T_i > t)$ de la forma siguiente:

Si $Z_i \leq t$, $\delta_i = 1$ y $F_i = 1$, resulta que

$$\Pr(T_i > t \mid \delta_i = 1,\ F_i = 1) = \Pr(T_i > t \mid T_i > Z_i,\ Y_i = Z_i,\ F_i = 1) = C_1(t, Z_i; \theta),$$

donde

$$C_1(t, Z_i; \theta) = \begin{cases} \dfrac{S(t;\theta)}{S(Z_i;\theta)} & si \quad S(t;\theta) > 0, \\ 0 & si \quad S(t;\theta) = 0, \end{cases}$$

mientras que, si $Z_i > t$ y $\delta_i = 2$, se verifica que

$$\Pr(T_i > t \mid \delta_i = 2) = \Pr(T_i > t \mid T_i < Z_i,\ X_i = Z_i) = C_2(t, Z_i; \theta),$$

donde

$$C_2(t, Z_i; \theta) = \begin{cases} \dfrac{S(t;\theta) - S(Z_i^-;\theta)}{1 - S(Z_i^-;\theta)} & si \quad S(t;\theta) < 1, \\ 1 & si \quad S(t;\theta) = 1, \end{cases}$$

supuesto que el modelo paramétrico es correcto.

Los resultados anteriores conducen a proponer el siguiente estimador de la función de supervivencia de T :

$$\widehat{S}_T(t) = \frac{1}{N} \sum_{i=1}^{N} \omega_i(t \mid S(\cdot;\widehat{\theta})), \tag{3.1}$$

donde

$$\omega_i(t \mid S(\cdot;\widehat{\theta})) = \begin{cases} 0 & si \quad Z_i \leq t, \delta_i = 0, 2 \text{ ó } \delta_i = 1, F_i = 0 \\ \left(1 + \dfrac{N_{10}}{N_{11}}\right) C_1(t, Z_i; \widehat{\theta}) & si \quad Z_i \leq t, \delta_i = 1, F_i = 1 \\ 1 & si \quad Z_i > t, \delta_i = 0 \\ \left(1 + \dfrac{N_{10}}{N_{11}}\right) & si \quad Z_i > t, \delta_i = 1, F_i = 1 \\ C_2(t, Z_i; \widehat{\theta}) & si \quad Z_i > t, \delta_i = 2 \end{cases}$$

Nótese que, si la variable de censura por la izquierda también se observa de forma parcial, la extensión del procedimiento de estimación utilizado es inmediata.

Ahora, se procede a comprobar que el estimador propuesto está bien definido.

Teorema 3.1

El estimador $\widehat{S}_T$ definido en (3.1) es una función de supervivencia.

Demostración

Puesto que

$$\begin{aligned}\widehat{S}_T(t) &= \frac{1}{N}\sum_{i=1}^{N} \mathrm{I}(Z_i > t, \delta_i = 0) + \frac{1}{N}\left(1+\frac{N_{10}}{N_{11}}\right)\sum_{i=1}^{N} \mathrm{I}(Z_i > t, \delta_i = 1, F_i = 1)\\ &+\frac{1}{N}\left(1+\frac{N_{10}}{N_{11}}\right)\sum_{i=1}^{N} C_1(t, Z_i; \widehat{\theta})\mathrm{I}(Z_i \leq t, \delta_i = 1, F_i = 1)\\ &+\frac{1}{N}\sum_{i=1}^{N} C_2(t, Z_i; \widehat{\theta})\mathrm{I}(Z_i > t, \delta_i = 2),\end{aligned}$$

resulta que $\widehat{S}_T(0^-) = \dfrac{1}{N}\left\{N_0 + \left(1+\dfrac{N_{10}}{N_{11}}\right)N_{11} + N_2\right\} = 1$ y $\lim_{t\to\infty} \widehat{S}_T(t) = 0.$

A su vez, el estimador propuesto puede expresarse como

$$\begin{aligned}\widehat{S}_T(t) &= \frac{1}{N}\sum_{i=1}^{N} \mathrm{I}(Z_i > t, \delta_i = 0) + \frac{1}{N}\left(1+\frac{N_{10}}{N_{11}}\right)\sum_{i=1}^{N} \mathrm{I}(\delta_i = 1, F_i = 1)\Lambda(t, Z_i; \widehat{\theta})\\ &+\frac{1}{N}\sum_{i=1}^{N} C_2(t, Z_i; \widehat{\theta})\mathrm{I}(Z_i > t, \delta_i = 2),\end{aligned}$$

donde

$$\Lambda(t, Z_i; \widehat{\theta}) = \left\{\left(1 - C_1(t, Z_i; \widehat{\theta})\right)\mathrm{I}(Z_i > t) + C_1(t, Z_i; \widehat{\theta})\right\}.$$

Por tanto, $\widehat{S}_T$ es no creciente y continua por la derecha en $[0, \infty)$ ya que, para $i = 1, .., N$, las funciones de t

$$\mathrm{I}(Z_i > t, \delta_i = 0), \quad \Lambda(t, Z_i; \widehat{\theta}) \;\text{ y }\; C_2(t, Z_i; \widehat{\theta})\mathrm{I}(Z_i > t, \delta_i = 2)$$

cumplen idénticas propiedades.

□

3.3 Propiedades asintóticas del estimador

En esta sección se supone que el modelo paramétrico seleccionado es correcto, es decir, que $S_T(t) = S(t;\theta)$ para todo $t \geq 0$. Antes de analizar algunas propiedades asintóticas del estimador descrito en la sección anterior, y con objeto de simplificar los cálculos, se procede a obtener una expresión alternativa de éste. Para ello, se consideran las funciones de subsupervivencia empírica $\widehat{S}_0(t)$ y $\widehat{S}_2(t)$, que se basan en los datos no censurados y censurados por la izquierda respectivamente, y la función de subsupervivencia pseudo-empírica $\widehat{S}_1^*(t)$, basada en las observaciones censuradas por la derecha con seguimiento, definidas por:

$$\widehat{S}_0(t) = \frac{1}{N}\sum_{i=1}^{N} \mathrm{I}(Z_i > t, \delta_i = 0),$$

$$\widehat{S}_2(t) = \frac{1}{N}\sum_{i=1}^{N} \mathrm{I}(Z_i > t, \delta_i = 2),$$

$$\widehat{S}_1^*(t) = \frac{1}{N}\left(1 + \frac{N_{10}}{N_{11}}\right)\sum_{i=1}^{N} \mathrm{I}(Z_i > t, \delta_i = 1, F_i = 1).$$

Como $\widehat{S}_1^*(t)$ y $\widehat{S}_2(t)$ son funciones no decrecientes y definidas a trozos con incrementos comunes, $(1 + N_{10}/N_{11})/N$ y $1/N$ respectivamente, en los correspondientes valores observados, es fácil comprobar que el estimador propuesto en este capítulo puede expresarse como:

$$\widehat{S}_T(t) = \widehat{S}_0(t) + \widehat{S}_1^*(t) - \int_0^t C_1(t, Z_i;\widehat{\theta})d\widehat{S}_1^*(u) - \int_t^\infty C_2(t, Z_i;\widehat{\theta})d\widehat{S}_2(u),$$

donde $\int_0^t$ y $\int_t^\infty$ representan las integrales de Lebesgue-Stieltjes en $[0,t]$ y (t,∞) respectivamente.

En el resto de esta sección se consideran los siguientes procesos

$$Q_{0N}(t) = \widehat{S}_0(t) - S_0(t)\,, \quad Q_{1N}(t) = \widehat{S}_1^*(t) - S_1(t)\,, \quad Q_{2N}(t) = \widehat{S}_2(t) - S_2(t)\,,$$

donde $t \geq 0$, y $S_j(t) = \Pr(Z > t,\, \delta = j)\,,\ j = 0, 1, 2$, son las funciones de subsupervivencia de Z,

$$W_{1N}(t,u) = C_1(t,u;\widehat{\theta}) - C_1(t,u;\theta)\,, \text{ para } 0 \leq u \leq t,$$

y

$$W_{2N}(t,u) = C_2(t,u;\widehat{\theta}) - C_2(t,u;\theta), \text{ para } 0 \leq t < u.$$

Lema 3.1

Los procesos estocásticos

$$\Phi_{1N}(t) = \int_0^t C_1(t,u;\theta)dQ_{1N}(u)\,, \qquad \Phi_{2N}(t) = \int_0^t W_{1N}(t,u)dQ_{1N}(u)\,,$$

$$\Phi_{3N}(t) = \int_t^\infty C_2(t,u;\theta)dQ_{2N}(u) \quad \text{y} \quad \Phi_{4N}(t) = \int_t^\infty W_{2N}(t,u)dQ_{2N}(u)$$

convergen a 0 uniformemente en $t \geq 0$ con probabilidad 1 cuando $N \to \infty$.

Demostración

Obsérvese que, por el teorema de Glivenko-Cantelli,

$$\widehat{S}_j(t) = \frac{1}{N}\sum_{i=1}^N I(Z_i > t, \delta_i = j) \xrightarrow[N\to\infty]{c.s.} S_j(t) = \Pr(Z > t, \delta = j)\,, \; j = 0,1,2,$$

y

$$\widehat{S}_{11}(t) = \frac{1}{N}\sum_{i=1}^N I(Z_i > t, \delta_i = 1, F_i = 1) \xrightarrow[N\to\infty]{c.s.} S_{11}(t) = \Pr(Z > t, \delta = 1, F = 1)$$

uniformemente en $t \geq 0$. Asimismo,

$$\widehat{S}_1^*(t) = \left(1 + \frac{N_{10}}{N_{11}}\right)\widehat{S}_{11}(t) \xrightarrow[N\to\infty]{c.s.} \frac{1}{f^*}S_{11}(t) = S_1(t)$$

uniformemente en $t \geq 0$ ya que, por la ley fuerte de los grandes números de Borel,

$$1 + \frac{N_{10}}{N_{11}} = \frac{N_1/N}{N_{11}/N} \xrightarrow[N\to\infty]{c.s.} \frac{\Pr(\delta = 1)}{\Pr(\delta = 1, F = 1)} = \frac{1}{\Pr(F = 1 \mid \delta = 1)} = \frac{1}{f^*}.$$

Luego, los procesos $Q_{0N}(t)$, $Q_{1N}(t)$ y $Q_{2N}(t)$ convergen a 0 uniformemente en $t \geq 0$ con probabilidad 1 cuando $N \to \infty$. Del mismo modo,

$$Q_{jN}(0^-) = \frac{N_j}{N} - \Pr(\delta = j) \xrightarrow[N\to\infty]{c.s.} 0, \;\; j = 0,1,2.$$

Por consiguiente, como $|C_1(t,u;\theta)| \leq 1$ y $|W_{1N}(t,u)| \leq 1$ para $0 \leq u \leq t$,

$$\sup_{t\geq 0} |\Phi_{jN}(t)| \leq \sup_{t\geq 0} \left| \int_0^t dQ_{1N}(u) \right| = \sup_{t\geq 0} \left| Q_{1N}(t) - Q_{1N}(0^-) \right| \xrightarrow[N\to\infty]{c.s.} 0, \quad j = 1, 2.$$

Analogamente,

$$\sup_{t\geq 0} |\Phi_{jN}(t)| \leq \sup_{t\geq 0} \left| \int_t^\infty dQ_{2N}(u) \right| = \sup_{t\geq 0} |Q_{2N}(t)| \xrightarrow[N\to\infty]{c.s.} 0, \quad j = 3, 4,$$

ya que $|C_2(t,u;\theta)| \leq 1$ y $|W_{2N}(t,u)| \leq 1$ para $0 \leq t < u$.

□

Lema 3.2

Si, para todo $t \geq 0$, $S_T(t) = S(t;\theta)$ y $S(t;.)$ es continua en θ, y $\widehat{\theta}$ converge casi seguramente a θ, entonces

$$\Phi_{5N}(t) = \int_0^t W_{1N}(t,u) dS_1(u) \quad \text{y} \quad \Phi_{6N}(t) = \int_t^\infty W_{2N}(t,u) dS_2(u)$$

convergen a 0 con probabilidad 1 cuando $N \to \infty$.

Demostración

Notese que, para $0 \leq u \leq t$, $W_{1N}(t,u)$ converge a 0 casi seguramente ya que, si $S_T(t) = 0$, $W_{1N}(t,u) = 0$ y, si $S_T(t) > 0$,

$$W_{1N}(t,u) = \frac{S(t;\widehat{\theta})}{S(u;\widehat{\theta})} - \frac{S_T(t)}{S_T(u)} \xrightarrow[N\to\infty]{c.s.} 0$$

por la continuidad de $S(t;.)$ en θ y la convergencia casi segura de $\widehat{\theta}$ a θ. De este modo, como $|W_{1N}(t,u)| \leq 1$ para $0 \leq u \leq t$, aplicando el teorema de la convergencia dominada, se obtiene que, para todo $t \geq 0$,

$$\Phi_{5N}(t) = \int_0^t W_{1N}(t,u) dS_1(u) \xrightarrow[N\to\infty]{c.s.} \int_0^t 0 dS_1(u) = 0.$$

Mediante un razonamiento análogo, se comprueba que

$$\Phi_{6N}(t) = \int_t^{\infty} W_{2N}(t,u)dS_2(u) \xrightarrow[N\to\infty]{c.s.} \int_t^{\infty} 0 dS_2(u) = 0.$$

□

Teorema 3.2

Si $\widehat{\theta}$ es un estimador fuertemente consistente de θ, $S(t;.)$ es continua en θ y $S_T(t) = S(t;\theta)$, para todo $t \geq 0$, entonces $\widehat{S}_T(t)$ es un estimador fuertemente consistente de $S_T(t)$, para todo $t \geq 0$. Si, además, S_T es continua en $[0,\infty)$, entonces $\widehat{S}_T$ es un estimador uniformemente fuertemente consistente de S_T.

Demostración

Puesto que $\widehat{\theta}$ converge casi seguramente a θ, es obvio que $\widehat{S}_T(t) = S_T(t)$ casi seguramente cuando $S_T(t)$ es 0 ó 1. Supóngase, por tanto, que $0 < S_T(t) < 1$. Como

$$S_T(t) = E\left[\mathrm{I}\left(T > t\right)\right] = E\left[E\left[\mathrm{I}\left(T > t\right) \mid (Z,\delta)\right]\right] = E\left[\Pr\left(T > t \mid (Z,\delta)\right)\right]$$

resulta que

$$\begin{aligned} S_T(t) &= -\sum_{j=0}^{2} \int_0^{\infty} \Pr\left(T > t \mid Z = u,\, \delta = j\right) dS_j(u) \\ &= -\sum_{j=0}^{2} \int_t^{\infty} \Pr\left(T > t \mid Z = u,\, \delta = j\right) dS_j(u) \\ &\quad -\sum_{j=0}^{2} \int_0^{t} \Pr\left(T > t \mid Z = u,\, \delta = j\right) dS_j(u). \end{aligned}$$

Luego,

$$S_T(t) = -\int_t^{\infty} dS_0(u) - \int_t^{\infty} dS_1(u) - \int_t^{\infty} \frac{S_T(t) - S_T(u^-)}{1 - S_T(u^-)} dS_2(u) - \int_0^t \frac{S_T(t)}{S_T(u)} dS_1(u),$$

de donde se obtiene la siguiente expresión

$$S_T(t) = S_0(t) + S_1(t) - \int_0^t \frac{S_T(t)}{S_T(u)} dS_1(u) - \int_t^{\infty} \frac{S_T(t) - S_T(u^-)}{1 - S_T(u^-)} dS_2(u).$$

Puesto que también

$$\widehat{S}_T(t) = \widehat{S}_0(t) + \widehat{S}_1^*(t) - \int_0^t \frac{S(t;\widehat{\theta})}{S(u;\widehat{\theta})} d\widehat{S}_1^*(u) - \int_t^\infty \frac{S(t;\widehat{\theta}) - S(u^-;\widehat{\theta})}{1 - S(u^-;\widehat{\theta})} d\widehat{S}_2(u),$$

se verifica que $\widehat{S}_T(t) - S_T(t)$ puede expresarse como

$$\left\{\widehat{S}_0(t) - S_0(t)\right\} + \left\{\widehat{S}_1^*(t) - S_1(t)\right\} - \int_0^t \frac{S(t;\widehat{\theta})}{S(u;\widehat{\theta})} d\widehat{S}_1^*(u) + \int_0^t \frac{S_T(t)}{S_T(u)} dS_1(u)$$

$$- \int_t^\infty \frac{S(t;\widehat{\theta}) - S(u^-;\widehat{\theta})}{1 - S(u^-;\widehat{\theta})} d\widehat{S}_2(u) + \int_t^\infty \frac{S_T(t) - S_T(u^-)}{1 - S_T(u^-)} dS_2(u),$$

o bien, como

$$Q_{0N}(t) + Q_{1N}(t) - \int_0^t \left\{W_{1N}(t,u) + \frac{S_T(t)}{S_T(u)}\right\} d\left(S_1(u) + Q_{1N}(u)\right) + \int_0^t \frac{S_T(t)}{S_T(u)} dS_1(u)$$

$$- \int_t^\infty \left\{W_{2N}(t,u) + \frac{S_T(t) - S_T(u^-)}{1 - S_T(u^-)}\right\} d\left(S_2(u) + Q_{2N}(u)\right)$$

$$+ \int_t^\infty \frac{S_T(t) - S_T(u^-)}{1 - S_T(u^-)} dS_2(u),$$

de lo cual resulta finalmente que

$$\begin{aligned}
\widehat{S}_T(t) - S_T(t) \quad = \quad & Q_{0N}(t) + Q_{1N}(t) - \int_0^t W_{1N}(t,u) dS_1(u) - \int_0^t \frac{S_T(t)}{S_T(u)} dQ_{1N}(u) \\
& - \int_0^t W_{1N}(t,u) dQ_{1N}(u) - \int_t^\infty W_{2N}(t,u) dS_2(u) \\
& - \int_t^\infty \frac{S_T(t) - S_T(u^-)}{1 - S_T(u^-)} dQ_{2N}(u) - \int_t^\infty W_{2N}(t,u) dQ_{2N}(u).
\end{aligned} \tag{3.2}$$

Luego, ya que $Q_{0N}(t)$ y $Q_{1N}(t)$ convergen casi seguramente a 0, se obtiene que

$$\widehat{S}_T(t) - S_T(t) = Q_{0N}(t) + Q_{1N}(t) - \sum_{j=1}^{6} \Phi_{jN}(t) \underset{N\to\infty}{\xrightarrow{c.s.}} 0\,, \text{ para todo } t \geq 0,$$

a partir de los lemas 3.1 y 3.2.

Si, además, S_T es continua en $[0,\infty)$, entonces dado un $\varepsilon > 0$ arbitrario, existen $(k+1)$ puntos $0 = t_0 < t_1 < ... < t_k = \infty$ tales que

$$\Lambda_j = S_T(t_{j-1}) - S_T(t_j) < \varepsilon/2,\ j = 1, 2, .., k.$$

A su vez, como $\widehat{S}_T(t_j) \xrightarrow[N\to\infty]{c.s.} S_T(t_j)$, $j = 0, .., k$, existe un número natural m tal que

$$\forall N \geq m:\ \max_{j=0,..,k} \left|\widehat{S}_T(t_j) - S_T(t_j)\right| < \varepsilon/2 \quad c.\,s.$$

Dado $t \geq 0$ arbitrario, existe $j \in \{1, 2, .., k\}$ tal que $t \in [t_{j-1}, t_j)$. Si ocurre que $\widehat{S}_T(t) - S_T(t) \geq 0$, entonces

$$\widehat{S}_T(t) - S_T(t) \leq \widehat{S}_T(t_{j-1}) - S_T(t_j) = \widehat{S}_T(t_{j-1}) - S_T(t_{j-1}) + S_T(t_{j-1}) - S_T(t_j),$$

de donde

$$\left|\widehat{S}_T(t) - S_T(t)\right| \leq \left|\widehat{S}_T(t_{j-1}) - S_T(t_{j-1})\right| + \left|S_T(t_{j-1}) - S_T(t_j)\right|.$$

De igual modo, si $\widehat{S}_T(t) - S_T(t) < 0$,

$$S_T(t) - \widehat{S}_T(t) \leq S_T(t_{j-1}) - \widehat{S}_T(t_j) = S_T(t_{j-1}) - S_T(t_j) + S_T(t_j) - \widehat{S}_T(t_j),$$

lo que implica que

$$\left|\widehat{S}_T(t) - S_T(t)\right| \leq \left|\widehat{S}_T(t_j) - S_T(t_j)\right| + \left|S_T(t_{j-1}) - S_T(t_j)\right|.$$

Por consiguiente, puesto que

$$\forall N \geq m:\ \sup_{t\geq 0} \left|\widehat{S}_T(t) - S_T(t)\right| \leq \max_{j=0,..,k} \left|\widehat{S}_T(t_j) - S_T(t_j)\right| + \max_{j=1,..,k} \Lambda_j < \varepsilon \quad c.\,s.,$$

se obtiene la consistencia uniforme fuerte de $\widehat{S}_T$, esto es,

$$\sup_{t\geq 0} \left|\widehat{S}_T(t) - S_T(t)\right| \xrightarrow[N\to\infty]{c.s.} 0.$$

□

Observación

Si se supone que $\widehat{\theta}$ converge en probabilidad a θ, la convergencia de $\Phi_{5N}(t)$ y $\Phi_{6N}(t)$ a 0 en el lema 3.2 sería en probabilidad, y la consistencia de $\widehat{S}_T(t)$ en el teorema 3.2 sería débil.

Lema 3.3

Para cada $t \geq 0$, sea $Q_N(t) = (Q_{0N}(t), Q_{1N}(t), Q_{2N}(t))$. Si $0 \leq t_1 < ... < t_m$, entonces la distribución asintótica de $\sqrt{N}\,(Q_N(t_1), ..., Q_N(t_m))$ es normal de media cero y matriz de varianzas-covarianzas

$$\Omega\,(t_1, ..., t_m) = \left(\Omega_{t_i t_j}\right)_{\substack{i=1,..,m,\\ j=1,..,m,}}$$

donde, para $s \leq t$, la matriz $\Omega_{ts} = \left(\omega_{ij}^{ts}\right)_{3\times 3}$ se define como

$$\Omega_{ts} = \begin{pmatrix} S_0(t)\left\{1 - S_0(s)\right\} & \dfrac{-S_0(t)S_{11}(s)}{f^*} & -S_0(t)S_2(s) \\ \dfrac{-S_{11}(t)S_0(s)}{f^*} & \dfrac{S_{11}(t)\left\{1 - S_{11}(s)\right\}}{f^{*2}} & \dfrac{-S_{11}(t)S_2(s)}{f^*} \\ -S_2(t)S_0(s) & \dfrac{-S_2(t)S_{11}(s)}{f^*} & S_2(t)\left\{1 - S_2(s)\right\} \end{pmatrix},$$

y Ω_{st} es la traspuesta de Ω_{ts}.

Demostración

En primer lugar, se consideran las matrices diagonales

$$A_N = diag(1, 1 + N_{10}/N_{11}, 1) \ \text{ y } \ A = diag(1, 1/f^*, 1),$$

y los vectores $B_N(t) = \dfrac{1}{N}\sum\limits_{j=1}^{N} D_j(t)$, donde

$$D_j(t) = (\mathrm{I}(Z_j > t, \delta_j = 0),\, \mathrm{I}(Z_j > t, \delta_j = 1, F_j = 1),\, \mathrm{I}(Z_j > t, \delta_j = 2))^{\mathrm{t}}\,,$$

para $j = 1, .., N$, son independientes e idénticamente distribuidos con media $\beta\,(t) = (S_0(t), S_{11}(t), S_2(t))^{\mathrm{t}}$. Obsérvese que, con esa notación,

$$\sqrt{N}\left(Q_N(t_1),...,Q_N(t_m)\right)=\sqrt{N}\left(\,A_N\left(B_N(t_1),...,B_N(t_m)\right)-A\left(\beta(t_1),...,\beta(t_m)\right)\,\right).$$

Puesto que $(B_N(t_1),...,B_N(t_m))=\dfrac{1}{N}\sum\limits_{j=1}^{N}D_j(t_1,...,t_m)$, donde

$$D_j(t_1,...,t_m)=\left(D_j(t_1),...,D_j(t_m)\right),\ j=1,..,N,$$

son vectores independientes e idénticamente distribuidos, entonces, por el teorema central del límite multidimensional, se obtiene que $\sqrt{N}\left(B_N(t_1),...,B_N(t_m)\right)$ es asintóticamente normal de media $\sqrt{N}\left(\beta(t_1),...,\beta(t_m)\right)$ y matriz de varianzas-covarianzas

$$\Sigma\left(t_1,..,t_m\right)=\left(\Sigma_{t_it_j}\right)_{\substack{i=1,..,m,\\ j=1,..,m,}}$$

donde, para $s\leq t$, $\Sigma_{ts}=diag\left(\beta\left(t\right)\right)-\beta\left(t\right)\beta\left(s\right)^{\mathrm{t}}$ y Σ_{st} es la traspuesta de Σ_{ts}. La expresión de la matriz $\Sigma_{ts}=\left(\sigma_{ij}^{ts}\right)_{3\times 3}$ puede obtenerse con facilidad. Por ejemplo:

$$\begin{aligned}
\sigma_{11}^{ts} &= cov\left[\mathrm{I}(Z>t,\delta=0),\mathrm{I}(Z>s,\delta=0)\right]\\
&= E\left[\mathrm{I}(Z>t,\delta=0)\mathrm{I}(Z>s,\delta=0)\right]-E\left[\mathrm{I}(Z>t,\delta=0)\right]E\left[\mathrm{I}(Z>s,\delta=0)\right]\\
&= S_0(t)-S_0(t)S_0(s),
\end{aligned}$$

$$\begin{aligned}
\sigma_{13}^{ts} &= cov\left[\mathrm{I}(Z>t,\delta=0),\mathrm{I}(Z>s,\delta=2)\right]\\
&= E\left[\mathrm{I}(Z>t,\delta=0)\mathrm{I}(Z>s,\delta=2)\right]-E\left[\mathrm{I}(Z>t,\delta=0)\right]E\left[\mathrm{I}(Z>s,\delta=2)\right]\\
&= 0-S_0(t)S_2(s),
\end{aligned}$$

$$\begin{aligned}
\sigma_{22}^{ts} &= cov\left[\mathrm{I}(Z>t,\delta=1,F=1),\mathrm{I}(Z>s,\delta=1,F=1)\right]\\
&= S_{11}(t)-S_{11}(t)S_{11}(s)
\end{aligned}$$

y, en último lugar,

$$\begin{aligned}
\sigma_{23}^{ts} &= cov\left[\mathrm{I}(Z>t,\delta=1,F=1)\mathrm{I}(Z>s,\delta=2)\right]\\
&= 0-E\left[\mathrm{I}(Z>t,\delta=1,F=1)\right]E\left[\mathrm{I}(Z>t,\delta=2)\right]=-S_{11}(t)S_2(t).
\end{aligned}$$

Por otra parte, como $1+N_{10}/N_{11}\xrightarrow[N\to\infty]{c.\,s.}1/f^{*}$, resulta que $A_N\xrightarrow[N\to\infty]{c.\,s.}A$. Luego,

$$\sqrt{N}\left(Q_N(t_1),...,Q_N(t_m)\right)$$

está asintóticamente idénticamente distribuido como

$$\sqrt{N} A\left(B_N(t_1)-\beta(t_1),...,B_N(t_m)-\beta(t_m)\right).$$

Por consiguiente, ya que $\Omega_{t_it_j}=A\cdot\Sigma_{t_it_j}\cdot A$, $i, j=1,..,m$, se verifica que

$$\sqrt{N}\left(Q_N(t_1),...,Q_N(t_m)\right)\underset{N\to\infty}{\xrightarrow{d}} N\left(0,\Omega\left(t_1,..,t_m\right)\right),$$

donde $\Omega\left(t_1,...,t_m\right)=\left(\Omega_{t_it_j}\right)_{\substack{i=1,..,m\\ j=1,..,m}}$. Nótese que $\Omega_{t_it_j}$ puede expresarse como

$$\Omega_{t_it_j}=diag(1,1/f^*,1)\left\{diag\left[\beta\left(\max\left\{t_i,t_j\right\}\right)\right]-\beta\left(t_i\right)\beta\left(t_j\right)^{t}\right\}diag(1,1/f^*,1).$$

$\square$

Teorema 3.3

Si las variables aleatorias T, X e Y son continuas, el soporte de X (o de T) está acotado, $\sqrt{N}(\widehat{\theta}-\theta)$ converge en distribución a una variable aleatoria normal de media cero y, para todo $t\geq 0$, $S_T(t)=S(t;\theta)$ y $S(t;.)$ es diferenciable con continuidad en un entorno de θ, entonces el proceso $\Psi_N=\left\{\Psi_N(t),\tau_1\leq t\leq\tau_2\right\}$, donde $\Psi_N(t)=\sqrt{N}\left(\widehat{S}_T(t)-S_T(t)\right)$ y $1>S_T(\tau_1)>S_T(\tau_2)>0$, converge débilmente a un proceso Gaussiano $\Psi=\left\{\Psi(t),\tau_1\leq t\leq\tau_2\right\}$ de media cero.

Demostración

Como $d(UV)=UdV+V^-dU$, se obtiene que $\int V^-dU=UV-\int UdV$. Por tanto, para $t\in[\tau_1,\tau_2]$ arbitrario, se verifica que

$$\begin{aligned}\Phi_{1N}(t) &= \int_0^t\frac{S_T(t)}{S_T(u)}dQ_{1N}(u)=\left.\frac{S_T(t)}{S_T(u)}Q_{1N}(u)\right]_{0^-}^{t}-\int_0^t Q_{1N}(u)S_T(t)dS_T^{-1}(u)\\ &= Q_{1N}(t)-S_T(t)Q_{1N}(0^-)-S_T(t)\int_0^t Q_{1N}(u)dS_T^{-1}(u),\end{aligned}$$

donde $V=S_T(t)/S_T(u)$, $dU=dQ_{1N}(u)$, $dV=S_T(t)dS_T^{-1}(u)$ y $U=Q_{1N}(u)$ en la fórmula de integración por partes. Del mismo modo, como

$$
\begin{aligned}
\Phi_{3N}(t) &= \int_t^\infty \frac{S_T(t) - S_T(u)}{1 - S_T(u)} dQ_{2N}(u) \\
&= S_T(t) \int_t^\infty \frac{1}{1 - S_T(u)} dQ_{2N}(u) - \int_t^\infty \frac{S_T(u)}{1 - S_T(u)} dQ_{2N}(u),
\end{aligned}
$$

e, integrando por partes, resulta que

$$
\begin{aligned}
\int_t^\infty \frac{1}{1 - S_T(u)} dQ_{2N}(u) &= \left. \frac{Q_{2N}(u)}{1 - S_T(u)} \right]_t^\infty - \int_t^\infty Q_{2N}(u) d\left(1 - S_T(u)\right)^{-1} \\
&= -\frac{Q_{2N}(t)}{1 - S_T(t)} - \int_t^\infty Q_{2N}(u) d\left(1 - S_T(u)\right)^{-1},
\end{aligned}
$$

donde en este caso $V = (1 - S_T(u))^{-1}$, $dU = dQ_{2N}(u)$, $dV = d\left(1 - S_T(u)\right)^{-1}$ y $U = Q_{2N}(u)$, y que

$$
\begin{aligned}
\int_t^\infty \frac{S_T(u)}{1 - S_T(u)} dQ_{2N}(u) = &\left. \frac{S_T(u)}{1 - S_T(u)} Q_{2N}(u) \right]_t^\infty - \int_t^\infty Q_{2N}(u) d\left\{ \frac{S_T(u)}{1 - S_T(u)} \right\} \\
&= -Q_{2N}(t) \frac{S_T(t)}{1 - S_T(t)} - \int_t^\infty Q_{2N}(u) d\left\{ \frac{S_T(u)}{1 - S_T(u)} \right\},
\end{aligned}
$$

donde $V = S_T(u)/\left(1 - S_T(u)\right)$, $dU = dQ_{2N}(u)$, $dV = d\left\{ S_T(u) \left(1 - S_T(u)\right)^{-1} \right\}$ y $U = Q_{2N}(u)$, se obtiene finalmente la siguiente representación de $\Phi_{3N}(t)$

$$
\Phi_{3N}(t) = \int_t^\infty Q_{2N}(u) d\left\{ S_T(u) \left(1 - S_T(u)\right)^{-1} \right\} - S_T(t) \int_t^\infty Q_{2N}(u) d\left(1 - S_T(u)\right)^{-1}.
$$

Por consiguiente, sustituyendo $\Phi_{1N}(t)$ y $\Phi_{3N}(t)$ en la expresión (3.2),

$$
\widehat{S}_T(t) - S_T(t) = Q_{0N}(t) + Q_{1N}(t) - \sum_{j=1}^{6} \Phi_{jN}(t),
$$

y multiplicando por $\sqrt{N}$, resulta que, para todo $t \in [\tau_1, \tau_2]$,

$$
\begin{aligned}
\Psi_N(t) = &\sqrt{N} Q_{0N}(t) - \sqrt{N} \int_0^t W_{1N}(t,u) dS_1(u) + \sqrt{N} S_T(t) Q_{1N}(0^-) \\
&+ \sqrt{N} S_T(t) \int_0^t Q_{1N}(u) dS_T^{-1}(u) - \sqrt{N} \int_0^t W_{1N}(t,u) dQ_{1N}(u)
\end{aligned}
$$

$$-\sqrt{N}\int_t^\infty W_{2N}(t,u)dS_2(u)-\sqrt{N}\int_t^\infty Q_{2N}(u)d\left\{S_T(u)\left(1-S_T(u)\right)^{-1}\right\}$$

$$+\sqrt{N}S_T(t)\int_t^\infty Q_{2N}(u)d(1-S_T(u))^{-1}-\sqrt{N}\int_t^\infty W_{2N}(t,u)dQ_{2N}(u)$$

$$=U_{1N}(t)-U_{2N}(t)+U_{3N}(t)+U_{4N}(t)-R_{1N}(t)-U_{5N}(t)-U_{6N}(t)$$

$$+U_{7N}(t)-R_{2N}(t).$$

Si se define $b_T=\sup\{t\geq 0:S_T(t)>0\}$ y $b_X=\sup\{x\geq 0:S_X(x)>0\}$, entonces, por la acotación del soporte de X (o de T), $b=\min\{b_X,b_T\}$ es finito.

Para todo $u>b_X$, $S_2(u)=0$ y $\widehat{S}_2(u)=0$ $c.s.$; para todo $u>b_T$, $S_T(u)=0$ y $W_{2N}(t,u)=0$. Luego, para todo $u>b$, o bien $S_2(u)=0$ y $Q_{2N}(u)=0$ $c.s.$, o bien $S_T(u)=0$ y $W_{2N}(t,u)=0$. Por tanto, dado $t\in[\tau_1,\tau_2]$ arbitrario, si $t<b$, las integrales en (t,∞) que aparecen en la expresión de $\Psi_N(t)$ coinciden con las correspondientes integrales en $(t,b]$, mientras que, si $t\geq b$, dichas integrales se anulan. Obviamente, se supone que $\tau_1<b$; si ocurre lo contrario, entonces $U_{5N}(t)=U_{6N}(t)=U_{7N}(t)=R_{2N}(t)=0$.

Por otra parte, como S_T, S_X, S_Y son continuas, las subsupervivencias

$$S_0(t)=\Pr(T>t,\ X\leq T\leq Y)=-\int_t^\infty\{S_Y(u)-S_X(u)\}\,dS_T(u),$$

$$S_1(t)=\Pr(Y>t,\ T>Y)=-\int_t^\infty S_T(u)dS_Y(u)$$

y

$$S_2(t)=\Pr(X>t,\ T<X)=-\int_t^\infty\{1-S_T(u)\}\,dS_X(t)$$

también lo son.

A partir del lema 3.3, se obtiene que $Q_N(t)=\sqrt{N}\left(Q_{0N}(t),Q_{1N}(t),Q_{2N}(t)\right)$ converge débilmente a un proceso Gaussiano $G(t)=(G_0(t),G_1(t),G_2(t))$ de media cero con trayectorias continuas $c.s.$ (S_0, S_1 y S_2 son continuas). Mediante las construcciones de Skorohod (1956, sec. 3.1.1), se comprueba que existen nuevas versiones (esto es, con idénticas distribuciones de dimensión finita) de los procesos

Q_N y G (que, por simplicidad, se denotan de igual forma) que convergen casi seguramente respecto a la métrica de Skorohod en intervalos compactos. Además, como la convergencia a un límite continuo (como lo son G_0, G_1 y G_2 $c.s.$) en intervalos compactos respecto a la métrica de Skorohod es equivalente a la convergencia respecto a la norma del supremo (Prokhorov, 1956, teorema 1 del apéndice 1), puede suponerse que

$$\sup_{0\le t\le \tau_2}\left|\sqrt{N}Q_{jN}(t)-G_j(t)\right| \xrightarrow[N\to\infty]{c.s.} 0,\ j=0,1, \ \text{ y } \sup_{\tau_1\le t\le b}\left|\sqrt{N}Q_{2N}(t)-G_2(t)\right| \xrightarrow[N\to\infty]{c.s.} 0.$$

Considerando a $W_{1N}(t,u)$ y $W_{2N}(t,u)$ como funciones de $\widehat{\theta}$, y desarrollándolas en series de Taylor en un entorno de θ, se verifica que

$$\sqrt{N}W_{1N}(t,u)=\sqrt{N}(\widehat{\theta}-\theta)^{\mathrm{t}}\frac{\partial}{\partial\widehat{\theta}}\left[\frac{S(t;\widehat{\theta})}{S(u;\widehat{\theta})}\right]_{\widehat{\theta}=\theta}+o_p(1)\,,\quad 0\le u\le t\le \tau_2,$$

$$\sqrt{N}W_{2N}(t,u)=\sqrt{N}(\widehat{\theta}-\theta)^{\mathrm{t}}\frac{\partial}{\partial\widehat{\theta}}\left[\frac{S(t;\widehat{\theta})-S(u;\widehat{\theta})}{1-S(u;\widehat{\theta})}\right]_{\widehat{\theta}=\theta}+o_p(1)\,,\quad \tau_1\le t<u\le b,$$

de donde se deduce que

$$\left\{\sqrt{N}W_{1N}(t,u),\ 0\le u\le t\le\tau_2\right\} \quad \text{y} \quad \left\{\sqrt{N}W_{2N}(t,u),\ \tau_1\le t<u\le b\right\}$$

convergen débilmente a sendos procesos Gaussianos de media cero,

$$\{V_1(t,u),0\le u\le t\le\tau_2\} \quad \text{y} \quad \{V_2(t,u),\tau_1\le t<u\le b\}\,,$$

cuyas trayectorias son uniformemente continuas casi seguramente. Utilizando las construcciones de Skorohod y la continuidad $c.s$ de V_1 y V_2, se demuestra que existen nuevas versiones de los procesos $\sqrt{N}W_{1N}$, $\sqrt{N}W_{2N}$, V_1 y V_2 (con idéntica notación, por simplicidad) tales que

$$\sup_{0\le u\le t\le\tau_2}\left|\sqrt{N}W_{1N}(t,u)-V_1(t,u)\right| \xrightarrow[N\to\infty]{c.s.} 0\,;\ \sup_{\tau_1\le t<u\le b}\left|\sqrt{N}W_{2N}(t,u)-V_2(t,u)\right| \xrightarrow[N\to\infty]{c.s.} 0.$$

Considérese ahora el proceso $\Psi = \{\Psi(t), \tau_1 \leq t \leq \tau_2\}$, donde

$$\begin{aligned}
\Psi(t) &= G_0(t) - \int_0^t V_1(t,u)dS_1(u) + S_T(t)G_1(0^-) + S_T(t)\int_0^t G_1(u)dS_T^{-1}(u) \\
&\quad - \int_t^\infty V_2(t,u)dS_2(u) - \int_t^\infty G_2(u)d\left\{S_T(u)\left(1-S_T(u)\right)^{-1}\right\} \\
&\quad + S_T(t)\int_t^\infty G_2(u)d(1-S_T(u))^{-1} \\
&= U_1(t) - U_2(t) + U_3(t) + U_4(t) - U_5(t) - U_6(t) + U_7(t).
\end{aligned}$$

Este proceso es Gaussiano de media cero al ser un promedio pesado de procesos Gaussianos de media cero. Se demostrará que las nueva versión de Ψ_N converge $c.s.$ respecto a la norma del supremo en $[\tau_1, \tau_2]$ a la nueva versión de Ψ. De esta forma, la versión original de Ψ_N convergerá débilmente a la versión original de Ψ, ya que la convergencia casi segura implica la convergencia en distribución o débil y, por la construcción de Skorohod, las versiones original y nueva tienen las mismas distribuciones de dimensión finita, para cada N. Como Ψ tiene $c.s.$ trayectorias continuas, se verificará la convergencia débil tanto en el sentido de la métrica de Skorohod como en el sentido de la norma del supremo.

En primer lugar,

$$\sup_{\tau_1 \leq t \leq \tau_2} |U_{1N}(t) - U_1(t)| \leq \sup_{0 \leq t \leq \tau_2} \left|\sqrt{N}Q_{0N}(t) - G_0(t)\right| \xrightarrow[N\to\infty]{c.s.} 0,$$

$$\sup_{\tau_1 \leq t \leq \tau_2} |U_{2N}(t) - U_2(t)| \leq \sup_{0 \leq u \leq t \leq \tau_2} \left|\sqrt{N}W_{1N}(t,u) - V_1(t,u)\right| \left|\int_0^{\tau_2} dS_1(u)\right| \xrightarrow[N\to\infty]{c.s.} 0$$

y

$$\sup_{\tau_1 \leq t \leq \tau_2} |U_{3N}(t) - U_3(t)| \leq \left|\sqrt{N}Q_{1N}(0^-) - G_1(0^-)\right| \xrightarrow[N\to\infty]{c.s.} 0.$$

Asimismo, puesto que

$$\sup_{0 \leq t \leq \tau_2} \left|S_T(t)\int_0^t dS_T^{-1}(u)\right| \leq \int_0^{\tau_2} dS_T^{-1}(u) = S_T^{-1}(\tau_2) - S_T^{-1}(0^-),$$

se obtiene que

$$\sup_{\tau_1 \le t \le \tau_2} |U_{4N}(t) - U_4(t)| \le \sup_{0 \le t \le \tau_2} \left|\sqrt{N} Q_{1N}(t) - G_1(t)\right| \left\{S_T^{-1}(\tau_2) - 1\right\} \xrightarrow[N\to\infty]{c.s.} 0.$$

En segundo lugar, como

$$U_{5N}(t) - U_5(t) = \int_t^b \left\{\sqrt{N} W_{2N}(t,u) - V_2(t,u)\right\} dS_2(u),$$

$$U_{6N}(t) - U_6(t) = \int_t^b \left\{\sqrt{N} Q_{2N}(u) - G_2(t)\right\} d\left\{S_T(u)\left(1 - S_T(u)\right)^{-1}\right\}$$

y

$$U_{7N}(t) - U_7(t) = \int_t^b \left\{\sqrt{N} Q_{2N}(u) - G_2(t)\right\} d\left(1 - S_T(u)\right)^{-1}$$

si $\tau_1 \le t < b$, y estas expresiones se anulan cuando $t \ge b$, resulta que

$$\sup_{\tau_1 \le t \le \tau_2} |U_{5N}(t) - U_5(t)| \le \sup_{\tau_1 \le t < u \le b} \left|\sqrt{N} W_{2N}(t,u) - V_2(t,u)\right| \left|\int_{\tau_1}^b dS_2(u)\right| \xrightarrow[N\to\infty]{c.s.} 0,$$

$$\sup_{\tau_1 \le t \le \tau_2} |U_{6N}(t) - U_6(t)| \le \sup_{\tau_1 \le t \le b} \left|\sqrt{N} Q_{2N}(t) - G_2(t)\right| \left|\int_{\tau_1}^b d\left\{\frac{S_T(u)}{1 - S_T(u)}\right\}\right| \xrightarrow[N\to\infty]{c.s.} 0,$$

y

$$\sup_{\tau_1 \le t \le \tau_2} |U_{7N}(t) - U_7(t)| \le \sup_{\tau_1 \le t \le b} \left|\sqrt{N} Q_{2N}(t) - G_2(t)\right| \left|\int_{\tau_1}^b d\left\{\frac{1}{1 - S_T(u)}\right\}\right| \xrightarrow[N\to\infty]{c.s.} 0.$$

Nótese que

$$\left|\int_{\tau_1}^b d\left\{\frac{S_T(u)}{1 - S_T(u)}\right\}\right| = \frac{S_T(\tau_1)}{1 - S_T(\tau_1)} - \frac{S_T(b)}{1 - S_T(b)}$$

y

$$\left|\int_{\tau_1}^b d\left\{\frac{1}{1 - S_T(u)}\right\}\right| = \frac{1}{1 - S_T(\tau_1)} - \frac{1}{1 - S_T(b)}$$

están acotadas ya que $S_T(\tau_1) < 1$.

Por otra parte,

$$|R_{1N}(t)| \le \left|\int_0^t \left\{\sqrt{N} W_{1N}(t,u) - V_1(t,u)\right\} dQ_{1N}(u)\right| + \left|\int_0^t V_1(t,u) dQ_{1N}(u))\right|$$

si $\tau_1 \leq t \leq \tau_2$, y

$$|R_{2N}(t)| \leq \left| \int_t^b \left\{ \sqrt{N} W_{2N}(t,u) - V_2(t,u) \right\} dQ_{2N}(u) \right| + \left| \int_t^b V_2(t,u) dQ_{2N}(u)) \right|$$

si $\tau_1 \leq t < b$, y se anula si $t \geq b$. Los primeros términos de los segundos miembros de las expresiones anteriores son menores o iguales que

$$2 \cdot \sup_{0 \leq u \leq t \leq \tau_2} \left| \sqrt{N} W_{1N}(t,u) - V_1(t,u) \right| \quad \text{y} \quad 2 \cdot \sup_{\tau_1 \leq t < u \leq b} \left| \sqrt{N} W_{2N}(t,u) - V_2(t,u) \right|$$

respectivamente. Para acotar los segundos términos, se considera un subconjunto Ω_0 del espacio muestral subyacente Ω con $\Pr(\Omega_0) = 1$ donde $V_1(t,u)$ y $V_2(t,u)$ sean uniformemente continuas en $0 \leq u \leq t \leq \tau_2$ y en $\tau_1 \leq t < u \leq b$, respectivamente, y

$$\sup_{0 \leq t \leq \tau_2} |Q_{1N}(t)| \xrightarrow[N\to\infty]{c.s.} 0 \quad \text{y} \quad \sup_{\tau_1 \leq t \leq b} |Q_{2N}(t)| \xrightarrow[N\to\infty]{c.s.} 0.$$

De esta forma, dado un $\varepsilon > 0$ arbitrario, existirá una partición de $[0, \tau_2]$ en k intervalos, $I_0 = [\alpha_0, \alpha_1]$, $I_1 = (\alpha_1, \alpha_2]$, ..., $I_{k-1} = (\alpha_{k-1}, \alpha_k]$, y una partición de $[\tau_1, b]$ en m intervalos, $J_0 = [\beta_0, \beta_1]$, $J_1 = (\beta_1, \beta_2]$, ..., $J_{m-1} = (\beta_{m-1}, \beta_m]$, tales que

$$\sup_{\substack{(t,u) \in I_i \times I_j \\ u \leq t}} |V_1(t,u) - V_1(\alpha_i, \alpha_j)| < \varepsilon \quad \text{y} \quad \sup_{\substack{(t,u) \in J_i \times J_j \\ t < u}} |V_2(t,u) - V_2(\beta_i, \beta_j)| < \varepsilon.$$

Luego,

$$\sup_{0 \leq u \leq t \leq \tau_2} |V_1(t,u) - V_1^*(t,u)| < \varepsilon \quad \text{y} \quad \sup_{\tau_1 \leq t < u \leq b} |V_2(t,u) - V_2^*(t,u)| < \varepsilon,$$

donde

$$\begin{cases} V_1^*(t,u) = V_1(\alpha_i, \alpha_j) & \text{si} \quad (t,u) \in I_i \times I_j, \ \ u \leq t, \\ V_2^*(t,u) = V_2(\beta_i, \beta_j) & \text{si} \quad (t,u) \in J_i \times J_j, \ \ t < u. \end{cases}$$

Por consiguiente, como

$$\left| \int_0^t V_1(t,u) dQ_{1N}(u)) \right| \leq \left| \int_0^t \{ V_1(t,u) - V_1^*(t,u) \} dQ_{1N}(u) \right| + \left| \int_0^t V_1^*(t,u) dQ_{1N}(u) \right|$$

y

$$\left|\int_t^b V_2(t,u)dQ_{2N}(u))\right| \leq \left|\int_t^b \{V_2(t,u) - V_2^*(t,u)\}\, dQ_{2N}(u)\right| + \left|\int_t^b V_2^*(t,u)dQ_{2N}(u)\right|,$$

se obtiene que

$$\left|\int_0^t V_1(t,u)dQ_{1N}(u))\right| \leq \left(\varepsilon + \max_{i,j=0,..,k} |V_1(\alpha_i,\alpha_j)|\right) \left|Q_{1N}(t) - Q_{1N}(0^-)\right|$$

y

$$\left|\int_t^\infty V_2(t,u)dQ_{2N}(u))\right| \leq \left(\varepsilon + \max_{i,j=0,..,m} |V_2(\beta_i,\beta_j)|\right) |Q_{2N}(b) - Q_{2N}(t)|\,.$$

Por tanto,

$$\sup_{0\leq t\leq \tau_2} \left|\int_0^t V_1(t,u)dQ_{1N}(u))\right| \xrightarrow[N\to\infty]{c.s.} 0 \quad \text{y} \quad \sup_{\tau_1\leq t} \left|\int_t^\infty V_2(t,u)dQ_{2N}(u))\right| \xrightarrow[N\to\infty]{c.s.} 0,$$

con lo cual,

$$\sup_{0\leq t\leq \tau_2} |R_{1N}(t)| \xrightarrow[N\to\infty]{c.s.} 0 \quad \text{y} \quad \sup_{\tau_1\leq t} |R_{2N}(t)| \xrightarrow[N\to\infty]{c.s.} 0.$$

En definitiva, como los procesos $U_{1N}, ..., U_{7N}$ originales convergen débilmente respecto a la norma del supremo en el intervalo $[\tau_1, \tau_2]$ a sendos procesos Gaussianos de media cero, $U_1, ..., U_7$, y los procesos R_{1N} y R_{2N} convergen a 0 casi seguramente en esta métrica, se obtiene como resultado final que la versión original de Ψ_N converge débilmente a la versión original de Ψ.

$\square$

Observación

La evaluación de la estructura de la covarianza asintótica del proceso Ψ_N incluye el cálculo de las covarianzas asintóticas de $\left(\sqrt{N}\left(\widehat{S}_0(t) - S_0(t)\right), \sqrt{N}(\widehat{\theta} - \theta)\right)$, $\left(\sqrt{N}\left(\widehat{S}_1^*(t) - S_1(t)\right), \sqrt{N}(\widehat{\theta} - \theta)\right)$ y $\left(\sqrt{N}\left(\widehat{S}_2(t) - S_2(t)\right), \sqrt{N}(\widehat{\theta} - \theta)\right)$, las cuales son difíciles de obtener cuando se han utilizado técnicas iterativas para determinar al estimador $\widehat{\theta}$.

3.4 Experiencia computacional

En esta sección se presenta un estudio computacional diseñado para comparar el comportamiento de los estimadores de la función de supervivencia propuestos en este capítulo y en el anterior, y del estimador no paramétrico autoconsistente. Se analiza el efecto de la proporción de seguimiento, f^*, sobre la eficacia de los estimadores precedentes, según el nivel de censura y el tamaño muestral, en los mismos dos casos estudiados en el capítulo anterior.

La bondad de ajuste global de un estimador genérico $\widetilde{S}_T$ de la función de supervivencia S_T se mide a través del estimador del error cuadrático medio integrado, $E\left[\int_0^\infty \left\{\widetilde{S}_T(t) - S_T(t)\right\}^2 dt\right]$, definido por

$$ECMI = \frac{1}{kN}\sum_{i=1}^{k}\sum_{j=1}^{N}\left\{\frac{\widetilde{S}_T(Z_{ij}^-) + \widetilde{S}_T(Z_{ij})}{2} - S_T(Z_{ij})\right\}^2,$$

donde k es el número de muestras de la simulación, N es su tamaño, y Z_{ij} es la j-ésima observación de la i-ésima muestra. Además, para aquellos valores de τ tales que $S_T(\tau) = 0.1, 0.5$ y 0.9 (esto es, para $\tau = t_{0.90}$, $t_{0.50}$ y $t_{0.10}$), se estima el *sesgo*, $E\left[\widetilde{S}_T(\tau)\right] - S_T(\tau)$, y el *error cuadrático medio* (*ecm*), $E\left[\left\{\widetilde{S}_T(\tau) - S_T(\tau)\right\}^2\right]$, de $\widetilde{S}_T(\tau)$ mediante los estadísticos

$$\frac{1}{k}\sum_{i=1}^{k}\widetilde{S}_{T_{(i)}}(\tau) - S_T(\tau) \quad \text{y} \quad \frac{1}{k}\sum_{i=1}^{k}\left\{\widetilde{S}_{T_{(i)}}(\tau) - S_T(\tau)\right\}^2$$

respectivamente, donde $\widetilde{S}_{T_{(i)}}(\tau)$ es el valor del estimador en τ para la i-ésima muestra generada.

El estudio consiste en la simulación de 5000 muestras de tamaños $25, 50$ y 100 para diferentes niveles de censura (aprox. 10%, 25% y 50%) y f^* variando de 0.1 a 1.0 con incrementos de 0.3, y el análisis comparativo de los estimadores de la función de supervivencia siguientes:

1. Estimador paramétrico de máxima pseudo-verosimilitud propuesto en el capítulo anterior.

2. Estimador no paramétrico autoconsistente (de máxima pseudo-verosimilitud).
3. Estimador semiparamétrico considerado en este capítulo con $\widehat{\theta} = \theta^*$.

en los *Casos I* y *II* estudiados en el capítulo precedente.

En el *Caso I* se asume que T sigue una distribución exponencial de parámetro λ, y que la ventana de observación es el intervalo de tiempo que transcurre entre el primer y s-ésimo eventos en un proceso homogéneo de Poisson de parámetro μ. Como en el capítulo anterior, se supone que $\lambda = 1$.

La tabla 3.1 muestra los valores del $ECMI\,(\times 10^5)$ para los tres estimadores considerados. Como parece lógico, los valores del $ECMI$ se reducen cuando aumenta el tamaño muestral o la proporción de seguimiento, y crecen cuando lo hace el nivel de censura. En este caso, el comportamiento global del estimador paramétrico propuesto en el capítulo 2 es mejor que el de los otros dos, mientras que los $ECMI$ del estimador semiparamétrico considerado en este capítulo son ligeramente menores que los correspondientes al no paramétrico.

		p_1=0.051, p_2=0.050			p_1=0.118, p_2=0.125			p_1=0.237, p_2=0.250		
f^*	N	25	50	100	25	50	100	25	50	100
0.1	$ECMI1$	455	202	83	504	225	113	742	404	217
	$ECMI2$	776	369	179	944	441	214	1449	701	345
	$ECMI3$	743	353	170	815	383	185	999	481	241
0.4	$ECMI1$	339	165	78	406	197	99	568	286	141
	$ECMI2$	734	356	177	886	418	203	1229	580	279
	$ECMI3$	689	337	167	742	359	177	805	395	199
0.7	$ECMI1$	338	164	77	401	193	95	528	263	130
	$ECMI2$	733	356	177	880	415	201	1191	560	269
	$ECMI3$	687	336	167	735	356	176	780	383	194
1.0	$ECMI1$	338	164	77	398	192	94	515	257	126
	$ECMI2$	732	355	176	878	414	200	1178	554	266
	$ECMI3$	687	336	166	732	355	175	771	380	192

Tabla 3.1 : Resultados del $ECMI \times 10^5$ en el *Caso I*

Las tablas 3.2 y 3.3 contienen respectivamente las estimaciones obtenidas del $sesgo\,(\times 10^4)$ y $ecm\,(\times 10^4)$ de los tres estimadores de $S_T(\tau)$, para $\tau = t_{0.90}$, $t_{0.50}$ y

$t_{0.10}$, cuando el tamaño muestral $N = 100$. Tanto los valores del *sesgo* como del *ecm* tienden a decrecer si aumenta f^* o el tamaño muestral, o se reduce el nivel de censura.

		p_1=0.051, p_2=0.050			p_1=0.118, p_2=0.125			p_1=0.237, p_2=0.250		
f^*	$S_T(\tau)$	0.1	0.5	0.9	0.1	0.5	0.9	0.1	0.5	0.9
0.1	*sesgo*1	-8	16	8	-10	26	11	-14	41	18
	*sesgo*2	32	-6	-24	58	9	-45	122	-14	-105
	*sesgo*3	13	-16	-9	15	20	12	19	37	17
0.4	*sesgo*1	-7	13	7	-9	22	10	-12	33	13
	*sesgo*2	25	-5	-22	43	8	-41	89	-12	-73
	*sesgo*3	12	-14	-8	13	18	11	17	32	16
0.7	*sesgo*1	-7	12	6	-9	20	9	-12	31	12
	*sesgo*2	22	-5	-20	39	8	-39	67	-11	-58
	*sesgo*3	11	-14	-8	12	17	9	16	30	16
1.0	*sesgo*1	-6	11	6	-8	19	8	-11	29	12
	*sesgo*2	20	-5	-19	37	8	-36	59	-11	-56
	*sesgo*3	11	-14	-8	12	17	10	16	28	15

Tabla 3.2 : Resultados del *sesgo* $\times 10^4$ en el *Caso I* para $N = 100$.

		p_1=0.051, p_2=0.050			p_1=0.118, p_2=0.125			p_1=0.237, p_2=0.250		
f^*	$S_T(\tau)$	0.1	0.5	0.9	0.1	0.5	0.9	0.1	0.5	0.9
0.1	*ecm*1	7	14	2	8	17	3	12	28	6
	*ecm*2	11	26	12	34	29	18	63	31	32
	*ecm*3	10	24	9	13	28	12	21	29	13
0.4	*ecm*1	6	13	2	7	15	3	8	18	5
	*ecm*2	10	25	11	18	26	17	41	30	31
	*ecm*3	10	24	9	10	27	10	16	28	12
0.7	*ecm*1	6	12	1	6	14	2	7	17	4
	*ecm*2	10	24	11	14	26	16	31	29	31
	*ecm*3	9	23	8	9	26	9	15	27	12
1.0	*ecm*1	5	12	1	6	14	2	7	17	4
	*ecm*2	10	23	10	13	25	16	27	28	30
	*ecm*3	9	22	8	11	25	9	14	27	11

Tabla 3.3 : Resultados del *ecm* $\times 10^4$ en el *Caso I* para $N = 100$.

En cuanto al *sesgo*, la bondad del estimador paramétrico es superior a la de los otros dos cuando $S_T(\tau) = 0.1$ ó 0.9, mientras que el estimador no paramétrico obtiene los mejores resultados si $S_T(\tau) = 0.5$.

Respecto al *ecm*, el estimador paramétrico presenta el mejor comportamiento, mientras que el semiparamétrico es ligeramente superior al no paramétrico.

En el *Caso II* se considera que la amplitud aleatoria de la ventana de observación, $Y - X$, y las variables T y X siguen sendas distribuciones de Weibull. De forma explícita, $Y - X \sim W(\mu, r)$, $T \sim W(\lambda, r)$ y $X \sim W(\sigma, r)$, donde $\lambda = 1$ y $r \in \{0.5,\ 1.0,\ 2.0\}$.

Los valores del $ECMI\,(\times 10^5)$ cuando $N = 100$ de los tres estimadores considerados aparecen en la tabla 3.4. Al igual que ocurría antes, los $ECMI$ se reducen cuando aumenta el tamaño muestral o la proporción de seguimiento, y crecen si se incrementa el nivel de censura (sobre todo para el estimador paramétrico). Para un nivel de censura del 10%, el comportamiento global del estimador paramétrico propuesto en el capítulo 2 es mejor que el de los otros dos, mientras que el estimador semiparamétrico es ligeramente superior al no paramétrico. Cuando se asume una censura más alta (25% ó 50 %), la superioridad del estimador semiparamétrico es manifiesta.

		10% *censura*			25% *censura*			50% *censura*		
f^*	r	0.5	1.0	2.0	0.5	1.0	2.0	0.5	1.0	2.0
0.1	$ECMI1$	137	138	142	425	410	420	924	852	869
	$ECMI2$	193	194	194	248	248	261	400	387	390
	$ECMI3$	181	182	183	219	223	231	331	321	322
0.4	$ECMI1$	134	133	135	383	392	402	817	780	812
	$ECMI2$	184	185	187	216	217	226	301	299	297
	$ECMI3$	172	174	174	188	193	198	250	251	247
0.7	$ECMI1$	128	131	133	372	386	377	776	763	775
	$ECMI2$	183	183	184	212	211	220	287	282	281
	$ECMI3$	171	171	172	184	188	192	241	237	236
1.0	$ECMI1$	127	130	131	369	383	375	762	756	758
	$ECMI2$	180	182	184	209	208	215	282	278	274
	$ECMI3$	169	170	171	182	186	190	234	233	229

Tabla 3.4 : Resultados del $ECMI \times 10^5$ en el *Caso II* para $N = 100$.

		10% *censura*			25% *censura*			50% *censura*		
f^*	$S_T(\tau)$	0.1	0.5	0.9	0.1	0.5	0.9	0.1	0.5	0.9
0.1	*sesgo*1	19	16	12	-32	39	46	-87	-83	124
	*sesgo*2	44	25	-25	64	28	-33	82	44	-106
	*sesgo*3	13	20	14	15	22	19	57	40	66
0.4	*sesgo*1	15	10	9	-30	35	43	-76	-68	109
	*sesgo*2	33	21	-20	53	23	-29	70	38	-95
	*sesgo*3	11	14	11	13	17	15	51	35	54
0.7	*sesgo*1	14	9	8	-29	33	39	-68	-57	100
	*sesgo*2	32	20	-19	50	22	-27	63	34	-91
	*sesgo*3	10	13	10	12	16	14	49	32	50
1.0	*sesgo*1	13	8	6	-28	31	36	-63	-49	96
	*sesgo*2	30	19	-19	48	20	-26	59	31	-88
	*sesgo*3	10	13	9	12	16	13	47	30	46

Tabla 3.5 : Resultados del *sesgo* $\times 10^4$ en el *Caso II* para $r = 1$ y $N = 100$.

		10% *censura*			25% *censura*			50% *censura*		
f^*	$S_T(\tau)$	0.1	0.5	0.9	0.1	0.5	0.9	0.1	0.5	0.9
0.1	*ecm*1	9	20	7	20	52	25	45	99	55
	*ecm*2	12	27	11	19	30	16	43	39	33
	*ecm*3	11	26	10	16	28	11	32	34	16
0.4	*ecm*1	8	20	6	15	50	23	35	90	49
	*ecm*2	11	26	10	13	28	15	26	31	32
	*ecm*3	10	26	9	12	26	11	19	28	15
0.7	*ecm*1	7	19	6	12	49	22	29	87	47
	*ecm*2	10	26	10	12	27	15	21	30	31
	*ecm*3	10	25	9	11	26	10	16	27	15
1.0	*ecm*1	7	19	6	12	48	22	28	84	45
	*ecm*2	10	26	10	12	27	15	19	30	31
	*ecm*3	10	25	9	11	25	10	15	27	15

Tabla 3.6 : Resultados del *ecm* $\times 10^4$ en el *Caso II* para $r = 1$ y $N = 100$.

Las tablas 3.5 y 3.6 muestran respectivamente los resultados del *sesgo* $(\times 10^4)$ y *ecm* $(\times 10^4)$ de los tres estimadores de $S_T(\tau)$, para $\tau = t_{0.90}$, $t_{0.50}$ y $t_{0.10}$, cuando $r = 1$ y el tamaño muestral $N = 100$. Como en el caso anterior, los valores del *sesgo* y *ecm* tienden a decrecer si aumenta la proporción de seguimiento o se reduce el nivel de censura.

En relación al *sesgo*, el estimador semiparamétrico presenta los mejores resultados cuando la censura es del 25% ó 50%, mientras que la bondad del estimador paramétrico es superior a la de los otros dos si la censura es del 10%.

Resultados similares a los obtenidos en la tabla 3.4 respecto al $ECMI$ se deducen para los ecm de los tres estimadores de $S_T(\tau)$, para $\tau = t_{0.90}$, $t_{0.50}$ y $t_{0.10}$, en la tabla 3.6.

En general, como parece lógico, los valores del *sesgo*, ecm y $ECMI$ correspondientes a los tres estimadores que se han considerado en el estudio comparativo se reducen cuando aumenta la proporción de seguimiento o el tamaño muestral, y crecen cuando se incrementa el nivel de censura.

Cuando la censura es baja (10%) o la función de pseudo-verosimilitud es relativamente simple (*Caso I*), el estimador paramétrico de la supervivencia de máxima pseudo-verosimilitud presenta el mejor comportamiento, mientras que el estimador semiparamétrico es ligeramente superior al no paramétrico. Si la función de pseudo-verosimilitud es compleja (*Caso II*) y el nivel de censura es alto ($\geq$ 25%), la bondad del estimador semiparamétrico es apreciablemente mejor que la de los otros dos, quizás debido a la inestabilidad de la solución de la ecuación de verosimilitud. En esta situación, el estimador propuesto en este capítulo parece combinar las mejores propiedades de los estimadores paramétrico y no paramétrico para obtener los mejores resultados.

Capítulo 4

Estimación de la supervivencia con proceso de Dirichlet a priori

4.1 Introducción

La estimación de la función de supervivencia de la variable T, *tiempo de vida*, tiene a menudo un interés especial en los estudios de biometría y fiabilidad. Cuando algunas observaciones no son exactas, la función de supervivencia empírica no es apropiada para la estimación de la función de supervivencia teórica de T. Para resolver este problema se han utilizado varios puntos de vista alternativos de acuerdo con el modelo supuesto y el tipo de información disponible.

En ocasiones es posible obtener cierta información a priori basada en opiniones subjetivas o en experiencias pasadas. Las técnicas Bayesianas transforman este conocimiento previo en una distribución de probabilidad (denominada "a priori") sobre el espacio paramétrico en cuestión. Además, proporcionan una metodología para combinar la información a priori con la que origina la propia muestra y obtener así una distribución "a posteriori", y tienen en cuenta las consecuencias de las decisiones a través de una función de pérdida, cuyo valor esperado se trata de minimizar.

En el contexto Bayesiano no paramétrico, donde la propia función de supervivencia es el "parámetro" a estimar, no es necesario someterse a restricciones impuestas por consideraciones teóricas, en muchos casos artificiales y alejadas del problema real. Sin embargo, cuenta con la dificultad de definir una medida de probabilidad sobre un espacio funcional que asigne probabilidad uno al conjunto de las funciones de supervivencia.

En 1973, Ferguson aborda el problema de la estimación Bayesiana no paramétrica en un importante artículo, en el cual asume que existe un proceso de Dirichlet a priori sobre el espacio de las funciones de distribución. Posteriormente, Susarla y Van Ryzin (1976) analizan el caso de censura aleatoria a la derecha mediante la utilización de procesos de Dirichlet a priori, y obtienen el estimador Bayesiano no paramétrico de la función de supervivencia considerando que la pérdida es cuadrática. Ferguson y Phadia (1979) generalizan este resultado al caso de procesos "neutrales a la derecha" (véase Doksum, 1974) cuando las variables de censura son degeneradas.

Las enormes complicaciones que se presentan en el caso de censura doble han propiciado que no se haya obtenido ningún resultado relevante hasta el momento. Estas dificultades han fomentado la aparición de procedimientos Bayesianos alternativos, tales como el que aquí se presenta, en cada caso particular.

En este capítulo se considera el siguiente esquema de doble censura:

Sean $T_1, ..., T_N$ variables aleatorias independientes con idéntica función de supervivencia, $S_T(t) = \Pr(T > t)$, que la variable aleatoria T. Las variables T_i, $i = 1, ..., N$, representan el tiempo de vida (o de fallo) de N individuos. El mecanismo de censura doble supone que existen dos variables no negativas de censura X e Y tales que X, Y y T son independientes. Por consiguiente, también se consideran dos muestras aleatorias $X_1, ..., X_N$ e $Y_1, ..., Y_N$ de X e Y respectivamente. De este modo, las observaciones consisten solamente de vectores aleatorios independientes (Z_i, δ_i), $i = 1, ..., N$, con la misma distribución que (Z, δ), donde

$$Z = \max\{\min(T, Y), X\} \tag{4.1}$$

y

$$\delta = \begin{cases} 0 & si \quad X \leq T \leq Y \quad (no\ censura), \\ 1 & si \quad X \leq Y < T \quad (censura\ a\ la\ derecha), \\ 2 & si \quad T < X \leq Y \quad o \quad Y < X \quad (censura\ a\ la\ izquierda). \end{cases} \tag{4.2}$$

Morales et al. (1991) proporcionan un ejemplo en el que es válido este modelo. Los autores analizan la supervivencia de cierto tipo de árboles plantados en una granja y estudian una determinada causa de muerte. Cuando comenzó el estudio, varios árboles ya habían muerto debido a la causa de interés o a otras causas (censura a la izquierda). Algunos árboles murieron por la causa de interés durante el periodo de observación (no censura), mientras que otros murieron debido a otras causas, o bien permanecían vivos al final del estudio (censura a la derecha).

Morales et al. (1990, 1991) analizan esta situación de doble censura desde el punto de vista Bayesiano no paramétrico, supuesto que existe un conocimiento a priori a través de un proceso de Dirichlet sobre la variable observable Z y sobre la variable de interés T respectivamente. En cambio, aquí se considera que el conocimiento a priori viene dado por un proceso de Dirichlet sobre el vector observable completo (Z, δ). Esto conduce directamente a tener un conocimiento a priori sobre las funciones de subsupervivencia de Z.

De este modo, se supone que $\mathcal{P}$ es una medida de probabilidad aleatoria a priori sobre (Ω, Λ), donde $\Omega = \mathbb{R}^+ \times \{0,1,2\}$ y $\Lambda = B \times C$, siendo B la σ-álgebra de Borel restringida a $\mathbb{R}^+ = (0, \infty)$ y C la σ-álgebra de las partes de $\{0,1,2\}$. Además, se asume que la medida aleatoria $\mathcal{P}$ es un proceso de Dirichlet sobre (Ω, Λ) con parámetro α, que es una medida finita no nula sobre (Ω, Λ), y que $(Z_1, \delta_1), ..., (Z_N, \delta_N)$ es una muestra aleatoria de tamaño N de este proceso de Dirichlet.

Por tanto, el estimador Bayes no paramétrico de cada función de subsupervivencia de Z, $S_j(t) = \Pr(Z > t, \delta = j)$ para $j = 0, 1, 2$, será

$$\widetilde{S}_j(t) = \frac{\alpha\left((t, \infty), \{j\}\right) + \sum_{i=1}^{N} \mathrm{I}\left(Z_i > t, \delta_i = j\right)}{\alpha\left(\Omega\right) + N},$$

bajo la función de pérdida cuadrática

$$L(\widetilde{S}_j, S_j) = \int_0^\infty \left\{\widetilde{S}_j(t) - S_j(t)\right\}^2 dw(t) \ ,$$

donde w es una determinada función de peso (finita, no decreciente y no negativa).

Además, si la función $\alpha\left((\cdot,\infty),\{j\}\right)$, $j=0,1,2$, es continua por la derecha en $t=0$, el estimador Bayes de $p_j=\Pr(\delta=j)$, supuesto que la pérdida es cuadrática, viene dado por

$$\widetilde{p}_j=\lim_{t\downarrow 0}\widetilde{S}_j(t)=\frac{\alpha(\mathbb{R}^+,\{j\})+\sum_{i=1}^{N}\mathrm{I}\left(\delta_i=j\right)}{\alpha\left(\Omega\right)+N}.$$

Si se conoce la función de subsupervivencia $S_j^o(t)$, que representa la estimación a priori de $S_j(t)$, para $j=0,1,2$, y el valor de $\alpha\left(\Omega\right)$, el cual mide el grado de creencia en la elección a priori de las subsupervivencias, entonces el proceso de Dirichlet $\mathcal{P}$ queda perfectamente determinado, donde

$$\alpha_j\left(t,\infty\right)=\alpha\left(\left(t,\infty\right),\{j\}\right)=\alpha\left(\Omega\right)S_j^o(t)\text{ para }j=0,1,2.$$

Este capítulo presenta un procedimiento Bayesiano semiparamétrico para la estimación de una curva de supervivencia cuando los datos son doblemente censurados e incompletos (la variable Y es parcialmente observable) y la pérdida es cuadrática, supuesto el modelo de tasas de fallo proporcionales y que el conocimiento a priori se expresa mediante un proceso de Dirichlet sobre el vector aleatorio observado (Z,δ). Así, en la sección 4.2, para cada $t>0$ fijo, se obtienen los estimadores de las funciones de subsupervivencia de Z en t, $\widehat{S}_j(t)$, $j=0,1,2$, bajo pérdida cuadrática. Asimismo, se define el estimador pseudo-Bayesiano $\widehat{S}_1^*(t)$ que sustituirá a $\widehat{S}_1(t)$ cuando Y sea parcialmente observable. En la siguiente sección se estiman los parámetros del modelo de tasas de fallo proporcionales y se demuestra su normalidad asintótica y, a partir de las estimaciones previamente obtenidas, se construye una estimación semiparamétrica de $S_T(t)$ y se comprueba que el estimador propuesto, $\widehat{S}_T$, es realmente una función de supervivencia. Algunas propiedades asintóticas de este estimador, como su consistencia fuerte y la convergencia débil a un proceso Gaussiano, se demuestran en la sección 4.4. Finalmente, la sección 4.5 contiene los resultados de un estudio de simulación diseñado para analizar y comparar el comportamiento del estimador propuesto, e investigar el efecto que produce la existencia de una relación parcial de datos censurados por la derecha.

4.2 Estimación de las subsupervivencias

En vez de tratar el problema de la estimación de las funciones de subsupervivencia S_j, $j = 0, 1, 2$, se considera, para cada $t > 0$, la estimación de los parámetros positivos $S_j(t)$ y, a partir de éstos, se estiman las probabilidades p_j.

Por las propiedades del proceso de Dirichlet, la distribución de probabilidad a priori de $S_j(t)$ es $Beta$ con parámetros $\alpha_j(t,\infty)$ y $\alpha(\Omega) - \alpha_j(t,\infty)$, para $j = 0, 1, 2$. Por tanto, se obtiene el teorema siguiente:

Teorema 4.1

Para cualquier $t > 0$ arbitrario y $j = 0, 1, 2$, el estimador Bayes de $S_j(t)$ bajo pérdida cuadrática y distribución a priori $Beta(\alpha_j(t,\infty), \alpha(\Omega) - \alpha_j(t,\infty))$ viene dado por la expresión

$$\widehat{S}_j(t) = \varphi_N S_j^e(t) + (1 - \varphi_N) S_j^o(t), \tag{4.3}$$

donde $\varphi_N = N/(\alpha(\Omega) + N)$, y $S_j^e(t) = N^{-1} \sum_{i=1}^{N} \mathrm{I}(Z_i > t, \delta_i = j)$ es el valor de la correspondiente función de subsupervivencia empírica de Z en el punto t. Además, si la función $\alpha_j(\cdot,\infty)$, para $j = 0, 1, 2$, es continua por la derecha en $t = 0$, el estimador Bayes de $p_j = \Pr(\delta = j)$ viene dado por

$$\widehat{p}_j = \lim_{t \downarrow 0} \widehat{S}_j(t) = \varphi_N \frac{N_j}{N} + (1 - \varphi_N) S_j^o(0),$$

donde $N_j = \sum_{i=1}^{N} \mathrm{I}(\delta_i = j)$.

Demostración

Como $V_j(t) = \sum_{i=1}^{N} \mathrm{I}(Z_i > t, \delta_i = j)$ sigue una distribución binomial con parámetros N y $S_j(t)$, y la distribución a priori de $S_j(t)$ es $Beta$ con parámetros $\alpha_j(t,\infty)$ y $\alpha(\Omega) - \alpha_j(t,\infty)$, se verifica que la distribución a posteriori de $S_j(t)$, condicionado a $V_j(t) = v_j$, es $Beta(v_j + \alpha_j(t,\infty), N - v_j + \alpha(\Omega) - \alpha_j(t,\infty))$. Por tanto, como la pérdida es cuadrática, el estimador Bayes de $S_j(t)$ coincide con la media de su distribución a posteriori, es decir,

$$\widehat{S}_j(t) = E\left[S_j(t) \mid V_j(t) = v_j\right] = \frac{v_j + \alpha_j(t,\infty)}{\alpha(\Omega) + N} = \frac{N S_j^e(t) + \alpha(\Omega) S_j^o(t)}{\alpha(\Omega) + N}.$$

Luego, la expresión (4.3) es correcta. El estimador Bayes de las probabilidad p_j, para $j = 0, 1, 2$, se obtiene de forma similar teniendo en cuenta que $S_j^e(0) = N_j/N$.

□

Observación

- A pesar de que la técnica precedente puede utilizarse para hallar estimadores con cualquier proceso a priori "neutral a la derecha", el estimador así obtenido y el estimador Bayes en sentido global coinciden únicamente cuando se consideran procesos de Dirichlet.

A partir de este punto, como en los capítulos anteriores, se supone que los valores de la variable de censura a la derecha, Y, sólo se conocen para las N_{11} unidades con seguimiento posterior. En ese caso, el segundo término de la expresión de la función de supervivencia empírica

$$S_1^e(t) = \frac{1}{N}\sum_{i=1}^{N} I(Z_i > t, \delta_i = 1, F_i = 1) + \frac{1}{N}\sum_{i=1}^{N} I(Z_i > t, \delta_i = 1, F_i = 0)$$

se desconoce. Asumiendo que la distribución de las N_{10} observaciones censuradas por la derecha sin seguimiento coincide con la de las N_{11} con seguimiento, es lógico redistribuir equitativamente la masa de los N_{10} datos desconocidos de la variable Y entre las N_{11} observaciones conocidas de ésta. Por tanto, se sustituye el término desconocido

$$\sum_{i=1}^{N} I(Z_i > t, \delta_i = 1, F_i = 0)$$

por

$$\frac{N_{10}}{N_{11}}\sum_{i=1}^{N} I(Z_i > t, \delta_i = 1, F_i = 1)$$

para obtener así la función de subsupervivencia empírica modificada

$$S_1^{em}(t) = \left(1 + \frac{N_{10}}{N_{11}}\right)\frac{1}{N}\sum_{i=1}^{N} I(Z_i > t, \delta_i = 1, F_i = 1)$$

que se utilizará en lugar de $S_1^e(t)$ en las fórmulas precedentes. Por consiguiente, ya que no se conoce el estimador Bayesiano $\widehat{S}_1(t)$, se define el siguiente estimador pseudo-Bayesiano de $S_1(t)$

$$\widehat{S}_1^*(t) = \frac{N S_1^{em}(t) + \alpha(\Omega)\, S_1^o(t)}{\alpha(\Omega) + N}$$

que sustituirá a $\widehat{S}_1(t)$ cuando sea necesario.

Nótese que, utilizando el teorema de Glivenko-Cantelli, al igual que ocurría con S_1^e, la función de subsupervivencia empírica modificada S_1^{em} es un estimador uniformemente fuertemente consistente de S_1.

4.3 Estimación del tiempo de vida

En esta sección se obtiene una estimación semiparamétrica de $S_T(t)$ cuando T, X e Y son absolutamente continuas y existe proporcionalidad entre las tasas de fallo de las variables de censura y de la variable *tiempo de vida*, esto es, $h_X(t) = \beta h_T(t)$ y $h_Y(t) = \gamma h_T(t)$ en su campo de definición común. Este hecho puede expresarse en términos de las funciones de supervivencia como:

$$\begin{cases} S_X(t) = S_T^{\beta}(t)\,, \\ S_Y(t) = S_T^{\gamma}(t)\,, \end{cases} \tag{4.4}$$

para todo $t \geq 0$, donde β y γ son parámetros positivos desconocidos.

Las constantes β y γ pueden interpretarse como los parámetros de censura a la izquierda y derecha respectivamente. Así pues, $(\beta = +\infty, \gamma = 0)$ corresponde al caso de no censura, $(\beta = +\infty, \gamma \neq 0)$ y $(\beta \neq +\infty, \gamma = 0)$ a los casos de censura a la derecha y a la izquierda respectivamente, y finalmente $(\beta \neq +\infty, \gamma \neq 0)$ corresponde al caso de doble censura.

Nótese que, en el caso de censura a la derecha, el supuesto de tasas de fallo proporcionales es equivalente a la independencia de Z y δ (Allen, 1963). Sin embargo, no es cierto cuando aparece censura a la izquierda.

El esquema de doble censura aleatoria (4.1)-(4.2)-(4.4) ha sido utilizado recientemente por Morales et al. (1990, 1991), Bravo y Esteban (1993a) y Fernández et al. (1995a, 1997) y puede considerarse la extensión natural del modelo de Koziol-Green (KG) al caso de doble censura. Entre los autores que analizan el modelo KG se encuentran, entre otros, Koziol y Green (1976), Csörgo y Horváth (1981), Ebrahimi (1985) y Herbst (1992a, 1992b, 1993).

Dada la proporcionalidad de las tasas de fallo, resulta que

$$\begin{cases} p_0 = \displaystyle\int_0^\infty S_Y(t)\left\{1 - S_X(t)\right\} d\left\{1 - S_T(t)\right\} = \frac{1}{\gamma+1} - \frac{1}{\beta+\gamma+1}, \\ p_2 = \displaystyle\int_0^\infty \left\{1 - S_T(x) S_Y(x)\right\} d\left\{1 - S_X(x)\right\} = 1 - \frac{\beta}{\beta+\gamma+1}. \end{cases}$$

Reemplazando en las ecuaciones anteriores p_0 y p_2 por sus estimaciones Bayesianas $\widehat{p}_0$ y $\widehat{p}_2$, dadas en el teorema 4.1, se obtienen las estimaciones de β y γ siguientes:

$$\widehat{\beta} = \frac{(\widehat{p}_0 + \widehat{p}_1)^2}{\widehat{p}_0 \widehat{p}_2} \quad ; \quad \widehat{\gamma} = \frac{\widehat{p}_1}{\widehat{p}_0}. \tag{4.5}$$

Además, utilizando que (N_0, N_1, N_2) sigue una distribución multinomial $M(N, p_0, p_1, p_2)$, el hecho que $\varphi_{_N} \to 1$ cuando $N \to \infty$, y el método delta (Rao, 1973), puede demostrarse el teorema siguiente:

Teorema 4.2

La distribución de $\sqrt{N}\left(\widehat{\beta} - \beta, \widehat{\gamma} - \gamma\right)$ es asintóticamente normal bivariada con media cero y matriz de varianzas-covarianzas

$$\Gamma = \begin{pmatrix} \sigma_{\beta\beta} & \sigma_{\beta\gamma} \\ \sigma_{\beta\gamma} & \sigma_{\gamma\gamma} \end{pmatrix},$$

donde $\sigma_{\beta\beta} = \beta(1+\beta+\gamma)\{\beta + (1+\gamma)^2\}/(1+\gamma)$, $\sigma_{\gamma\gamma} = \gamma(1+\gamma)^2(1+\beta+\gamma)/\beta$ y $\sigma_{\beta\gamma} = \gamma(1+\gamma)(1+\beta+\gamma)$.

Demostración

Obsérvese que $\left(\widehat{\beta},\widehat{\gamma}\right)$ puede expresarse como una función de $\widehat{P}=(\widehat{p}_0,\widehat{p}_1,\widehat{p}_2)^{\mathrm{t}}$. Sea $\left(\widehat{\beta},\widehat{\gamma}\right)=F\left(\widehat{p}_0,\widehat{p}_1,\widehat{p}_2\right)=\left(F_1\left(\widehat{p}_0,\widehat{p}_1,\widehat{p}_2\right),F_2\left(\widehat{p}_0,\widehat{p}_1,\widehat{p}_2\right)\right)$, donde

$$F_1\left(\widehat{p}_0,\widehat{p}_1,\widehat{p}_2\right)=\frac{(\widehat{p}_0+\widehat{p}_1)^2}{\widehat{p}_0\widehat{p}_2}\ ;\quad F_2\left(\widehat{p}_0,\widehat{p}_1,\widehat{p}_2\right)=\frac{\widehat{p}_1}{\widehat{p}_0}.$$

Como F es derivable con continuidad en un entorno del punto $P=(p_0,p_1,p_2)^{\mathrm{t}}$, ya que p_0 y p_2 son positivos, puede considerarse el siguiente desarrollo de Taylor en un entorno de P

$$F\left(\widehat{p}_0,\widehat{p}_1,\widehat{p}_2\right)=F(p_0,p_1,p_2)+\sum_{j=0}^{2}\left(\widehat{p}_j-p_j\right)\left(\frac{\partial}{\partial p_j}F(p_0,p_1,p_2)+R_j\right),$$

donde $R_j\longrightarrow 0$ cuando $\widehat{p}_j\longrightarrow p_j$.

Se verifica que $\widehat{p}_j\xrightarrow[N\to\infty]{\Pr}p_j$, esto es, $\forall\,\varepsilon>0:\lim\limits_{N\to\infty}\Pr\left(|\widehat{p}_j-p_j|\geq\varepsilon\right)=0$. Nótese que, por la desigualdad de Chebychev,

$$\Pr\left(|\widehat{p}_j-p_j|\geq\varepsilon\right)=\Pr\left(\left|\varphi_{\scriptscriptstyle N}\left(\frac{N_j}{N}-p_j\right)+(1-\varphi_{\scriptscriptstyle N})\left(S_j^o(0)-p_j\right)\right|\geq\varepsilon\right)$$

es menor o igual que

$$\frac{E\left[\left\{\varphi_{\scriptscriptstyle N}\left(\frac{N_j}{N}-p_j\right)+(1-\varphi_{\scriptscriptstyle N})\left(S_j^o(0)-p_j\right)\right\}^2\right]}{\varepsilon^2}.$$

Luego,

$$\Pr\left(|\widehat{p}_j-p_j|\geq\varepsilon\right)\leq\frac{\varphi_{\scriptscriptstyle N}^2}{N\varepsilon^2}p_j(1-p_j)+\frac{(1-\varphi_{\scriptscriptstyle N})^2}{\varepsilon^2}\left(S_j^o(0)-p_j\right)^2\xrightarrow[N\to\infty]{}0.$$

Por consiguiente, $R_j\xrightarrow[N\to\infty]{\Pr}0$, para $j=0,1,2$, de donde resulta que

$$\sqrt{N}\left(\widehat{\beta}-\beta,\widehat{\gamma}-\gamma\right)=\sqrt{N}\left(F_1\left(\widehat{P}\right)-F_1\left(P\right),\,F_2\left(\widehat{P}\right)-F_2\left(P\right)\right)$$

está asintóticamente idénticamente distribuido como

$$\left(G_1^{\mathrm{t}}\sqrt{N}(\widehat{P}-P),\,G_2^{\mathrm{t}}\sqrt{N}(\widehat{P}-P)\right),$$

donde G_j es el gradiente de F_j en P, para $j = 1, 2$. Además, como (N_0, N_1, N_2) sigue una distribución multinomial $M(N, p_0, p_1, p_2)$ y $\varphi_N \to 1$ cuando $N \to \infty$ se obtiene, por el teorema central del límite, que $\sqrt{N}(\widehat{P} - P)$ es asintóticamente normal con media cero y matriz de varianzas-covarianzas $\Pi = (\pi_{ij})_{\substack{i=0,1,2 \\ j=0,1,2}}$, donde

$$\pi_{ij} = p_i \mathrm{I}(i = j) - p_i p_j\,, \quad \text{para } \; i,\, j = 0, 1, 2.$$

Por tanto, en definitiva, $\sqrt{N}\left(\widehat{\beta} - \beta, \widehat{\gamma} - \gamma\right)$ es asintóticamente normal bivariada con media cero y matriz de varianzas-covarianzas $\Gamma = (G_i^{\mathrm{t}} \Pi G_j)_{\substack{i=1,2 \\ j=1,2}}$.

Los elementos de la matriz Γ se obtienen con facilidad considerando que

$$G_1 = \left(\frac{p_0^2 - p_1^2}{p_0^2 p_2}, \frac{2(p_0 + p_1)}{p_0 p_2}, \frac{-(p_0 + p_1)^2}{p_0 p_2^2}\right)^{\mathrm{t}}, \quad G_2 = \left(\frac{-p_1}{p_0}, \frac{1}{p_0}, 0\right)^{\mathrm{t}},$$

$$p_0 = \frac{1}{\gamma + 1} - \frac{1}{\beta + \gamma + 1}\,, \quad p_1 = \frac{\gamma}{\gamma + 1} - \frac{\gamma}{\beta + \gamma + 1}\,, \quad \text{y} \quad p_2 = 1 - \frac{\beta}{\beta + \gamma + 1}.$$

□

Este resultado puede utilizarse para construir intervalos de confianza asintóticos de los parámetros β and γ.

Supuesto el modelo de tasas de fallo proporcionales, Bravo y Esteban (1993a) demostraron que las funciones de subsupervivencia de Z venían dadas por las expresiones siguientes:

$$S_0(t) = \Pr(Z > t, \delta = 0) = \frac{S_T^{\gamma+1}(t)}{\gamma + 1} - \frac{S_T^{\beta+\gamma+1}(t)}{\beta + \gamma + 1},$$

$$S_1(t) = \Pr(Z > t, \delta = 1) = \gamma\left(\frac{S_T^{\gamma+1}(t)}{\gamma + 1} - \frac{S_T^{\beta+\gamma+1}(t)}{\beta + \gamma + 1}\right), \tag{4.6}$$

$$S_2(t) = \Pr(Z > t, \delta = 2) = \beta\left(\frac{S_T^{\beta}(t)}{\beta} - \frac{S_T^{\beta+\gamma+1}(t)}{\beta + \gamma + 1}\right).$$

Reemplazando en las expresiones (4.6) $S_j(t)$, $j = 0, 1, 2$, por sus estimaciones respectivas $\widehat{S}_0(t)$, $\widehat{S}_1^*(t)$ y $\widehat{S}_2(t)$, definidas en el apartado anterior, y los parámetros β y γ por sus correspondientes estimaciones $\widehat{\beta}$ y $\widehat{\gamma}$, dadas en (4.5), se obtienen tres estimaciones de $S_T(t)$ que se denotan por $\widehat{S}_{T0}(t)$, $\widehat{S}_{T1}(t)$ y $\widehat{S}_{T2}(t)$. Este hecho lleva a proponer, de forma natural, el siguiente estimador de la función de supervivencia de T:

$$\widehat{S}_T(t) = \widehat{p}_0 \widehat{S}_{T0}(t) + \widehat{p}_1 \widehat{S}_{T1}(t) + \widehat{p}_2 \widehat{S}_{T2}(t). \tag{4.7}$$

Observaciones

- En el contexto no Bayesiano $(\alpha(\Omega) = 0)$ con datos completos $(f^* = 1)$, $\widehat{S}_T(t)$ resulta ser el estimador propuesto por Bravo y Esteban (1993a).

- La expresión (4.7) puede justificarse de la siguiente forma: la estructura del modelo que aparece en (4.4) implica que cualquier estimador natural de S_T debe presentar saltos tanto en los tiempos de vida como en los tiempos de censura, y esos saltos deben ser idénticos a los de las funciones de subsupervivencia estimadas de Z. Además, la probabilidad de que $\widehat{S}_0$, $\widehat{S}_1^*$ o $\widehat{S}_2$ provoque el salto de cualquier estimador de S_T puede estimarse por $\widehat{p}_0$, $\widehat{p}_1$ y $\widehat{p}_2$ respectivamente.

Ahora, se procede a comprobar que el estimador propuesto, $\widehat{S}_T$, está bien definido.

Teorema 4.3

El estimador $\widehat{S}_T$ definido en (4.7) es una función de supervivencia.

Demostración

La unicidad de $\widehat{S}_T$ se obtiene ya que

$$g_0(x) = \frac{x^{\widehat{\gamma}+1}}{\widehat{\gamma}+1} - \frac{x^{\widehat{\beta}+\widehat{\gamma}+1}}{\widehat{\beta}+\widehat{\gamma}+1},$$

$$g_1(x) = \widehat{\gamma} \left(\frac{x^{\widehat{\gamma}+1}}{\widehat{\gamma}+1} - \frac{x^{\widehat{\beta}+\widehat{\gamma}+1}}{\widehat{\beta}+\widehat{\gamma}+1} \right),$$

$$g_2(x) = \widehat{\beta} \left(\frac{x^{\widehat{\beta}}}{\widehat{\beta}} - \frac{x^{\widehat{\beta}+\widehat{\gamma}+1}}{\widehat{\beta}+\widehat{\gamma}+1} \right)$$

son funciones crecientes para $x \in [0,1]$.

A partir de las definiciones de $\widehat{S}_0$, $\widehat{S}_1^*$ y $\widehat{S}_2$, resulta que $\widehat{S}_T$ es continua por la derecha, $\widehat{S}_T(0) = 1$ y $\lim_{t \to +\infty} \widehat{S}_T(t) = 0$.

Finalmente, también $\widehat{S}_T$ es no creciente ya que

$$\widehat{S}_T(t) = \widehat{p}_0 g_0^{-1} \left\{ \widehat{S}_0(t) \right\} + \widehat{p}_1 g_1^{-1} \left\{ \widehat{S}_1^*(t) \right\} + \widehat{p}_2 g_2^{-1} \left\{ \widehat{S}_2(t) \right\},$$

donde g_j^{-1}, $j = 0, 1, 2$, son funciones crecientes, y $\widehat{S}_0$, $\widehat{S}_1^*$ y $\widehat{S}_2$ son no crecientes.

$\square$

El siguiente corolario proporciona las expresiones del estimador propuesto para los casos sin censura y con sólo censura a la derecha.

Corolario 4.1

(a) Si no existe censura y $\alpha_1(\mathbb{R}^+)=\alpha_2(\mathbb{R}^+)=0$, resulta que $\widehat{S}_T$ coincide con el estimador $\widehat{S}_0$ propuesto por Ferguson (1973).

(b) En el caso de censura a la derecha con $\alpha_2(\mathbb{R}^+) = 0$,

$$\widehat{S}_T(t) = \widehat{p}_0 \exp \left\{ \widehat{p}_0 \log \left(\widehat{S}_0(t)/\widehat{p}_0 \right) \right\} + \widehat{p}_1 \exp \left\{ \widehat{p}_0 \log \left(\widehat{S}_1^*(t)/\widehat{p}_1 \right) \right\}, \qquad (4.8)$$

que es la versión Bayesiana con información incompleta del estimador propuesto por Ebrahimi (1985).

Demostración

(a) El resultado se obtiene inmediatamente a partir de la definición de $\widehat{S}_T$ en (4.7).

(b) Si no existe censura a la izquierda, las soluciones de las ecuaciones

$$\widehat{S}_0(t) = \frac{1}{\widehat{\gamma}+1} S_T^{\widehat{\gamma}+1}(t) \quad \text{y} \quad \widehat{S}_1^*(t) = \frac{\widehat{\gamma}}{\widehat{\gamma}+1} S_T^{\widehat{\gamma}+1}(t)$$

son respectivamente

$$\widehat{S}_{T0}(t) = \left[(\widehat{\gamma}+1)\widehat{S}_0(t)\right]^{1/(\widehat{\gamma}+1)} \quad \text{y} \quad \widehat{S}_{T1}(t) = \left[(\widehat{\gamma}+1)\widehat{S}_1^*(t)/\widehat{\gamma}\right]^{1/(\widehat{\gamma}+1)}.$$

La expresión (4.8) puede comprobarse con facilidad utilizando que $\widehat{\gamma} = \widehat{p}_1/\widehat{p}_0$ y la definición de $\widehat{S}_T(t)$ en (4.7).

En este caso, si la información es completa (esto es, $f^* = 1$) y se reemplazan las subsupervivencias estimadas por las empíricas, y $\widehat{p}_j$ por N_j/N para $j = 0, 1$, se obtiene el estimador de Ebrahimi (1985).

□

4.4 Propiedades asintóticas del estimador

Teorema 4.4

El estimador $\widehat{S}_T$ definido en (4.7) es un estimador uniformemente fuertemente consistente de la función de supervivencia de T.

Demostración

A partir de la relación (4.5), $\widehat{S}_{T0}$, $\widehat{S}_{T1}$ y $\widehat{S}_{T2}$ pueden expresarse unívocamente en términos de $\widehat{S}_0(t)$, $\widehat{S}_1^*(t)$ y $\widehat{S}_2(t)$ respectivamente, y de las probabilidades estimadas $\widehat{p}_1$ y $\widehat{p}_2$. Por tanto:

$$\begin{aligned}\widehat{S}_T(t) &= (1-\widehat{p}_1-\widehat{p}_2)\Psi_0\left(\widehat{p}_1,\widehat{p}_2,\widehat{S}_0(t)\right) + \widehat{p}_1\Psi_1\left(\widehat{p}_1,\widehat{p}_2,\widehat{S}_1^*(t)\right) + \widehat{p}_2\Psi_2\left(\widehat{p}_1,\widehat{p}_2,\widehat{S}_2(t)\right)\\ &= \Psi\left(\widehat{p}_1,\widehat{p}_2,\widehat{S}_0(t),\widehat{S}_1^*(t),\widehat{S}_2(t)\right),\end{aligned}$$

donde $\Psi_j\left(\widehat{p}_1,\widehat{p}_2,\widehat{S}_j(t)\right) = \widehat{S}_{Tj}(t)$, para $j = 0, 2$, y $\Psi_1\left(\widehat{p}_1,\widehat{p}_2,\widehat{S}_1^*(t)\right) = \widehat{S}_{T1}(t)$.

Utilizando que $\varphi_{_N} = N/(N + \alpha(\Omega)) \to 1$ y que $1 + N_{10}/N_{11} \overset{c.s.}{\to} 1/f^*$ cuando $N \to \infty$, y el teorema de Glivenko-Cantelli, se obtiene que $\widehat{S}_0$, $\widehat{S}_1^*$ y $\widehat{S}_2$ son estimadores uniformemente fuertemente consistentes de S_0, S_1 y S_2 respectivamente. Además, como

$$\widehat{p}_j = \varphi_{_N} \frac{N_j}{N} + (1 - \varphi_{_N}) S_j^o(0) \underset{N\to\infty}{\overset{c.s.}{\longrightarrow}} p_j \,,\; j = 0, 1 \,,$$

ya que N_j/N converge casi seguramente a p_j por la ley fuerte de los grandes números, $\varphi_{_N} \to 1$ cuando $N \to \infty$, y las funciones Ψ_j , $j = 0, 1, 2$, son continuas y están definidas en un subconjunto del compacto

$$\{(q_1, q_2, q_3) : q_1 \geq 0,\; q_2 \geq 0,\; q_1 + q_2 \leq 1,\; 0 \leq q_3 \leq 1\} \,,$$

resulta que Ψ es suma de funciones uniformemente continuas. Por consiguiente, se verifica que $\widehat{S}_T(t) = \Psi\left(\widehat{p}_1, \widehat{p}_2, \widehat{S}_0(t), \widehat{S}_1^*(t), \widehat{S}_2(t)\right)$ converge casi seguramente a $\Psi(p_1, p_2, S_0(t), S_1(t), S_2(t)) = S_T(t)$ uniformemente en $t \geq 0$, esto es,

$$\Pr\left(\| \widehat{S}_T - S_T \|_\infty \underset{N\to\infty}{\longrightarrow} 0\right) = \Pr\left(\lim_{N\to\infty} \sup_{t\geq 0} \left|\widehat{S}_T(t) - S_T(t)\right| = 0\right) = 1.$$

□

De aquí en adelante, se considera que τ_1 y τ_2 son números reales que verifican

$$\inf\{t \geq 0 : S_T(t) < 1\} < \tau_1 < \tau_2 < \sup\{t \geq 0 : S_T(t) > 0\} \,.$$

Lema 4.1

Para cada $t \in [\tau_1, \tau_2]$ arbitrario, se define $\widehat{\xi}_t = \left(\, \widehat{p}_1, \widehat{p}_2, \widehat{S}_0(t), \widehat{S}_1^*(t), \widehat{S}_2(t)\,\right)^{\rm t}$ y $\xi_t = (\, p_1, p_2, S_0(t), S_1(t), S_2(t)\,)^{\rm t}$. Si $\tau_1 \leq t_1 < .... < t_m \leq \tau_2$, entonces $\sqrt{N}\left(\widehat{\xi}_{t_1} - \xi_{t_1}, ..., \widehat{\xi}_{t_m} - \xi_{t_m}\right)$ es asintóticamente normal con media cero y matriz de varianzas-covarianzas

$$\Sigma(t_1, .., t_m) = \left(\Sigma_{t_i t_j}\right)_{\substack{i=1,..,m,\\ j=1,..,m,}}$$

donde, para $\tau_1 \leq s \leq t \leq \tau_2$, la matriz Σ_{ts} se define como

$$\begin{pmatrix} p_1(1-p_1) & -p_1p_2 & -p_1S_0(s) & \dfrac{(1-p_1)S_{11}(s)}{f^*} & -p_1S_2(s) \\ -p_1p_2 & p_2(1-p_2) & -p_2S_0(s) & \dfrac{-p_2S_{11}(s)}{f^*} & (1-p_2)S_2(s) \\ -p_1S_0(t) & -p_2S_0(t) & S_0(t)\{1-S_0(s)\} & \dfrac{-S_0(t)S_{11}(s)}{f^*} & -S_0(t)S_2(s) \\ \dfrac{(1-p_1)S_{11}(t)}{f^*} & \dfrac{-p_2S_{11}(t)}{f^*} & \dfrac{-S_0(s)S_{11}(t)}{f^*} & \dfrac{S_{11}(t)\{1-S_{11}(s)\}}{f^{*2}} & -\dfrac{S_{11}(t)S_2(s)}{f^*} \\ -p_1S_2(t) & (1-p_2)S_2(t) & -S_0(s)S_2(t) & \dfrac{-S_1(s)S_2(t)}{f^*} & S_2(t)\{1-S_2(s)\} \end{pmatrix}, \quad (4.9)$$

la matriz Σ_{st} es la traspuesta de Σ_{ts} y $S_{11}(t) = \Pr(Z > t,\, \delta = 1,\, F = 1)$

Demostración

Nótese que los vectores

$$\mathrm{H_{jt}} = (\mathrm{I}(\delta_j = 1)\,, \mathrm{I}(\delta_j = 2), \mathrm{I}(Z_j > t, \delta_j = 0)\,, \mathrm{I}(Z_j > t, \delta_j = 1, F_j = 1)\,, \mathrm{I}(Z_j > t, \delta_j = 2))^{\mathrm{t}}\,,$$

para $j = 1, .., N$, son independientes e idénticamente distribuidos con media

$$\mu_t = (\, p_1, p_2, S_0(t), S_{11}(t), S_2(t)\,)^{\mathrm{t}}\,,$$

y que

$$\widehat{\xi}_t = \varphi_{_N} A D_t + (1 - \varphi_{_N}) C_t\,,$$

donde

$$A = diag(1, 1, 1, 1 + N_{10}/N_{11}, 1),\;\; D_t = \frac{1}{N}\sum_{j=1}^{N} H_{jt}$$

y

$$C_t = (\, S_1^o(0),\, S_2^o(0),\, S_0^o(t),\, S_1^o(t),\, S_2^o(t)\,)^{\mathrm{t}}\,.$$

Si $\tau_1 \leq t_1 < ... < t_m \leq \tau_2$, entonces

$$\left(\widehat{\xi}_{t_1}, ..., \widehat{\xi}_{t_m}\right) = \varphi_{_N} A\,(D_{t_1}, ..., D_{t_m}) + (1 - \varphi_{_N})\,(C_{t_1}, ..., C_{t_m})\,.$$

Como $\varphi_{_N} \underset{N\to\infty}{\longrightarrow} 1$ y $1 + N_{10}/N_{11} \underset{N\to\infty}{\overset{c.\,s.}{\longrightarrow}} 1/f^*$, resulta que el vector $\left(\widehat{\xi}_{t_1}, ..., \widehat{\xi}_{t_m}\right)$ está asintóticamente idénticamente distribuido como $A^*\,(D_{t_1}, ..., D_{t_m})$, donde la matriz $A^* = diag(1, 1, 1, 1/f^*, 1)$.

Además, $(D_{t_1}, ..., D_{t_m}) = \dfrac{1}{N} \sum\limits_{j=1}^{N} H_j(t_1, ..., t_m)$, donde

$$H_j(t_1, ..., t_m) = (H_{jt_1}, ..., H_{jt_m}) \ , \ j = 1, .., N,$$

son independientes e idénticamente distribuidos. Luego, por el teorema central del límite multidimensional, se obtiene que $\sqrt{N}\,(D_{t_1}, ..., D_{t_m})$ es asintóticamente normal de media $\sqrt{N}\,(\mu_{t_1}, ..., \mu_{t_m})$ y matriz de varianzas-covarianzas

$$\Omega\,(t_1, .., t_m) = \left(\Omega_{t_i t_j}\right)_{\substack{i=1,..,m, \\ j=1,..,m,}}$$

donde, para $\tau_1 \le s \le t \le \tau_2$, $\Omega_{ts} = diag(\mu_t) - \mu_t \mu_s^{\mathrm{t}} + R_{ts}$, $R_{ts} = (r_{ij})_{5\times5}$ es una matriz 5×5 cuyos únicos elementos no nulos son $r_{14} = S_{11}(s)$, $r_{25} = S_2(s)$, $r_{41} = S_{11}(t)$ y $r_{52} = S_2(t)$, $\Omega_{st} = \Omega_{ts}^{\mathrm{t}}$ y $R_{st} = R_{ts}^{\mathrm{t}}$. Por tanto, como $\xi_{t_i} = A^* \mu_{t_i}$ y $\Sigma_{t_i t_j} = A^* \cdot \Omega_{t_i t_j} \cdot A^*$, para $i, j = 1, .., m$, se obtiene en definitiva que

$$\sqrt{N}\left(\widehat{\xi}_{t_1} - \xi_{t_1}, ..., \widehat{\xi}_{t_m} - \xi_{t_m}\right) \xrightarrow[N\to\infty]{d} N\left(0, \Sigma\,(t_1, .., t_m)\right),$$

donde $\Sigma\,(t_1, .., t_m) = \left(\Sigma_{t_i t_j}\right)_{\substack{i=1,..,m, \\ j=1,..,m,}}$. Nótese que $\Sigma_{t_i t_j}$ puede expresarse como

$$\Sigma_{t_i t_j} = diag(1, 1, 1, 1/f^*, 1)\left\{diag\left(\mu_{\max\{t_i, t_j\}}\right) - \mu_{t_i}\mu_{t_j}^{\mathrm{t}} + R_{t_i t_j}\right\} diag(1, 1, 1, 1/f^*, 1).$$

Sólamente se hallarán algunos términos de la matriz $\Omega_{ts} = \left(\omega_{ij}^{ts}\right)_{\substack{i=1,..,5 \\ j=1,..,5}}$. El resto se calcula de forma similar. Por ejemplo:

$$\omega_{11}^{ts} = var\left[\mathrm{I}(\delta = 1)\right] = E\left[\mathrm{I}(\delta = 1)\right] - \left\{E\left[\mathrm{I}(\delta = 1)\right]\right\}^2 = p_1 - p_1^2,$$

$$\begin{aligned}
\omega_{14}^{ts} &= cov\left[\mathrm{I}(\delta = 1), \mathrm{I}(Z > s, \delta = 1, F = 1)\right] \\
&= E\left[\mathrm{I}(\delta = 1)\mathrm{I}(Z > s, \delta = 1, F = 1)\right] - E\left[\mathrm{I}(\delta = 1)\right] E\left[\mathrm{I}(Z > s, \delta = 1, F = 1)\right] \\
&= E\left[\mathrm{I}(Z > s, \delta = 1, F = 1)\right](1 - p_1) = (1 - p_1) S_{11}(s),
\end{aligned}$$

$$\begin{aligned}
\omega_{23}^{ts} &= cov\left[\mathrm{I}(\delta = 2), \mathrm{I}(Z > s, \delta = 0)\right] \\
&= E\left[\mathrm{I}(\delta = 2)\mathrm{I}(Z > s, \delta = 0)\right] - E\left[\mathrm{I}(\delta = 2)\right] E\left[\mathrm{I}(Z > s, \delta = 0)\right] \\
&= 0 - \Pr(\delta = 2)\Pr(Z > s, \delta = 0) = -p_2 S_0(s),
\end{aligned}$$

$$\omega_{35}^{ts} = cov\left[\mathrm{I}(Z > t, \delta = 0), \mathrm{I}(Z > s, \delta = 2)\right]$$

$$= E\left[\mathrm{I}(Z > t, \delta = 0)\mathrm{I}(Z > s, \delta = 2)\right] - E\left[\mathrm{I}(Z > t, \delta = 0)\right] E\left[\mathrm{I}(Z > s, \delta = 2)\right]$$

$$= 0 - S_0(t)S_2(s).$$

y, en último lugar,

$$\omega_{44}^{ts} = cov\left[\mathrm{I}(Z > t, \delta = 1, F = 1), \mathrm{I}(Z > s, \delta = 1, F = 1)\right] = S_{11}(t) - S_{11}(t)S_{11}(s).$$

$\square$

Teorema 4.5

El proceso $\{W_N(t), \tau_1 \leq t \leq \tau_2\}$, donde $W_N(t) = \sqrt{N}\left(\widehat{S}_T(t) - S_T(t)\right)$, converge débilmente a un proceso Gaussiano de media cero.

Demostración

Para cada $t \in [\tau_1, \tau_2]$ arbitrario y fijo, como se comprobó en el teorema 4.4, se verifica que $\widehat{S}_T(t)$ puede expresarse como una función de $\widehat{\xi}_t$, esto es, $\widehat{S}_T(t) = \Psi(\widehat{\xi}_t)$.

Como Ψ es diferenciable con continuidad en un entorno del punto ξ_t (ya que $0 < S_j(t) < p_j$, $j = 0, 1, 2$), puede considerarse el desarrollo de Taylor en dicho punto

$$\widehat{S}_T(t) = \Psi(\widehat{\xi}_t) = \Psi(\xi_t) + \left(G_t^{\mathrm{t}} + R_N^{\mathrm{t}}\right)(\widehat{\xi}_t - \xi_t),$$

donde G_t denota al gradiente de Ψ en ξ_t y $R_N = (R_{1N}, ..., R_{5N})^{\mathrm{t}}$ converge a cero cuando $\widehat{\xi}_t$ converge a ξ_t. Además, puesto que $\widehat{\xi}_t$ converge en probabilidad a ξ_t, ya que $\sqrt{N}(\widehat{\xi}_t - \xi_t)$ es asintóticamente normal de media cero por el lema 4.1, se verifica que $R_N \xrightarrow[N\to\infty]{\mathrm{Pr}} 0$. Por tanto,

$$W_N(t) = \sqrt{N}\left(\widehat{S}_T(t) - S_T(t)\right) = \sqrt{N}\left(\Psi(\widehat{\xi}_t) - \Psi(\xi_t)\right)$$

está asintóticamente idénticamente distribuido como $G_t^{\mathrm{t}}\sqrt{N}(\widehat{\xi}_t - \xi_t)$, y éste, por el lema 4.1, es asintóticamente normal con media cero y matriz de varianzas-covarianzas $G_t^{\mathrm{t}}\Sigma_{tt}G_t$. La matriz Σ_{tt} puede obtenerse con facilidad de acuerdo con la expresión (4.9).

Análogamente, en el caso m-dimensional, se obtiene que, para cualquier $\tau_1 \leq t_1 < ... < t_m \leq \tau_2$, $(W_N(t_1), ..., W_N(t_m))$ está asintóticamente idénticamente distribuido como $\sqrt{N}\left(G_{t_1}^{\mathrm{t}}(\widehat{\xi}_{t_1} - \xi_{t_1}), ..., G_{t_m}^{\mathrm{t}}(\widehat{\xi}_{t_m} - \xi_{t_m})\right)$ que, por el lema 4.1, es asintóticamente normal con media cero y matriz de varianzas-covarianzas

$$\Phi\left(t_1, ..., t_m\right) = \left(G_{t_i}^{\mathrm{t}} \Sigma_{t_i t_j} G_{t_j}\right)_{\substack{i=1,..,m \\ j=1,..,m}} .$$

$\square$

Observaciones

- Dada la normalidad asintótica de $\sqrt{N}\left(\widehat{S}_T(t) - S_T(t)\right)$, es posible construir intervalos de confianza asintóticos para $S_T(t)$ y efectuar contrastes de hipótesis sobre el valor de éste.

- Aunque el proceso considerado, $\{W_N(t), \tau_1 \leq t \leq \tau_2\}$, es asintóticamente normal, no tiene asintóticamente incrementos independientes. Basta observar que, para $\tau_1 \leq s < t \leq \tau_2$, se verifica que

$$\begin{aligned}\lim_{N\to\infty} cov\left[W_N(t) - W_N(s), W_N(s)\right] &= \lim_{N\to\infty} cov\left[W_N(t), W_N(s)\right] - \lim_{N\to\infty} var\left[W_N(s)\right] \\ &= G_t^{\mathrm{t}}\Sigma_{ts}G_s - G_s^{\mathrm{t}}\Sigma_{ss}G_s \neq 0.\end{aligned}$$

4.5 Experiencia computacional

En esta sección se presenta un estudio computacional diseñado para analizar el comportamiento del estimador de la función de supervivencia $\widehat{S}_T$ definido en (4.7). En primer lugar, se pretende estudiar la bondad de ajuste de $\widehat{S}_T$, analizar su robustez frente a violaciones de los supuestos del modelo, y compararlo con el estimador no paramétrico autoconsistente (Turnbull, 1974; Mykland y Ren, 1996) y el de Morales et al. (1990) considerando datos completos (esto es, con $f^* = 1$). En segundo lugar, se evalúa el efecto de f^* en el comportamiento del estimador propuesto $\widehat{S}_T$.

La bondad de ajuste global de un estimador genérico $\widetilde{S}_T$ de la función de supervivencia S_T se mide a través del estimador del error cuadrático medio integrado,

$E\left[\int_0^\infty \left\{\widetilde{S}_T(t) - S_T(t)\right\}^2 dt\right]$, definido por

$$ECMI = \frac{1}{kN}\sum_{i=1}^{k}\sum_{j=1}^{N}\left\{\frac{\widetilde{S}_T(Z_{ij}^-) + \widetilde{S}_T(Z_{ij})}{2} - S_T(Z_{ij})\right\}^2$$

donde k es el número de muestras de la simulación, N es su tamaño, y Z_{ij} es la j-ésima observación de la i-ésima muestra.

Además, para aquellos valores de τ tales que $S_T(\tau) = 0.1, 0.5$ y 0.9 (esto es, para $\tau = t_{0.9}$, $t_{0.5}$ y $t_{0.1}$), se estiman el *sesgo*, $E\left[\widetilde{S}_T(\tau)\right] - S_T(\tau)$, y el error cuadrático medio (*ecm*), $E\left[\left\{\widetilde{S}_T(\tau) - S_T(\tau)\right\}^2\right]$, de $\widetilde{S}_T(\tau)$ mediante los estadísticos

$$B = \frac{1}{k}\sum_{i=1}^{k}\widetilde{S}_{T_{(i)}}(\tau) - S_T(\tau) \quad \text{y} \quad m = \frac{1}{k}\sum_{i=1}^{k}\left\{\widetilde{S}_{T_{(i)}}(\tau) - S_T(\tau)\right\}^2$$

respectivamente, donde $\widetilde{S}_{T_{(i)}}(\tau)$ es el valor del estimador en τ para la i-ésima muestra de la simulación.

El estudio consiste en la simulación de 2000 muestras de tamaños 100 y 25 con un 10% ó 50% de censura, considerando que las tasas de fallo son proporcionales y distribuciones de Weibull. Específicamente, se supone que las funciones de supervivencia de T, X, e Y son

$$S_T(t) = \exp\left\{-(t/\theta)^\lambda\right\}, \quad S_X(x) = \exp\left\{-\gamma(x/\theta)^\lambda\right\} \quad \text{y} \quad S_Y(t) = \exp\left\{-\beta(x/\theta)^\lambda\right\},$$

para t, x, $y \geq 0$ respectivamente, donde β y γ son los parámetros de censura que se obtienen como solución del siguiente sistema:

$$\begin{cases} p_0 = \dfrac{1}{\gamma+1} - \dfrac{1}{\beta+\gamma+1}, \\[2ex] p_2 = \dfrac{1-p_0}{2} = 1 - \dfrac{\beta}{\beta+\gamma+1}, \end{cases}$$

esto es, $\beta = (1+p_0)^2/\left\{2p_0(1-p_0)\right\}$ y $\gamma = (1-p_0)/\left(2p_0\right)$. Consecuentemente, para un bajo nivel de censura ($p_0 = 0.9$) se obtiene que $\beta = 361/18$ y $\gamma = 1/18$, mientras que para un alto nivel de censura ($p_0 = 0.5$), $\beta = 9/2$ y $\gamma = 1/2$. En ambos casos, θ y λ toman los valores 0.5, 1.0, 1.5 y 2.0.

La elección del parámetro de Dirichlet α es una cuestión importante. Si se cuenta con datos previos, pueden utilizarse para estimar α (en un contexto Bayes empírico). En ausencia de tales datos, la elección del parámetro sólo puede hacerse subjetivamente. Es conveniente especificar una familia paramétrica a la que pertenezca α. En nuestro caso, un método apropiado sería el estimar subjetivamente ciertos percentiles de las subsupervivencias a priori, y de ahí elegir los parámetros de forma que los percentiles de las subsupervivencias sean tan próximos como sea posible a los percentiles estimados. No obstante, en este estudio se considera que las subsupervivencias a priori son correctas o uniformes.

Tanto para un nivel de censura bajo como alto, se analizan los cuatro casos correspondientes a unas subsupervivencias correctas o uniformes, y un grado de creencia en las elecciones a priori fuerte ($\alpha(\Omega) = 5N$) o débil ($\alpha(\Omega) = N/5$).

La tabla 4.1 muestra los valores del $ECMI$ de $\widehat{S}_T$ cuando se asume una correcta elección de las subsupervivencias a priori. Como era de esperar, los valores del $ECMI$ crecen cuando disminuye el tamaño muestral, y decrecen cuando se considera un menor nivel de censura. Asimismo, los valores del $ECMI$ son menores cuando se supone una fuerte creencia en las subsupervivencias a priori.

	θ	0.5		1.0		1.5		2.0	
λ	$\alpha(\Omega)$	*Baja cens.*	*Alta cens.*	*Baja cens.*	*Alta cens.*	*Baja cens.*	*Alta cens.*	*Baja cens.*	*Alta cens.*
0.5	$n/5$	16-55	20-81	16-56	22-80	18-51	21-87	16-54	19-85
	$5n$	9-21	11-61	9-22	10-61	9-21	11-63	9-22	10-59
1.0	$n/5$	16-50	19-90	18-50	21-74	16-54	20-76	16-52	21-73
	$5n$	9-21	10-60	9-21	10-55	9-21	10-57	9-21	10-57
1.5	$n/5$	16-54	19-74	18-55	19-75	16-56	19-73	17-54	18-80
	$5n$	9-21	10-59	9-21	10-59	9-22	10-57	9-23	10-59
2.0	$n/5$	17-50	19-70	16-52	19-85	17-53	19-72	16-51	18-76
	$5n$	9-21	10-59	9-21	10-59	9-21	10-58	9-20	10-58

Tabla 4.1 : Resultados del $ECMI \times 10^4$ de $\widehat{S}_T$ con subsupervivencias a priori correctas para $N = 100$ - 25.

	θ	0.5		1.0		1.5		2.0	
	$\alpha\,(\Omega)$	$n/5$	$5n$	$n/5$	$5n$	$n/5$	$5n$	$n/5$	$5n$
λ	$S_T(\tau)$	$B\,(m)$	$B\,(m)$	$B\,(m)$	$B\,(m)$	$B\,(m)$	$B\,(m)$	$B\,(m)$	$B\,(m)$
0.5	0.1	-33 (17)	-53 (30)	-31 (15)	-51 (28)	-32 (18)	-51 (28)	-33 (17)	-51 (28)
	0.5	-11 (18)	-31 (12)	-16 (16)	-32 (12)	-14 (19)	-32 (12)	-12 (17)	-31 (12)
	0.9	0 (6)	-11 (4)	-3 (6)	-10 (4)	-2 (7)	-12 (4)	-4 (7)	-13 (4)
1.0	0.1	-34 (18)	-52 (29)	-34 (18)	-52 (28)	-32 (16)	-50 (27)	-32 (17)	-52 (28)
	0.5	-13 (17)	-31 (12)	-15 (20)	-30 (11)	-15 (18)	-31 (12)	-9 (16)	-30 (11)
	0.9	2 (7)	-9 (3)	-1 (7)	-11 (4)	-4 (7)	-12 (5)	1 (7)	-11 (4)
1.5	0.1	-31 (17)	-51 (28)	-33 (17)	-51 (28)	-32 (16)	-52 (29)	-31 (16)	-51 (29)
	0.5	-13 (17)	-30 (11)	-11 (21)	-29 (11)	-11 (19)	-31 (12)	-10 (20)	-29 (10)
	0.9	-1 (5)	-10 (3)	-5 (7)	-10 (4)	-2 (6)	-12 (4)	0 (7)	-11 (4)
2.0	0.1	-35 (19)	-53 (30)	-33 (17)	-51 (28)	-33 (18)	-52 (29)	-32 (19)	-53 (30)
	0.5	-18 (17)	-31 (11)	-12 (18)	-29 (10)	-15 (20)	-31 (12)	-11 (17)	-30 (12)
	0.9	-3 (5)	-13 (4)	0 (6)	-11 (4)	0 (6)	-10 (4)	-2 (6)	-11 (4)

Tabla 4.2 : Resultados del *sesgo* $\times\, 10^3$ (*ecm* $\times\, 10^4$) de $\widehat{S}_T(\tau)$ en el caso de baja censura con subsupervivencias a priori correctas para $N = 100$.

	θ	0.5		1.0		1.5		2.0	
	$\alpha\,(\Omega)$	$n/5$	$5n$	$n/5$	$5n$	$n/5$	$5n$	$n/5$	$5n$
λ	$S_T(\tau)$	$B\,(m)$	$B\,(m)$	$B\,(m)$	$B\,(m)$	$B\,(m)$	$B\,(m)$	$B\,(m)$	$B\,(m)$
0.5	0.1	-35 (23)	-80 (70)	-37 (27)	-83 (76)	-36 (26)	-81 (73)	-33 (23)	-78 (68)
	0.5	0 (22)	-18 (5)	-3 (25)	-19 (6)	2 (24)	-18 (5)	6 (21)	-17 (5)
	0.9	-12 (11)	-29 (14)	-16 (10)	-32 (14)	-8 (9)	-27 (11)	-9 (9)	-29 (12)
1.0	0.1	-33 (24)	-79 (69)	-36 (26)	-81 (73)	-32 (23)	-77 (67)	-36 (24)	-84 (78)
	0.5	5 (21)	-18 (5)	-5 (21)	-20 (6)	0 (21)	-19 (5)	-1 (23)	-18 (5)
	0.9	-10 (10)	-29 (13)	-15 (10)	-33 (15)	-16 (11)	-34 (16)	-16 (11)	-33 (15)
1.5	0.1	-32 (22)	-78 (69)	-35 (24)	-81 (73)	-34 (24)	-81 (74)	-32 (23)	-82 (74)
	0.5	1 (18)	-18 (5)	-1 (21)	-19 (6)	-1 (19)	-21 (6)	2 (20)	-19 (6)
	0.9	-12 (10)	-29 (13)	-18 (11)	-33 (15)	-14 (11)	-32 (15)	-13 (10)	-31 (14)
2.0	0.1	-34 (24)	-80 (71)	-34 (23)	-82 (75)	-35 (24)	-83 (77)	-33 (20)	-78 (69)
	0.5	3 (20)	-18 (6)	1 (22)	-18 (5)	-2 (20)	-19 (5)	5 (19)	-17 (5)
	0.9	-14 (9)	-32 (14)	-17 (11)	-33 (15)	-16 (10)	-32 (14)	-9 (10)	-29 (13)

Tabla 4.3 : Resultados del *sesgo* $\times\, 10^3$ (*ecm* $\times\, 10^4$) de $\widehat{S}_T(\tau)$ en el caso de alta censura con subsupervivencias a priori correctas para $N = 100$.

Las tablas 4.2 y 4.3 contienen las estimaciones del *sesgo* y *ecm* de $\widehat{S}_T(\tau)$ (con $\tau = t_{0.9}$, $t_{0.5}$ y $t_{0.1}$) para los casos de baja y alta censura respectivamente. En ambos se considera una correcta elección a priori y un tamaño muestral $N = 100$. En general, se observa que los *sesgos* (absolutos) y *ecm* son menores en el caso de baja censura. Nótese, sin embargo, que los valores del *sesgo* parecen ser mayores cuando se supone una creencia fuerte en las subsupervivencias a priori que cuando se asume que la creencia es débil.

Por otra parte, la tabla 4.4 presenta los valores del $ECMI$ cuando las subsupervivencias a priori son uniformes. Obviamente, estos valores son mucho mayores que los que se obtuvieron cuando las subsupervivencias a priori eran correctas. Puede observarse que, si se asume una creencia débil en la correcta elección de las subsupervivencias a priori, los valores del $ECMI$ son mayores para muestras pequeñas. En general, se verifica lo contrario cuando el grado de creencia es fuerte.

	θ	0.5		1.0		1.5		2.0	
λ	$\alpha(\Omega)$	*Baja cens.*	*Alta cens.*	*Baja cens.*	*Alta cens.*	*Baja cens.*	*Alta cens.*	*Baja cens.*	*Alta cens.*
0.5	$n/5$	10-13	10-16	10-13	10-16	10-13	10-16	10-12	10-15
	$5n$	134-98	95-81	135-99	96-80	135-98	95-81	134-99	96-79
1.0	$n/5$	7-9	7-12	6-9	7-13	6-10	7-13	6-9	7-13
	$5n$	73-48	52-44	72-47	51-46	72-46	51-45	72-47	51-47
1.5	$n/5$	5-8	5-11	4-8	5-11	5-8	5-12	4-8	5-12
	$5n$	40-26	30-29	39-26	30-29	39-25	30-30	39-25	30-30
2.0	$n/5$	3-7	4-11	3-7	4-11	4-7	4-10	4-8	5-11
	$5n$	23-17	20-22	23-16	21-21	23-16	20-22	24-16	21-23

Tabla 4.4 : Resultados del $ECMI \times 10^3$ de $\widehat{S}_T$ con subsupervivencias a priori uniformes para $N = 100$ - 25.

En resumen, los resultados de la simulación indican que el estimador propuesto en este capítulo presenta un comportamiento suficientemente bueno, incluso cuando el nivel de censura es alto. Como resulta del todo lógico, un fuerte grado de creencia en las subsupervivencias a priori es mucho mejor cuando éstas son correctas.

Con objeto de analizar la robustez del estimador propuesto frente a violaciones de los supuestos del modelo, se considera que las variables T e $Y' = Y - \omega\theta$ tienen sendas distribuciones de Weibull,

$$S_T(t) = \exp\left\{-(t/\theta)^{\lambda}\right\}, \, t \geq 0, \quad \text{y} \quad S_{Y'}(y') = \exp\left\{-\gamma(y'/\theta)^{\lambda}\right\}, y' \geq 0,$$

donde $\lambda = 1$, y que la variable X sigue una distribución uniforme en el intervalo $(0, \omega\theta)$. La generación de datos para evaluar los $ECMI$ en este contexto (*Caso R*) se realizó de forma similar al caso en el que el modelo se ajustaba a los datos.

Resolviendo el sistema $p_0 = \Pr(\delta = 0)$ y $(1 - p_0)/2 = \Pr(\delta = 2)$ resulta que ω es la solución de la ecuación $\{1 - \exp(-\omega)\}/\omega = (1 + p_0)/2$, mientras que $\gamma = 1/\{1 + 1/\omega + (p_0 - 1/\omega)\exp(\omega)\} - 1$.

De este modo, se verifica que $\omega = 0.60586$ y $\gamma = 0.84572$ en el caso de alta censura, mientras que $\omega = 0.10348$ y $\gamma = 0.05871$ en el de baja censura.

En esta simulación:

$$S_0(t) = \begin{cases} \dfrac{1}{\omega}\left\{e^{-t/\theta}(1 + t/\theta) - e^{-\omega}(1 + \omega)\right\} + \dfrac{e^{-\omega}}{1+\gamma} & si \;\; t < \omega\theta, \\[2ex] \dfrac{e^{\gamma\omega - (1+\gamma)t/\theta}}{1+\gamma} & si \;\; t \geq \omega\theta, \end{cases}$$

$$S_1(t) = \begin{cases} \dfrac{\gamma e^{-\omega}}{1+\gamma} & si \;\; t < \omega\theta, \\[2ex] \dfrac{\gamma e^{\gamma\omega - (1+\gamma)t/\theta}}{1+\gamma} & si \;\; t \geq \omega\theta, \end{cases}$$

$$S_2(t) = \begin{cases} 1 - \dfrac{t}{\omega\theta} + \dfrac{1}{\omega}\left(e^{-\omega} - e^{-t/\theta}\right) & si \;\; t < \omega\theta, \\[2ex] 0 & si \;\; t \geq \omega\theta. \end{cases}$$

La tabla 4.5 muestra los valores del $ECMI$ para el *Caso R* cuando se asume una elección correcta de las subsupervivencias a priori. A partir de los resultados expuestos en esta tabla y los de la tabla 4.1 (para $\lambda = 1$), puede concluirse que las desviaciones de los datos de los supuestos del modelo no parecen ser relevantes, excepto cuando el nivel de censura es alto. La conclusión es similar cuando las subsupervivencias a priori son uniformes.

θ		0.5		1.0		1.5		2.0	
λ	$\alpha(\Omega)$	*Baja cens.*	*Alta cens.*	*Baja cens.*	*Alta cens.*	*Baja cens.*	*Alta cens.*	*Baja cens.*	*Alta cens.*
1.0	$n/5$	19-57	91-176	19-53	89-181	17-57	93-180	16-54	91-186
	$5n$	10-23	87-163	10-23	86-162	10-24	85-163	10-24	85-173

Tabla 4.5 : Resultados del $ECMI \times 10^4$ de $\widehat{S}_T$ con subsupervivencias a priori correctas y $N = 100$ - 25 para el *Caso R*.

En este punto, se pretende realizar algunas comparaciones con otros estimadores. En primer término, se considera el estimador no paramétrico autoconsistente, obtenido primeramente por Turnbull (1974), el cual, bajo condiciones bastante generales, es el estimador no paramétrico máximo verosímil de la función de supervivencia para datos doblemente censurados. Además, se verifica que este estimador es uniformemente fuertemente consistente, y asintóticamente eficiente y normalmente distribuido (véase Gu y Zhang (1993) y sus referencias).

Cuando se comparan los resultados del $ECMI$ de la tabla 4.1 con los de la tabla 4.6, que corresponden al estimador no paramétrico autoconsistente, puede apreciarse que el estimador $\widehat{S}_T$ es mejor que el autoconsistente si se supone una creencia débil en las subsupervivencias a priori, y mucho mejor cuando se asume una creencia fuerte. Las diferencias son aún mayores cuando el nivel de censura es alto.

En segundo término, se considera el estimador propuesto en el artículo de Morales et al. (1990). La tabla 4.7 muestra los valores del $ECMI$ de este estimador supuesta una correcta elección de las subsupervivencias a priori . A partir de los resultados de la simulación, se observa que $\widehat{S}_T$ y el estimador de Morales et al. presentan un comportamiento similar. En algunas ocasiones, el estimador $\widehat{S}_T$ parece ser ligeramente mejor; en otras, se observa exactamente lo contrario.

θ	0.5		1.0		1.5		2.0	
λ	*Baja cens.*	*Alta cens.*	*Baja cens.*	*Alta cens.*	*Baja cens.*	*Alta cens.*	*Baja cens.*	*Alta cens.*
0.5	19-75	46-125	18-75	47-128	19-67	47-138	17-74	49-135
1.0	18-71	50-144	19-66	42-126	17-71	47-137	18-65	46-132
1.5	18-81	44-137	20-81	45-130	19-75	42-132	21-83	48-134
2.0	17-73	45-134	18-77	46-143	21-74	40-120	18-73	48-128

Tabla 4.6 : Resultados del $ECMI \times 10^4$ del estimador autoconsistente para $N = 100$ - 25.

	θ	0.5		1.0		1.5		2.0	
λ	$\alpha\,(\Omega)$	*Baja cens.*	*Alta cens.*	*Baja cens.*	*Alta cens.*	*Baja cens.*	*Alta cens.*	*Baja cens.*	*Alta cens.*
0.5	$n/5$	17-52	20-77	16-52	22-75	17-53	22-81	16-50	21-81
	$5n$	5-13	11-41	5-13	11-41	5-12	12-47	5-13	12-44
1.0	$n/5$	16-51	21-91	17-51	21-76	15-52	20-74	16-52	21-69
	$5n$	5-13	11-49	5-13	12-44	5-13	10-40	5-13	12-38
1.5	$n/5$	16-55	19-76	17-52	19-74	16-52	17-74	15-55	19-78
	$5n$	5-13	10-41	5-13	10-38	5-13	9-43	5-21	10-49
2.0	$n/5$	15-51	19-74	16-53	20-84	18-53	17-73	16-52	18-72
	$5n$	5-14	10-40	3-21	10-45	5-13	9-40	5-13	10-42

Tabla 4.7 : Resultados del $ECMI \times 10^4$ del estimador de Morales et al. con subsupervivencias a priori correctas para $N = 100$ - 25.

En último lugar, se analiza el efecto de la proporción de seguimiento, f^*, en el comportamiento del estimador propuesto. En las tablas 4.8 y 4.9 se presentan los resultados de la eficiencia relativa de $\widehat{S}_T$, supuesto que las subsupervivencias a priori son correctas, cuando el tamaño muestral es 25 y 100 respectivamente. La eficiencia relativa se obtiene multiplicando por 100 la razón entre ($ECMI$ cuando $f^* = 1$) y ($ECMI$ cuando $f^* = i$) para $i = 0.1, 0.4$ y 0.7. En la mayoría de las ocasiones, basta un nivel $f^* = 0.4$ para obtener una eficiencia aceptable, sobre todo si el nivel de censura no es alto. Como resulta lógico, la eficiencia aumenta cuando f^* crece o cuando disminuye el nivel de censura. A pesar de que el $ECMI$ de $\widehat{S}_T$ disminuye si aumenta el grado de creencia en las subsupervivencias a priori, no ocurre lo mismo con la eficiencia relativa. Además, se aprecia que la eficiencia relativa para $N = 25$ y $N = 100$ es bastante similar, excepto cuando f^* es pequeño.

		θ	0.5		1.0		1.5		2.0	
			Baja	*Alta*	*Baja*	*Alta*	*Baja*	*Alta*	*Baja*	*Alta*
λ	$\alpha(\Omega)$	f^*	*cens.*	*cens.*	*cens.*	*cens.*	*cens.*	*cens.*	*cens.*	*cens.*
0.5	$n/5$	0.1	59.2	52.6	58.0	49.9	58.4	51.3	57.1	51.2
		0.4	86.3	78.5	83.9	77.2	85.0	77.7	84.6	78.2
		0.7	96.4	93.0	96.3	92.1	96.2	91.8	95.7	92.9
	$5n$	0.1	66.8	40.3	66.4	39.9	66.7	39.3	65.3	39.3
		0.4	80.3	67.2	79.3	67.9	79.5	67.1	79.9	67.4
		0.7	95.1	87.3	95.0	87.4	95.8	87.2	94.3	86.6
1.0	$n/5$	0.1	58.5	51.1	56.8	51.9	59.0	50.2	57.5	50.6
		0.4	86.0	76.8	84.8	77.8	85.6	78.7	84.5	77.0
		0.7	96.7	91.8	95.5	92.6	96.7	93.4	96.0	92.0
	$5n$	0.1	65.2	39.6	65.4	39.2	66.7	38.9	65.7	40.2
		0.4	79.7	67.9	79.7	66.5	80.0	67.9	79.4	67.6
		0.7	94.3	87.2	95.4	86.5	94.7	87.9	95.1	87.2
1.5	$n/5$	0.1	58.2	50.3	59.6	51.9	58.0	52.7	57.5	52.4
		0.4	84.8	78.1	86.1	79.1	85.0	78.9	83.6	78.5
		0.7	95.8	93.1	97.0	93.2	96.3	93.3	96.7	92.1
	$5n$	0.1	66.6	39.2	65.9	39.5	66.1	40.5	64.9	39.4
		0.4	80.3	68.2	80.7	67.1	79.6	67.7	78.9	67.4
		0.7	95.2	88.3	94.8	87.2	94.8	87.4	95.0	87.1
2.0	$n/5$	0.1	57.2	53.5	57.7	51.6	58.4	52.3	57.9	50.8
		0.4	83.9	79.9	84.5	77.7	84.9	79.6	83.8	78.4
		0.7	95.3	93.4	96.0	92.0	96.3	94.0	95.9	93.2
	$5n$	0.1	65.6	40.2	66.1	40.2	65.7	39.4	65.4	39.5
		0.4	80.2	68.3	80.1	67.4	80.1	67.7	78.6	67.4
		0.7	94.6	87.4	94.9	87.6	94.7	87.9	95.3	87.1

Tabla 4.8 : Resultados de la eficiencia relativa de $\widehat{S}_T$ con subsupervivencias a priori correctas para $N = 25$.

Para finalizar, la tabla 4.10 presenta los resultados de las estimaciones de los errores cuadráticos medios de los tres estimadores considerados (autoconsistente, de Morales et al. y el propuesto en este capítulo) de $S_T(\tau)$ cuando τ coincide con la mediana poblacional, supuesto que las subsupervivencias a priori son correctas y que $N = 100$ - 25. Desde luego, los *ecm* disminuyen cuando aumenta tanto N como f^* en todos los casos. También, como las subsupervivencias a priori son correctas, una fuerte creencia en dichas subsupervivencias proporciona unos *ecm* más reducidos. Se observa a simple vista que los *ecm* del estimador autoconsistente en $t_{0.5}$ son mayores que los de los otros dos estimadores, sobre todo si existe un fuerte grado de creencia en las subsupervivencias a priori. A partir de los resultados de la simulación, puede apreciarse que el comportamiento de $\widehat{S}_T$ y del estimador de Morales et al. en $t_{0.5}$ es bastante similar. En algunas ocasiones, los *ecm* de $\widehat{S}_T\,(t_{0.5})$ son ligeramente menores; en otras, ocurre lo contrario. En los otros dos casos, para $\tau = t_{0.1}$ y $t_{0.9}$, se observan resultados parecidos a los obtenidos para $\tau = t_{0.5}$.

		θ	0.5		1.0		1.5		2.0	
λ	$\alpha\,(\Omega)$	f^*	*Baja cens.*	*Alta cens.*	*Baja cens.*	*Alta cens.*	*Baja cens.*	*Alta cens.*	*Baja cens.*	*Alta cens.*
0.5	$n/5$	0.1	58.7	39.4	59.2	41.7	58.8	40.6	58.6	40.3
		0.4	85.1	78.7	86.4	80.0	85.8	78.3	84.8	78.2
		0.7	95.3	92.4	95.5	93.3	95.5	92.7	95.0	92.8
	$5n$	0.1	46.1	23.5	46.1	23.4	46.1	23.9	45.9	24.0
		0.4	71.8	63.8	72.3	64.3	72.1	63.3	72.0	64.2
		0.7	88.7	86.4	88.3	86.7	88.5	86.3	88.1	86.6
1.0	$n/5$	0.1	58.9	39.9	58.2	41.0	59.8	39.5	58.7	40.3
		0.4	84.6	78.5	84.7	79.9	85.9	79.5	85.5	80.0
		0.7	95.5	93.0	94.9	93.6	95.7	94.0	95.1	93.9
	$5n$	0.1	46.0	23.4	46.1	24.4	46.0	23.4	46.5	23.5
		0.4	72.1	63.9	72.3	64.7	71.9	63.9	72.5	63.9
		0.7	88.4	86.6	88.6	87.2	88.3	87.0	89.4	86.9
1.5	$n/5$	0.1	58.5	41.5	59.2	40.4	59.2	39.3	59.8	39.5
		0.4	84.8	80.0	86.3	80.4	86.5	78.4	86.1	78.3
		0.7	95.7	93.2	95.8	93.1	95.4	92.1	94.9	92.4
	$5n$	0.1	45.6	23.9	46.3	24.4	46.3	23.6	46.1	23.8
		0.4	71.3	64.3	72.0	65.4	72.4	64.2	72.3	63.7
		0.7	87.9	87.0	88.7	87.5	88.7	86.5	89.2	86.1
2.0	$n/5$	0.1	57.8	39.1	58.9	40.4	57.8	41.1	59.7	39.8
		0.4	84.7	78.2	85.7	80.1	84.3	79.4	85.8	78.5
		0.7	94.9	93.5	94.6	93.1	94.6	93.1	95.6	92.3
	$5n$	0.1	45.6	23.5	45.5	23.6	45.6	24.0	46.7	23.6
		0.4	71.9	63.7	71.7	64.4	71.8	64.2	71.4	63.5
		0.7	88.4	86.8	88.9	86.9	88.2	86.9	88.6	86.7

Tabla 4.9 : Resultados de la eficiencia relativa de $\widehat{S}_T$ con subsupervivencias a priori correctas para $N = 100$.

		θ	1.0				2.0			
			Baja censura		*Alta censura*		*Baja censura*		*Alta censura*	
λ	f^*	$\alpha(\Omega)$	$n/5$	$5n$	$n/5$	$5n$	$n/5$	$5n$	$n/5$	$5n$
1.0	0.1	$ecm1$	26-113	26-113	61-199	61-199	27-114	27-114	60-198	60-198
		$ecm2$	21-72	3-26	43-161	14-75	22-76	3-26	44-161	14-69
		$ecm3$	22-68	17-55	48-174	16-162	24-71	18-54	50-176	15-163
	0.4	$ecm1$	25-112	25-112	50-161	50-161	27-113	27-113	49-154	49-154
		$ecm2$	18-71	3-26	27-116	13-68	19-75	3-25	28-115	13-64
		$ecm3$	18-67	14-54	25-107	10-90	20-70	15-53	26-108	11-94
	0.7	$ecm1$	25-111	25-111	48-154	48-154	26-112	26-112	48-146	48-146
		$ecm2$	17-70	3-25	24-106	12-62	18-73	3-24	25-105	13-60
		$ecm3$	17-66	12-53	21-93	7-72	19-68	12-51	22-93	7-73
	1.0	$ecm1$	25-111	25-111	48-151	48-151	26-112	26-112	47-144	47-144
		$ecm2$	17-70	3-25	23-103	12-62	18-73	3-24	24-101	12-58
		$ecm3$	17-66	11-53	21-89	6-63	18-68	12-51	21-88	6-63
2.0	0.1	$ecm1$	27-109	27-109	63-195	63-195	27-112	27-112	62-203	62-203
		$ecm2$	23-72	3-26	46-153	14-70	22-75	3-27	47-153	14-71
		$ecm3$	23-68	18-54	52-164	15-160	23-72	18-55	54-165	14-161
	0.4	$ecm1$	26-108	26-108	52-151	52-151	26-111	26-111	50-159	50-159
		$ecm2$	18-70	3-25	29-109	13-64	18-73	3-26	28-109	12-65
		$ecm3$	19-66	14-53	26-99	11-91	19-71	14-54	25-100	11-92
	0.7	$ecm1$	26-107	26-107	51-143	51-143	26-110	26-110	48-153	48-153
		$ecm2$	18-69	3-25	26-100	13-61	18-72	3-25	24-100	12-61
		$ecm3$	18-64	12-52	23-87	7-73	18-69	12-52	22-88	7-73
	1.0	$ecm1$	26-107	26-107	50-141	50-141	26-110	26-110	48-152	48-152
		$ecm2$	17-69	3-25	25-97	12-59	18-71	3-25	23-97	12-59
		$ecm3$	17-64	11-52	22-84	6-64	18-68	11-52	21-84	6-64

Tabla 4.10 : Resultados del $ecm \times 10^4$ de los estimadores de $S_T(t_{0.5})$ con subsupervivencias a priori correctas para $N = 100$ - 25.

Bibliografía

Aalen, O.(1978). "Nonparametric inference for a family of counting processes", *Ann. Statist.*, **6**, 701-726.

Allen, W. R. (1963). "A note on conditional probability of failure when hazards are proportional", *Operations Research*, **11**, 658-659.

Andersen, P. K., Borgan, O., Gill, R. D. y Keiding, N. (1993). *Statistical Models based on Counting Processes*, Springer-Verlag, New York.

Andersen, P. K. y Gill, R. D. (1982). "Cox's regression model for counting processes: a large sample study ", *Ann. Statist.*, **10**, 1100-1120.

Antoniak, C. E. (1974). "Mixtures of Dirichlet processes with application to Bayesian nonparametric problems", *Ann. Statist.*, **2**, 1152-1174.

Ascher, H. y Feingold, H. (1984). *Repairable Systems Reliability: Modeling, Inference, Misconceptions and Their Causes*, Lectures Notes in Statistics, Vol. 7, Marcel Dekker, New York.

Bain, L. J. (1978). *Statistical Analysis of Reliability and Life-Testing Models*, Marcel Dekker, New York.

Bain, L. J., y Engelhardt, M. (1980). "Probability of correct selection of Weibull versus gamma based on likelihood ratio", *Comm. Statist.-Theory Meth.*, **9**, 375-381.

Barlow, R. E. y Proschan, F. (1965). *Mathematical Theory of Reliability,* John Wiley, New York.

Barlow, R. E. y Proschan, F. (1975). *Statistical Theory of Reliability and Life-Testing: Probability Models,* Holt, Rinehart and Wilson, New York.

Bartholomew, D. J. (1957). "A problem in life testing", *J. Amer. Statist. Assoc.*, **52**, 350- 355.

Bartholomew, D. J. (1963). "The sampling distribution of an estimate arising in life testing", *Technometrics*, **5**, 361-374.

Bickel, P. J., Klaassen, C. A. J., Ritov, I. y Wellner, J. A. (1993). *Efficient and Adaptive Estimation for Semiparametric Models*, Johns Hopkins Springer Univ. Press.

Bravo, J. I. y Esteban, M. D. (1993a). "Nonparametric estimation of survival functions for doubly censored observations when the lifetime hazard rate is proportionally related to the censoring hazard rates", *Comm. Statist.-Theory Meth.*, **22**, 3237-3253.

Bravo, J. I. y Esteban, M. D. (1993b). "Partially parametric estimation of survival functions for doubly censored observations", ASMDA 93: *Proc. Int. Symp. Appl. Stochastic Models and Data Analysis* (Eds. J. Janssen and C. H. Skiadas), World Scientific Publ. Co., 106-113.

Bravo, J. I., De Fuentes, I. y Fernández, A. J. (1995a). "A semi-parametric estimation of a survival function from incomplete and doubly censored data", *Comm. Statist.-Theory Meth.*, **24**, 2735-2752.

Bravo, J. I., De Fuentes, I. y Fernández, A. J. (1995b). "Nonparametric estimation of a survival function from incomplete and doubly censored data under a proportional hazard rates model", *ASA Proceedings of the Biometrics Section*, American Statistical Association, Alexandria, 309-314.

Bravo, J. I., De Fuentes, I. y Fernández, A. J. (1995c). "A partially parametric estimation of a survival function from incomplete and doubly censored data", *ASA Proceedings of the Biometrics Section*, American Statistical Association, Alexandria, 315-320.

Breslow, N. y Crowley, J. (1974). "A large sample study of the life table and product limit estimates under random censorship", *Ann. Statist.*, **2**, 437-453.

Buckley, J. y James, I. (1979). "Linear regression with censored data", *Biometrika*, **66**, 429–436.

Campbell, G. (1981). "Nonparametric bivariate estimation with randomly censored data", *Biometrika*, **68**, 417-422.

Chang, M. N. (1990). "Weak convergence of self-consistent estimator of the survival function with doubly censored data", *Ann. Statist.*, **18**, 391-404.

Chang, M. N. y Yang, G. L. (1987). "Strong consistency of a nonparametric estimator of the survival function with doubly censored data", *Ann. Statist.*, **15**, 1536-1547.

Cheng, P. E. y Lin, G. D. (1987). "Maximum likelihood estimation of a survival function under the Koziol-Green proportional hazards model", *Statist. & Probab. Letters*, **5**, 75-80.

Chhikara, R. S. y Folks, J. L. (1977). "The inverse Gaussian distribution as a lifetime model", *Technometrics*, **19**, 461-468.

Cohen, A. C. (1961). "Tables for maximum likelihood estimation: singly truncated and singly censored samples", *Technometrics*, **3**, 535-541.

Cohen, A. C. (1965). "Maximum likelihood estimation in the Weibull distribution based on complete and censored samples", *Technometrics*, **17**, 579-588.

Cox, D. R. (1972). "Regression models and life-tables (with discussion)", *J. Roy. Statist. Soc. B*, **34**, 187-220.

Cox, D. R. y Hinkley, D. V. (1974). *Theoretical Statistics,* Chapman and Hall, London.

Cox, D. R. y Oakes, D. (1984). *Analysis of Survival Data,* Chapman and Hall, London.

Csörgo, S. y Horváth, L. (1981). "On the Koziol-Green model for random censorship", *Biometrika*, **68**, 391-401.

Davis, D. J. (1952). "An analysis of some failure data", *J. Amer. Statist. Assoc.*, **47**, 113-150.

Davis, H. T. y Feldstein, M. L. (1979). "The generalized Pareto law as a model for progressively censored survival data", *Biometrika*, **66**, 299-306.

Dempster, A. P., Laird, N. M. y Rubin, D. B. (1977). "Maximun likelihood from incomplete data via the EM algorithm (with discussion)", *J. Roy. Statist. Soc. B*, **39**, 1-38.

Dewanji, A. (1992). "A note on a test for competing risks with missing failure time, *Biometrika",* **79**, 855-857.

Diamond, I. D. y Mc Donald, J. W. (1991). "Analysis of current status data", In *Demographic applications of event history analysis*, (J. Trusell, R. Hankinson and J. Tilton, eds.), Oxford Univ. Press, 231-252.

Diamond, I. D., Mc Donald, J. W. y Shah, I. H. (1986). "Proportional hazars models for current status data: application to the study of differentials in age at weaning in Pakistan", *Demography*, **23**, 607-620.

Dinse, G. E. (1982). "Nonparametric estimation for partially-complete time and type of failure data", *Biometrics*, **38**, 417-431.

Dinse, G. E. (1986). "Nonparametric prevalence and mortality estimators for animal experiments with incomplete causa-of-death data", *J. Amer. Statist. Assoc.*, **81**, 328-336.

Dinse, G. E. (1988). "Estimating tumor incidence rates in animal carcinogenicity experiments", *Biometrics*, **44**, 405-415.

Dinse, G. E. y Lagakos, S. W. (1983). "Regression analysis of tumor prevalence data", *J. Roy. Statist. Soc. C*, **32**, 236-248.

Dixon, W. J. (1960). "Simplified estimation from censored normal samples", *Ann. Math. Statist.*, **31**, 385-391.

Doksum, K. A. (1974). "Tailfree and neutral random probabilities and their posterior distributions", *Ann. Probab.*, **2**, 183-201.

Dykstra, R. L. y Laud, P. (1981). "A Bayesian nonparametric approach to reliability", *Ann. Statist.*, **9**, 356-367.

Ebrahimi, N. (1985). "Nonparametric estimation of survival functions for incomplete observations when the life time distribution is proportionally related to the censoring time distribution", *Comm. Statist.-Theory Meth.*, **14** , 2887-2898.

Efron, B. (1967). "The two-sample problem with censored data", *Proc. Fith Berkeley Symp. Math. Statist. Probab.*, Univ. California Press, **4**, 831-853.

Epstein, B. (1958). "The exponential distribution and its role in life-testing", *Ind. Qual. Control*, **15**, 2-7.

Epstein, B. y Sobel, M. (1953). "Life testing", *J. Amer. Statist. Assoc.*, **48**, 486-502.

Epstein, B. y Sobel, M. (1954). "Some theorems relevant to life testing from an exponential distribution", *Ann. Math. Statist.*, **25**, 373-381.

Epstein, B. y Tsao, C. K. (1953). "Some test based on ordered observations from two exponential populations", *Ann. Math. Statist.*, **24**, 456-466.

Feinleib, M. (1960). "A method of analyzing lognormally distributed survival data with incomplete follow-up", *J. Amer. Statist. Assoc.*, **55**, 534-545.

Ferguson, T. S. (1973). "A Bayesian analysis of some nonparametric problems", *Ann. Statist.*, **1**, 209-230.

Ferguson, T. S. y Phadia, E. G. (1979). "Bayesian nonparametric estimation based on censored data", *Ann. Statist.*, **7**, 163-186.

Fernández, A. J., Bravo, J. I. y De Fuentes, I. (1995a). "Estimación de una curva de supervivencia bajo un modelo de doble censura y tasas de fallo proporcionales", *Actas de la V Conferencia Española de Biometría*, Valencia, 87-90.

Fernández, A. J., Bravo, J. I. y De Fuentes, I. (1995b). "Estimation of a lifetime curve from incomplete field data and a Dirichlet process prior knowledge on the observable random vector", *ASA Proceedings of the Section on Bayesian Statistical Science*, American Statistical Association, Alexandria, 218-223.

Fernández, A. J., Bravo, J. I. y De Fuentes, I. (1996). "Modelos de supervivencia en doble censura con parámetros variables en el tiempo y covariables", *Memoria del X Foro Nacional de Estadística y II Congreso Iberoamericano de Estadística* (Eds. M. L. de la Fuente, E. Gutiérrez y J. Olguín), 109-113.

Fernández, A. J., Bravo, J. I. y De Fuentes, I. (1997). "Lifetime estimation from doubly censored data and Dirichlet process prior knowledge on the observable random vector", *Comm. Statist.-Theory Meth.*, **26**, 1541-1558.

Finkelstein, D. M. (1986). "A proportional hazards model for interval-censored failure time data", *Biometrics*, **42**, 845-854.

Finkelstein, D. M. y Wolfe, R. A. (1985). "A semiparametric model for regression analysis of interval-censored failure time data", *Biometrics*, **41**, 933-945.

Fleming, T. R. y Harrington, D. P. (1991). *Counting Processes and Survival Analysis ,* John Wiley, New York.

Gehan, E. A. (1965a). "A generalized two-sample Wilcoxon test for comparing arbitrarily singly censored samples", *Biometrika*, **52**, 203-223.

Gehan, E. A. (1965b). "A generalized two-sample Wilcoxon test for doubly censored data", *Biometrika*, **52**, 650–653.

Genest, C., Ghoudi, K. y Rivest, L. P. (1995). " A semiparametric estimation procedure of dependence parameters in multivariate families of distributions", *Biometrika*, **82**, 543-552.

Gerstbakh, I. B. (1989). *Statistical Reliability Theory,* Marcel Dekker, New York.

Gill, R. D. (1989). "Non- and semi-parametric maximum likelihood estimators and the von Mises method", *Scand. J. Statist.*, **16**, 97–128.

Goetghebeur, E. y Ryan, L. (1990). "A modified logrank test for competing risks with missing failure type", *Biometrika*, **77**, 207-211.

Goetghebeur, E. y Ryan, L. (1995). "Analysis of competing risks survival data when some failure types are missing", *Biometrika*, **82**, 821-833.

Gross, A. J. y Clark, V. A. (1975). *Survival Distributions: Reliability Applications in the Biomedical Sciences*, John Wiley, New York.

Gu, M. G. y Zhang C. H. (1993). "Asymptotic properties of self-consistent estimators based on doubly censored data", *Ann. Statist.*, **21**, 611-624.

Gupta, S. S. y Groll, P. A. (1961). "Gamma distribution in acceptance sampling based on life tests", , *J. Amer. Statist. Assoc.*, **65**, 942-970.

Harter, H. L. (1967). "Maximum-likelihood estimation of the parameters of a four-parameter generalized gamma population for complete and censored samples", *Technometrics*, **9**, 159-165.

Harter, H. L. y Moore, A. H. (1965). "Maximum likelihood estimation of the parameters of the gamma and Weibull populations from complete and from censored samples", *Technometrics*, **7**, 639-643.

Herbst, T. (1992a). "Test of fit with the Koziol-Green model for random censorship", *Statistics & Decisions*, **10**, 163-171.

Herbst, T. (1992b). "Estimation of moments under Koziol-Green model for random censorship", *Comm. Statist.-Theory Meth.*, **21**, 613-624.

Herbst, T. (1993). "On estimation of residual moments under Koziol-Green model of random censorship", *Comm. Statist.-Theory Meth.*, **22**, 2403-2419.

Hill, B. M. (1963). "The three-parameter lognormal distribution and Bayesian analysis of a point-source epidemic", *J. Amer. Statist. Assoc.*, **58**, 72-84.

Hoel, D. G. y Walburg, R. A. (1972). "Statistical analysis of survival experiments", *J. Nat. Cancer Inst.*, **49**, 361-372.

Huang, J. (1996). "Efficient estimation for the proportional hazards model with interval censoring", *Ann. Statist.*, **24**, 540-568.

Jewell, N. P., Malahi, H. M. y Vittinghoff, E. (1994). "Nonparametric estimation for a form of doubly censored data, with applications to two problems in AIDS", *J. Amer. Statist. Assoc.*, **89**, 425, 7-18.

Johansen, S. (1978). "The product limit estimate as a maximum likelihood estimate", *Scand. J. Statist.*, **5**, 195-199.

Kalbfleisch, J. D. (1978). "Non-parametric Bayesian analysis of survival time data", *J. R. Statist. Soc. B*, **40**, 214-221.

Kalbfleisch , J. D. y Lawless, J. F. (1988). "Estimation of reliability in field-perfomance studies", *Technometrics*, **30**, 365-377.

Kalbfleisch , J. D. y Prentice, R. L. (1980). *The Statistical Analysis of Failure Time Data*, John Wilew, New York.

Kao, J. H. K. (1959). "A graphical estimation of mixed Weibull parameters in life testing of electron tubes", *Technometrics*, **1**, 389-407.

Kaplan, E. L. y Meier, P. (1958). "Nonparametric estimation from incomplete observations", *J. Amer. Statist. Assoc.*, **53**, 457-481.

Klassen, B. y van Peppen, J. C. L. (1989). *System Reliability: Concepts and Applications,* Eduard Arnold, New York.

Klein, J. P., Lee, S. C. y Moeschberger, M. L. (1990). "A partially parametric estimator of survival in the presence of randomly censored data", *Biometrics*, **46**, 795-811.

Kodell, R. L. y Chen, J. J. (1987). "Handling cause of death in equivocal cases using the EM algorithm (with rejoinder), *Comm. Statist.-Theory Meth.*, **16**, 2567-2585.

Korwar, R. M. (1987). "Nonparametric estimation of a bivariate survivorship function with doubly censored data", *Statist. & Probab. Letters*, **5**, 119-124.

Koziol, J. A. y Green, S. B. (1976). "A Cramér-von Mises statistic for randomly censored data", *Biometrika*, **63**, 465-474.

Lagakos, S. W. (1982). "An evaluation of some two-sample tests to analyze aninal carcinogenicity experiments", *Utilitas Math. B*, **21**, 239-260.

Lagakos, S. W. y Louis, T. A. (1988). "Use of tumour lethality to interpret tumorigenicity experiments lacking cause of death data", *Appl. Statist.*, **37**, 169-179.

Lawless, J. F. (1982). *Statistical Models and Methods for Lifetime Data*, John Wiley, New York.

Le Cam, L. (1953). "On some asymptotic properties of maximum likelihood estimates and related Bayes' estimates", *Univ. of Calif. Publ. in Statist.*, **1**, 277-330.

Le Cam, L. (1970). "On the assumptions used to prove asymptotic normality of maximum likelihood estimates", *Ann. Math. Statist.*, **41**, 802-828.

Lehmann, E. L. (1983). *Theory of Point Estimation*, John Wiley, New York.

Leiderman, P. H., Babu, D., Kagia, J., Kraemer, H. C. y Leiderman, G. F. (1973). "Africant infant precocity and some social influences during the first year", *Nature*, **242**, 247-249.

Lieblein, J. y Zelen, M. (1956). "Statistical investigation of the fatigue life of deep groove ball bearings", *J. Res. Nat. Bur. Stand.*, **57**, 273-316.

Little, R. J. y Rubin, D. B. (1987). *Statistical Analysis with Missing Data*, John Wiley, New York.

Mäkeläinen, T., Schmidt, K. y Styan, G. P. H. (1981). "On the existence and uniqueness of the maximum likelihood estimate of a vector-valued parameter in fixed-sized samples", *Ann. Statist.*, **9**, 758.767.

Maguluri, G. (1993). "Semiparametric estimation of association in a bivariate survival function", *Ann. Statist.*, **21**, 1648-1662.

Mann, N. R., Schafer, R. E., y Singpurwalla, N. D. (1974). *Methods for Statistical Analysis of Reliability and Life Data,* John Wiley, New York.

Mantel, N. (1967). "Ranking procedures for arbitrarily restricted observations", *Biometrics*, **23**, 65–78.

Martz, H. F. y Waller, R. A. (1982). *Bayesian Reliability Analysis,* John Wiley, New York.

Miller, R. G. (1976). "Least squares regression with censored data", *Biometrika*, **63**, 447-464.

Miller, R. G. (1981). *Survival Analysis,* John Wiley, New York.

Miller, R. G. (1983). "What price Kaplan-Meier?", *Biometrics*, **39**, 1077-1081.

Miyakawa, M (1984). "Analysis of incomplete data in competing risks models", *IEEE Trans. Reliab.*, **33**, 293-296.

Morales, D., Pardo, L. y Quesada, V. (1990). "Estimation of a survival function with doubly censored data and Dirichlet proccess prior knowledge on the observable variable", *Comm. Statist.-Simul.*, **19**, 349-361.

Morales, D., Pardo, L. y Quesada, V. (1991). "Bayesian survival estimation for incomplete data when the life distribution is proportionally related to the censoring time distributions, *Comm. Statist.-Theory Meth.*, **20**, 831-850.

Mykland, P. A. y Ren, J.-J. (1996). "Algorithms for computing self-consistent and maximum likelihood estimators with doubly censored data", *Ann. Statist.*, **24**, 1740-1764.

Nelson, W. (1982). *Applied Life Data Analysis,* John Wiley, New York.

Nelson, W. (1990). *Accelerated Testing: Statistical Models, Test Plans and Data Analyses,* John Wiley, New York.

Oakes, D. (1986). "Semiparametric inference in a model for association in bivariate survival data", *Biometrika*, **73**, 353-361.

Oakes, D. (1994). "Multivariate survival distributions", *J. Nonparam. Statist.*, **3**, 343-354.

Padgett, W. J. (1979). "Confidence bounds on reliability for the inverse Gaussian model", *IEEE Trans. Reliab.*, **28**, 165-168.

Peto, R. (1973). "Experimental survival curves for interval-censored data", *Appl. Statist.*, **22**, 86–91.

Peterson, A. V.(1977). "Expressing the Kaplan-Meier estimate as a function of empirical subsurvival funcions", *J. Amer. Statist. Assoc.*, **82**, 854-858.

Prokhorov, Yu. V. (1956). "Convergence of random processes and limit theorems in probability theory", *Theor. Prob. and Appl.* (translated by SIAM), **1**, 157-224.

Proschan, F. (1963). "Theoretical explanation of observed decreasing failure rate", *Technometrics*, **5**, 375-383.

Rabinowitz, D., Tsiatis, A. y Aragon, J. (1995). "Regression with interval-censored data", *Biometrika*, **82**, 501-513.

Racine-Poon, A. H. y Hoel, D. G. (1984). "Nonparametric estimation of the survival function when cause of death is uncertain", *Biometrics*, **40**, 1151-1158.

Rao, C. R. (1973). *Linear Statistical Inference and Its Applications*, John Wiley, New York.

Ren, J. (1995). "Generalized Cramér-von Mises tests of goodness of fit for doubly censored data", *Ann. Inst. Statist. Math.*, **47**, 525-549.

Samuelson, S. O. (1989). "Asymptotic theory for non-parametric estimators from doubly censored data", *Scand. J. Statist.*, **16**, 1–21.

Shiboski, S. C. y Jewell, N. P. (1992). "Statistical analysis of the time dependence of HIV infectivity based on partner study data", *J. Amer. Statist. Assoc.*, **87**, 360-372.

Skorohod, A. V. (1956). "Limit theorems for stochastic processes", *Theor. Prob. and Appl.* (translated by SIAM), **1**, 261-290.

Stacy, E. W. y Mihram, G. A. (1965). "Parameter estimation for a generalized gamma distribution", *Technometrics*, **7**, 349-358.

Susarla, V. y Van Ryzin, J. (1976). "Nonparametric Bayesian estimation of survival curves from incomplete observations", *J. Amer. Statist. Assoc.*, **71**, 897-902.

Susarla, V. y Van Ryzin, J. (1978). "Large sample theory for a Bayesian nonparametric survival curve estimator based on censored data", *Ann. Statist.*, **6**, 755-768.

Suzuki, K. (1985a). "Nonparametric estimation of lifetime distributions from a record of failures and follow-ups", *J. Amer. Statist. Assoc.*, **80**, 68-72.

Suzuki, K. (1985b). "Estimation of lifetime parameters from incomplete field data", *Technometrics*, **27**, 263-271.

Tiwari, R. C. y Zalkikar, J. N. (1993). "Nonparametric Bayesian estimation of survival function under random left truncation", *J. Statist. Plann. Infer.*, **35**, 31-45.

Tang, M.-X., Tsai, W.-Y., Marder, K. y Mayeaux, R. (1995). "Linear rank test for doubly censored data", *Statistics in Medicine*, **14**, 2555-2563.

Tsai, W. Y. y Crowley, J. (1985). "A large sample study of generalized maximum likelihood estimators from incomplete data via self-consistency", *Ann. Statist.*, **13**, 1317-1334.

Tsokos, C. P. y Shimi, I. N. (1977). *The Theory and Applications of Reliability with Emphasis on Bayesian and Non-parametric Methods,* Academic Press, New York.

Turnbull, B. W. (1974). "Nonparametric estimation of a survivorship function with doubly censored data", *J. Amer. Statist. Assoc.*, **69**, 169-173.

Turnbull, B. W. (1976). "The empirical distribution function with arbitrarily grouped, censored and truncated data", *J. Roy. Statist. Soc. B*, **38**, 290-295.

Turnbull, B. W. y Weiss, L. (1978). "A likelihood ratio statistic for testing goodness of fit with randomly censored data", *Biometrics*, **34**, 367-375.

Wang, Z. y Gardiner, J. C. (1996). "A class of estimators of the survival function from interval-censored data", *Ann. Statist.*, **24**, 647-658.

Weibull, W. (1951). "A statistical distribution function of wide applicability", *J. Appl. Mech.*, **18**, 293-297.

Wellner, J. A. (1982). "Asymptotic optimality of the product limit estimator", *Ann. Statist.*, **10**, 595-602.

Wilk, M. B., Gnanadesikan, R. y Huyett, M. J. (1962a). "Probability plots for the gamma distribution", *Technometrics*, **4**, 1-20.

Wilk, M. B., Gnanadesikan, R. y Huyett, M. J. (1962b). "Estimation of parameters of the gamma distribution using order statistics", *Biometrika*, **49**, 525-545.

Wu, C. F. (1983). "On the convergence properties of the EM algorithm", *Ann. Statist.*, **11**, 95-103.

Zelen, M. y Dannemiller, M. (1961). "The robustness of life testing procedures derived from the exponential distribution", *Technometrics*, **3**, 29-49.

Zhang C.-H. y Li, X. (1996). "Linear regression with doubly censored data", *Ann. Statist.*, **24**, 2720-2743.