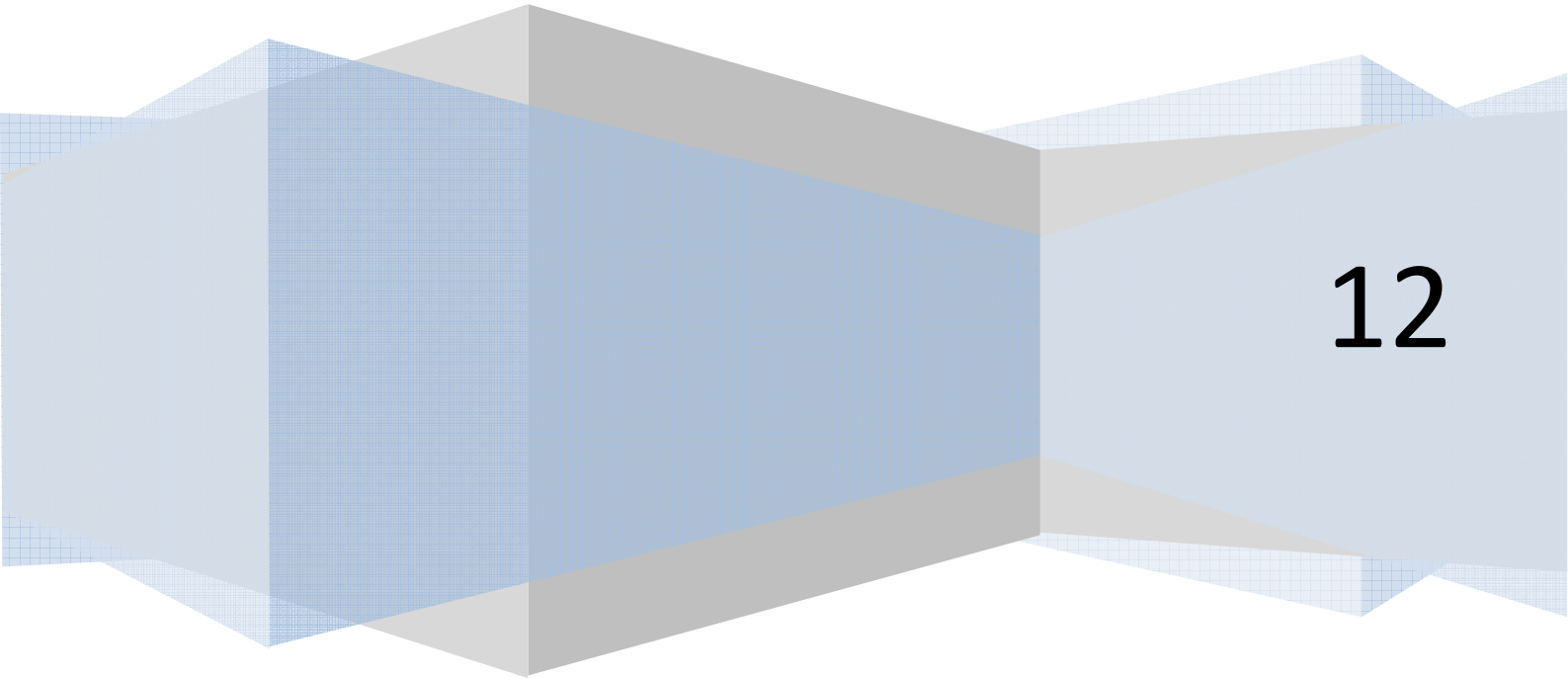


**Máster Oficial en Estadística Aplicada**  
Universidad de Granada

# **Modelos de riesgos competitivos**

**Proyecto Fin de Máster**

**Autora: Susana Alba Gutiérrez**



**12**



## Contenido

1. Introducción .....	5
2. El modelo de riesgos competitivos.....	7
3 Interpretación del modelo.....	14
4. Modelos de regresión de riesgos competitivos.....	18
5. Análisis estadístico mediante el uso del paquete “cmprsk” de R.....	24
6. Referencias .....	37



## 1. Introducción

El análisis de supervivencia es el conjunto de técnicas que permiten estudiar la variable “tiempo hasta que ocurre un evento”, por ejemplo el tiempo hasta la muerte, curación, probabilidad de fallo en cada instante, riesgo de fallo, etc. Dentro del análisis de supervivencia los modelos de riesgos competitivos son apropiados para analizar el comportamiento de un sujeto que puede fallar por diferentes causas y para cada unidad en estudio se observa tanto el tiempo hasta el fallo, como el tipo de fallo.

Los modelos de riesgos competitivos fueron introducidos por David Bernoulli, quien realizó intentos en 1760 de averiguar el riesgo de morir a causa de la viruela o de otras causas. Este es de hecho un ejemplo clásico de los riesgos competitivos, donde los individuos están sujetos a múltiples causas de muerte. Mucho más adelante, se han publicado trabajos en el campo de la estadística o la matemática que han incorporado nuevos avances, entre ellos la monografía de David y Moeschberger (1978).

Con la disponibilidad de datos más amplios, claros y precisos, los riesgos competitivos volvieron a surgir como un tipo crucial de análisis dentro de los análisis de tiempo de supervivencia. Varios autores han contribuido a mejorar el conocimiento de las situaciones de riesgos competitivos. Otros, por su parte, han mejorado y desarrollado técnicas y en algunos casos han proporcionado códigos de programas informáticos preparados para el uso en el análisis estadístico aplicado.

Algunos ejemplos de datos de riesgos competitivos se encuentran hoy día en muchos campos. Por ejemplo, en el campo de la fiabilidad, un caso de riesgo competitivo puede ser el estudio de la ruptura de un componente mecánico, donde hay varias causas posibles de fallo (vibración, corrosión, etc.). Otro ejemplo puede ser un estudio demográfico donde se registran las principales causas de muerte de la población (enfermedades del corazón, cáncer,...). Aquí puede ser interesante centrarse en analizar cada una de ellas separadamente. Otro caso de riesgo competitivo puede darse en un ensayo clínico realizado para encontrar los beneficios de un nuevo medicamento para prevenir el infarto de miocardio. El suceso de interés es el infarto de miocardio, pero los pacientes en estudio (personas con enfermedades en las arterias coronarias), pueden morir por otras causas tras un seguimiento de dos años.

Existen otras aplicaciones similares tanto en demografía como en ciencias actuariales, bajo la denominación de análisis de decremento múltiple. También, actualmente, en el campo de la fiabilidad, existen modernos bancos de datos que suelen distinguir entre un gran número de causas de fallo, lo que indica el uso y aplicación de métodos de la teoría de riesgos competitivos.

Por lo tanto, la característica distintiva de un ajuste de un modelo de riesgos competitivos es que cada individuo está caracterizado por un par  $(T, C)$  donde  $T > 0$  es el tiempo hasta la ocurrencia del fenómeno de interés que, en fiabilidad suele ser el fallo de un dispositivo, y  $C \in \{1, \dots, k\}$  representa la causa de fallo y por eso, se requiere de un modelo conjunto para  $T$  y  $C$ .

Por tanto, disponemos de  $k$  diferentes causas de fallo y cada causa de fallo se puede clasificar como un elemento perteneciente al conjunto  $\{1, \dots, k\}$ .

Hay que tener en cuenta que el término “fallo” se utiliza como un término genérico y en la práctica se corresponde con cualquier evento de interés dependiendo del estudio que estemos realizando. Igualmente sucede con el término “tiempo”. No significa necesariamente el tiempo cronológico, sino que puede ser cualquier medida adecuada no decreciente en relación con el tiempo cronológico: tiempo de funcionamiento, número de ciclos, número de kilómetros recorridos, la longitud de una grieta, etc.

La distribución conjunta de  $(T, C)$  puede especificarse completamente a través del llamado *riesgo de causa-específica*, es decir, el riesgo instantáneo de fallo (o muerte) en un momento dado por una causa dada, entre todos los individuos en riesgo en ese momento. También, la distribución conjunta puede especificarse a través de la *función de incidencia acumulada*, que representa la probabilidad de fallo por una causa dada, antes de un tiempo de vencimiento específico. Estas dos funciones representan cantidades distintas, estimar una u otra depende de las cuestiones de interés en la investigación.

Cabe mencionar que las probabilidades de fallo, dada la eliminación de alguna o de todas las otras causas, ha sido un tema central y clásico en el análisis de riesgos competitivos. En artículos recientes esas probabilidades aparecen con otros nombres, pero anteriormente se referían a esas cantidades como:

- *Probabilidades netas*. A excepción de la  $j$ -ésima causa, todas las otras causas son eliminadas.
- *Probabilidades crudas (parciales)*. Todas las causas están presentes (algunas pero no todas las causas son eliminadas).

En el artículo de Prentice y Kalbfleisch (1978) se encuentra una buena discusión de tales problemas.

A continuación, en la sección 2 vamos a especificar los modelos de riesgos competitivos. En las subsecciones 2.1 y 2.2 introduciremos la notación y los conceptos clave. En las subsecciones 2.3 y 2.4, derivaremos la función de verosimilitud dada una muestra aleatoria y obtendremos estimadores no paramétricos. En la sección 3 se definen las funciones de supervivencia en el marco de los riesgos competitivos, y discutiremos cómo interpretar correctamente estas cantidades. En la sección 4, se revisan los modelos de regresión de riesgos competitivos cuando hay presente un conjunto de covariables. En concreto, en la subsección 4.1, se consideran los modelos basados en los riesgos por causas específicas y en la subsección 4.2, se explicará con algunos detalles el modelo más popular basado en la función de incidencia acumulada, modelo de Fine y Gray (1999).

## 2. El modelo de riesgos competitivos

Estamos interesados en el análisis del tiempo hasta que ocurre un determinado suceso de interés en un grupo de sujetos del mismo tipo. El suceso de interés se refiere en general al fallo de un dispositivo en el ámbito del análisis de fiabilidad y a la muerte de un sujeto, en análisis de supervivencia.

En este trabajo nos referimos indistintamente al suceso de interés como fallo o como muerte.

### 2.1 Especificación del modelo

Se define, para cada individuo, el par  $(T, C)$ , donde  $T$  es el tiempo de fallo, y  $C$  es la causa del fallo.  $T$  es una variable aleatoria continua y positiva, mientras que  $C$  toma valores en el conjunto finito  $\{1, \dots, k\}$ . Se asume que el individuo falla debido a una y sólo una causa. La distribución conjunta de  $(T, C)$  se especifica completamente a través de cualquiera de las funciones de riesgo de causa-específica,  $\lambda_j(t)$ , o a través de las funciones de incidencias acumulada,  $F_j(t)$ .

Se define la función de riesgo de causa-específica en el tiempo  $t$  para la  $j$ -ésima causa en presencia de todas las otras causas actuando simultáneamente en la población como

$$\lambda_j(t) = \lim_{\Delta t \rightarrow 0} \frac{\Pr(T < t + \Delta t, C = j | T \geq t)}{\Delta t} \quad j = 1, \dots, k$$

es decir,

$$\lambda_j(t) = \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \Pr(\text{morir solamente por la causa } j \text{ en } (t, t + \Delta t) \text{ y actuando todos los riesgos en } (t, t + \Delta t) \mid \text{sobrevivo a todas las causas hasta } t)$$

y representa la tasa instantánea para el  $j$ -ésimo fallo en el tiempo  $t$ , y en presencia de todos los otros tipos de fallos, condicionada a que no ha ocurrido ninguno de ellos hasta ese tiempo.

Esta función aparece con el nombre de función de riesgo en trabajos anteriores a Crowder (1994) quien la llamó función de sub-riesgo.

Se define la función de incidencia acumulada o función de sub-distribución en el tiempo  $t$  por la causa  $j$  y en presencia de todas las causas actuando simultáneamente como

$$F_j(t) = \Pr(T \leq t, C = j) \quad j = 1, \dots, k,$$

De manera análoga, se define la función de sub-supervivencia como

$$S_j(t) = \Pr(T > t, C = j) \quad j = 1, \dots, k$$

la cual nos describe la probabilidad de morir eventualmente debido a la causa  $j$  en un tiempo mayor que  $t$ .

Debemos indicar (aunque la demostración no se expone aquí) que las funciones  $F_j(t)$  y  $S_j(t)$  no son distribuciones propias. Sin embargo, se pueden definir distribuciones condicionales propias en términos de esas sub-distribuciones. Por ejemplo, la función de distribución de supervivencia (FDS) condicional propia (en presencia de todas las causas) asociada con la causa  $j$  y denotada por  $S_j^*(t)$ , es

$$S_j^*(t) = \Pr(T > t \mid C = j)$$

entonces

$$F_j^*(t) = 1 - S_j^*(t) = \Pr(T \leq t \mid C = j)$$

denota la probabilidad condicional de presentar un fallo antes del tiempo  $t$  y en presencia de todas las causas, dado que el individuo tendrá un fallo por la causa  $j$ .

Los riesgos acumulados de causa-específica  $\Lambda_j(t)$ , el riesgo general  $\lambda(t)$ , el riesgo acumulado total  $\Lambda(t)$  y la función general de supervivencia  $S(t)$  se definen, respectivamente, como:

$$\begin{aligned}\Lambda_j(t) &= \int_0^t \lambda_j(u) du \quad j = 1, \dots, k, \\ \lambda(t) &= \lim_{\Delta t \rightarrow 0} \frac{\Pr(T < t + \Delta t \mid T \geq t)}{\Delta t} = \sum_{j=1}^k \lambda_j(t), \\ \Lambda(t) &= \int_0^t \lambda(u) du = \sum_{j=1}^k \Lambda_j(t), \\ S(t) &= \Pr(T > t) = e^{-\Lambda(t)}.\end{aligned}$$

La función de supervivencia puede factorizarse como el producto de las  $k$  funciones de sub-supervivencia,  $S_j(t) = e^{-\Lambda_j(t)}$ , como sigue:

$$S(t) = e^{-\sum_{j=1}^k \Lambda_j(t)} = \prod_{j=1}^k e^{-\Lambda_j(t)} = \prod_{j=1}^k S_j(t).$$

Como ya hemos comentado, es necesario tener precaución al interpretar las funciones  $S_j(t)$ . A pesar de tener las propiedades matemáticas de las funciones de supervivencia continuas, no son las funciones de supervivencia de ninguna de las variables aleatorias observables y ya sabemos que,  $S_j(t) \neq 1 - F_j(t)$ , (en las secciones 2.3 y 3, se ofrecen más detalles sobre su interpretación).

Las funciones de sub-densidad  $f_j(t)$  asociadas a la causa  $j$ , la distribución marginal  $F(t)$  de  $T$ , y la distribución de  $C$  se dan respectivamente por:



$$f_j(t) = \frac{d}{dt}F_j(t) = \lambda_j(t)S(t),$$

$$F(t) = P(T \leq t) = \sum_{j=1}^k F_j(t),$$

$$\pi_j(t) = Pr(C = j) = \lim_{t \rightarrow \infty} F_j(t) \quad j = 1, \dots, k.$$

La función de incidencia acumulada por la causa  $j$ ,  $F_j(t)$ , puede obtenerse del riesgo de causa-específica  $\lambda_j$  y la función de supervivencia general  $S(t)$  de la relación:

$$F_j(t) = \int_0^t \lambda_j(u)S(u)du \quad j = 1, \dots, k.$$

Igualmente, podemos obtener la función de sub-supervivencia como

$$S_j(t) = \int_t^\infty \lambda_j(u)S(u) du \quad j = 1, \dots, k$$

## 2.2 Representaciones alternativas de los riesgos competitivos

Un modo diferente de describir una situación de riesgos competitivos con  $k$  causas de fallo, es considerar un tiempo de fallo  $T_j$  para cada causa,  $j \in \{1, \dots, k\}$ . Estos tiempos son variables latentes correspondientes a tiempos de fallo hipotéticos, si las otras causas de fallo no estuvieran presentes.

Sería interesante poder especificar un modelo multivariante  $F(t_1, \dots, t_k)$  para la distribución conjunta de  $T_1, \dots, T_k$  (ver Kalbfleisch and Prentice, 2002; Lawless, 2003; Andersen et al., 2002 para otras referencias). Sin embargo, cuando todos los riesgos están presentes sólo se puede observar  $T = \min(T_1, \dots, T_k)$  junto con  $C = j$ , tal que  $T = T_j$ , y tenemos entonces un problema de identificabilidad del modelo (Tsiatis, 1975; Cox and Oakes, 1984).  $F(t_1, \dots, t_k)$  es un valor incalculable que está basado en estas observaciones. Dos distribuciones diferentes para  $F(t_1, \dots, t_k)$  pueden resultar con la misma marginal para  $T_j$ . Sólo bajo fuertes supuestos tales como la independencia, es identificable la distribución multivariante. Sin embargo, este supuesto no es comprobable basándonos únicamente en los datos de riesgo competitivos observados, y en determinadas circunstancias resulta una hipótesis que no puede sostenerse.

Por otra parte, se pueden ver los riesgos competitivos como un caso especial de un modelo multi-estado (Andersen et al., 2002). En este caso, el modelo multi-estado tiene un estado transitorio "vivo" y  $k$  estados absorbentes "fracaso debido a la causa  $j$ ", con  $j \in \{1, \dots, k\}$ . En esta configuración, el objetivo es ajustar las probabilidades de transición desde el estado "vivo" a cualquier estado absorbentes, a través de las funciones  $P_{hj}(s, t)$ , probabilidad de estar en el estado  $j$  en el momento  $t$ , dado que en el momento  $s$  el proceso ocupaba el estado  $h$ . En este caso notamos  $P_{0j}(0, t) = P(T \leq t, C = j)$  son las funciones de incidencia acumulada definidas en la sección 2.1, mientras que las funciones de intensidad de transición son los riesgos de causa-específica.

## 2.3 Función de verosimilitud

Consideremos una muestra aleatoria de  $n$  individuos, y sea  $(T_1, C_1), \dots, (T_n, C_n)$ , donde  $T_i$  es el tiempo de fallo y  $C_i$  es la causa de fallo para el sujeto  $i$ . Para cada individuo, existe un tiempo no negativo de censura a la derecha  $V_i$ . Sea  $\delta_i = I(T_i \leq V_i)$  el indicador de censura, y definimos  $\tilde{C}_i = \delta_i C_i$ .  $\tilde{C}_i$  es la causa de fallo para individuos que han fallado ó 0 para individuos censurados. Los datos observados para cada individuo son:

$$\{Y_i = \min(T_i, V_i), \delta_i, \tilde{C}_i, i = 1, \dots, n\}.$$

En lo que sigue, asumiremos estas condiciones:

H1.  $V_i$  es independiente de  $(T_i, C_i)$ .

H2. Si  $Y_i = T_i$  ( $T_i$  no está censurado), entonces se observa  $C_i$  (excluimos los casos donde el tiempo de fallo se observa, pero la información sobre la causa del fallo no está disponible).

H3. Los dominios de  $T$  y  $V$  son disjuntos.

Para obtener la función de verosimilitud, debe tenerse en cuenta la contribución de cada individuo. Para aclarar mejor este punto, primero desarrollaremos la función de verosimilitud para  $n = 1$ , y entonces para  $n > 1$ .

Muestra de tamaño  $n = 1$ :

Solamente se observa un individuo con los datos  $(Y, \delta, \tilde{C})$ . Existen dos posibilidades:

**A)** El individuo no está censurado,  $\delta = 1 \Rightarrow (Y, \delta, \tilde{C}) = (y, 1, j)$ ,  $j \neq 0$  (H2):

Para el caso discreto,

$$\begin{aligned} P\{y, 1, j\} &= P(Y = y, \delta = 1, C = j) = P(T = y, T \leq V, C = j) = P(T = y, V \geq y, C = j) \\ &\stackrel{H1}{=} P(T = y, C = j)P(V \geq y). \end{aligned}$$

Si  $T$  y  $V$  son variables aleatorias continuas, y  $q(v)$  y  $Q(v)$  son las funciones de densidad y supervivencia para  $V$ , la contribución del individuo a la función de verosimilitud vendría dada por  $f_j(y)Q(y)$ .

**B)** Si el individuo está censurado,  $\delta = 0 \Rightarrow (Y, \delta, \tilde{C}) = (y, 0, 0)$ :

Para el caso discreto,

$$\begin{aligned} P\{y, 0, 0\} &= P(Y = y, \delta = 0, \tilde{C} = 0) = P(Y = y, \delta = 0) = P(V = y, T > V) = P(V = y, T > y) \\ &\stackrel{H1}{=} P(T > y)P(V = y). \end{aligned}$$

Si  $T$  y  $V$  son variables aleatorias continuas como se definen en A), la contribución del individuo a la función de verosimilitud será  $S(y)q(y)$ .

Por lo tanto, la función de verosimilitud para un determinado individuo tiene la forma

$$\mathcal{L} = (f_c(y)Q(y))^\delta (S(y)q(y))^{1-\delta}$$

Muestra de tamaño  $n > 1$ :

Evaluamos la contribución de cada individuo en la verosimilitud. Si el sujeto  $i$  falla en  $y_i$  por la causa  $c_i = j$ , entonces su contribución a la verosimilitud vendrá dada por  $f_j(y_i)$ , la función de densidad de causa-específica para la causa  $j$ . Por otra parte, si el individuo  $i$  está censurado en el tiempo  $y_i$ , el individuo está todavía en riesgo para cualquier causa, y por lo tanto su contribución a la verosimilitud viene dada por la supervivencia global  $S(y_i)$ .

La función de verosimilitud para esta muestra viene dada entonces por

$$\mathcal{L} = \prod_{i=1}^n f_{c_i}(y_i)^{\delta_i} S(y_i)^{1-\delta_i} \prod_{i=1}^n Q(y_i)^{\delta_i} q(y_i)^{1-\delta_i}$$

Puesto que el tiempo de censura  $V$  es independiente del tiempo de fallo  $T$ , y sus dominios son disjuntos (hipótesis H1 y H3, respectivamente), los términos de censura en la verosimilitud no proporcionan información sobre el proceso de fallo, y se pueden eliminar. La función de verosimilitud es entonces proporcional a

$$\mathcal{L} \propto L = \prod_{i=1}^n f_{c_i}(y_i)^{\delta_i} S(y_i)^{1-\delta_i}$$

Denotemos por  $\delta_{ij} = I(C_i = j)$ , donde  $\delta_i = \sum_{j=1}^k \delta_{ij}$ . Si  $\delta_i = 1$ , entonces existe alguna  $j$  donde sea  $\delta_{ij} = 1$ . Considerando la factorización de la supervivencia  $S(t) = \prod_{j=1}^k S_j(t)$  (ver sección 2.1), y definiendo  $g_j(t) = -S'_j(t) = \lambda_j(t)S_j(t)$ , la función de verosimilitud puede reescribirse como producto de  $k$  componentes separados para cada causa de fallo:

$$\begin{aligned} L &= \prod_{i=1}^n \left( \prod_{j=1}^k f_j(y_i)^{\delta_{ij}} \right) S(y_i)^{1-\delta_i} = \prod_{i=1}^n \left( \prod_{j=1}^k (\lambda_j(y_i)S(y_i))^{\delta_{ij}} \right) S(y_i)^{1-\delta_i} \\ &= \prod_{i=1}^n \left\{ \left( \prod_{j=1}^k (\lambda_j(y_i)S(y_i))^{\delta_{ij}} \right) \left( \prod_{j=1}^k S_j(y_i)^{1-\delta_{ij}} \right) \right\} \\ &= \prod_{i=1}^n \prod_{j=1}^k \left( \lambda_j(y_i)^{\delta_{ij}} \left[ \prod_{\ell=1}^k S_\ell(y_i)^{\delta_{ij}} \right] S_j(y_i)^{1-\delta_{ij}} \right) \\ &= \prod_{i=1}^n \prod_{j=1}^k \lambda_j(y_i)^{\delta_{ij}} S_j(y_i)^{\delta_{ij}} S_j(y_i)^{\sum_{\ell \neq j} \delta_{i\ell} + 1 - \delta_i} \\ &= \prod_{i=1}^n \prod_{j=1}^k g_j(y_i)^{\delta_{ij}} S_j(y_i)^{1-\delta_{ij}} = \prod_{j=1}^k \left( \prod_{i=1}^n g_j(y_i)^{\delta_{ij}} S_j(y_i)^{1-\delta_{ij}} \right) = \prod_{j=1}^k L_j. \end{aligned}$$

Esta expresión proporciona una factorización de la verosimilitud global  $L$  en términos de las verosimilitudes de causa-específica  $L_j$ . Notemos que  $L_j$  corresponde a la

verosimilitud que se obtendría de la muestra  $\{(Y_i, \delta_{ij}), i = 1 \dots, n\}$  si el tiempo de fallo de otras causas fuera considerado como tiempo de censura, y donde el riesgo correspondiente, y las funciones de densidad y supervivencia son, respectivamente,  $\lambda_j(t)$ ,  $g_j(t)$  y  $S_j(t)$ . Esta factorización muestra cómo  $\lambda_j(t)$  y  $\Lambda_j(t)$  se estiman directamente de los datos  $(Y_i, \delta_{ij})$ , tratando los fallos por otras causas como si fuesen observaciones censuradas en  $Y_i$ . Sin embargo, como ya hemos mencionado anteriormente,  $S_j(t)$  **no** corresponde a **ninguna** variable aleatoria observable, es decir, no existe ningún variable aleatoria observada  $\mathcal{U}_j$  tal que  $P(\mathcal{U}_j > t) = S_j(t)$ .

## 2.4 Estimación no paramétrica

Consideremos una muestra aleatoria con los datos observados  $\{(Y_i, \delta_i, \tilde{C}_i), i = 1 \dots, k\}$ , definida como en la sección previa 2.3. Sean  $0 < y_1 < \dots < y_N$  los distintos puntos de tiempo observados y ordenados. Definimos las siguientes cantidades:

- $d_{ij}$  es el número de sujetos que fallan por la causa  $j$  en el tiempo  $y_i$ .
- $d_i = \sum_{j=1}^k d_{ij}$  es el número de sujetos que fallan en el tiempo  $y_i$  por cualquier causa.
- $n_i = \sum_{l=1}^n I_l(y_i)$ , con  $I_l(t) = I(y_l \geq t)$ , es el número de individuos en riesgo en  $y_i$ , es decir, vivos y sin censura justo antes de este tiempo.

El riesgo de causa-específica estimado para la causa  $j$  en el tiempo  $y_i$  viene dado por  $\hat{\lambda}_j(y_i) = \frac{d_{ij}}{n_i}$ , y es 0 en cualquier otro tiempo. Por lo tanto, el estimador de Nelson-Aalen de la función de riesgo acumulada de causa-específica de tipo  $j$ , viene dado por

$$\hat{\Lambda}_j(t) = \sum_{i: t_i \leq t} \frac{d_{ij}}{n_i} \quad j = 1, \dots, k,$$

con varianza estimada por

$$\widehat{\text{Var}}[\hat{\Lambda}_j(t)] = \sum_{i: t_i \leq t} \frac{d_{ij}}{n_i^2} \quad j = 1, \dots, k.$$

De esta manera, el estimador de Nelson-Aalen para la función de riesgo global  $\Lambda(t)$  está dado por

$$\hat{\Lambda}(t) = \sum_{j=1}^k \hat{\Lambda}_j(t)$$

La función de supervivencia global para  $T$  puede obtenerse usando el estimador de Kaplan-Meier:

$$\hat{S}(t) = \prod_{i: y_i < t} \left(1 - \frac{d_i}{n_i}\right)^{\delta_i}$$

Alternativamente, el estimador de Nelson-Aalen para la función de riesgo global, proporciona un estimador alternativo de  $S(t)$ , que es  $\hat{S}(t) = \exp[-\sum_{j=1}^k \hat{\Lambda}_j(t)]$ .

Así, dado que la función de incidencia acumulada por la causa  $j$  puede obtenerse a través del riesgo de causa-específica  $F_j(t) = \int_0^t \lambda_j(u)S(u)du$ , un estimador natural no paramétrico de  $F_j(t)$  sería:

$$\hat{F}_j(t) = \int_0^t \hat{\lambda}_j(u)\hat{S}(u)du = \sum_{i:y_i \leq t} \frac{d_{ij}}{n_i} \hat{S}(y_i^-) \quad j = 1, \dots, k.$$

donde  $\hat{S}(y_i^-) = \lim_{\varepsilon \rightarrow 0^+} \hat{S}(y_i - \varepsilon) = \hat{S}(y_{i-1})$ .

### 3 Interpretación del modelo

En las secciones anteriores hemos mostrado cómo especificar un modelo de riesgo competitivo a través de los riesgos de causa-específica  $\lambda_j(t)$  o mediante las funciones de incidencia acumulada  $F_j(t)$ , y cómo estimarlos de forma no paramétrica. En comparación con el análisis de supervivencia clásico, parece extraño utilizar la función de incidencia acumulada  $F_j(t)$ , en lugar de algún tipo de función de supervivencia de causa-específica para la causa  $j$ . En el análisis de supervivencia clásico, un tiempo de vida  $T$  está descrito usualmente por su función de supervivencia  $S(t) = P(T > t)$ , que satisface que  $S(t) = 1 - F(t)$ , siendo  $F(t)$  su función de distribución. La función de supervivencia podría obtenerse de la función de riesgo de  $T$ ,  $\lambda(t)$ , por

$$S(t) = e^{-\int_0^t \lambda(u) du}$$

Por analogía, en riesgos competitivos, dado el riesgo de causa-específica para la causa  $j$ ,  $\lambda_j(t)$ , se podría considerar una función similar  $S_j(t)$ , para cada causa de fallo:

$$S_j(t) = e^{-\int_0^t \lambda_j(u) du} = e^{-\Lambda_j(t)}$$

Sin embargo, como hemos mencionado antes, estas funciones no tienen el significado habitual de una función de supervivencia en el enfoque clásico. Además, las funciones  $S_j(t)$  no se corresponden con las complementarias de las funciones de incidencia  $F_j(t)$ , es decir,  $S_j(t) \neq 1 - F_j(t)$ , tampoco con la probabilidad conjunta de fallo debido a la causa  $j$  después del tiempo  $t$ ,  $P(T > t, C = j)$ . Estas consideraciones nos llevan a definir dos funciones más que pueden desempeñar el papel de supervivencia por causa-específica.

Por un lado, definimos  $S^*(t)$  como la complementaria de la función de incidencia acumulada

$$S_j^*(t) = 1 - F_j(t)$$

por otro lado, y por analogía con la definición de  $F_j(t)$ , definimos

$$\tilde{S}_j(t) = P(T > t, C = j)$$

Estas dos funciones nuevas son, como era  $S_j(t)$ , funciones que satisfacen las propiedades matemáticas de una función de supervivencia. En lo que sigue, vamos a profundizar en la interpretación de estas tres funciones,  $S_j(t)$ ,  $S_j^*(t)$  y  $\tilde{S}_j(t)$ , y argumentaremos por qué no son funciones propias de supervivencia, y cuál es la relación entre ellas. Para esta cuestión vale la pena recordar que una función  $S(t)$  es una función de supervivencia si

- i. Está definida en  $[0, \infty]$ .
- ii. Es no negativa y no creciente
- iii. Es continua por la derecha

iv.  $S(0) = 1$  y  $\lim_{t \rightarrow \infty} S(t) = 0$

Además,  $S(t)$  es una función de supervivencia de una variable aleatoria  $T$  si  $S(t) = P(T > t)$ .

### 3.1 La función $S_j^*(t) = 1 - F_j(t)$

$S_j^*(t) = 1 - F_j(t)$  representa la probabilidad de no fallar debido a la causa  $j$  antes del tiempo  $t$ . No es una función de supervivencia propia porque

$$\lim_{t \rightarrow \infty} S_j^*(t) = 1 - \lim_{t \rightarrow \infty} F_j(t) = 1 - P(C = j),$$

lo cual es estrictamente positivo si hay al menos dos causas de fallo. Por otra parte,

$$S_j^*(t) = 1 - F_j(t) = 1 - F(t) + \sum_{\ell \neq j} F_\ell(t) = S(t) + \sum_{\ell \neq j} P(T \leq t, C = \ell)$$

Es decir, la probabilidad de no fallar debido a la causa  $j$  antes del tiempo  $t$  es la suma de la probabilidad de no haber fallado por ninguna causa en  $t$  más la probabilidad de haber fallado antes de  $t$  debido a otras causas distintas a  $j$ . Esta probabilidad  $S_j^*(t)$  se usa para construir el modelo de regresión de Fine y Gray (1999) para la función de incidencia acumulada.

### 3.2 La función $\tilde{S}_j(t) = P(T > t, C = j)$

Por analogía con el modo en que las funciones de incidencia acumulada  $F_j$  fueron definidas,  $\tilde{S}_j(t) = P(T > t, C = j)$  representa la probabilidad de fallar debido a la causa  $j$  después del tiempo  $t$ . No es una función propia de supervivencia porque

$$\tilde{S}_j(0) = P(C = j)$$

que es estrictamente inferior a 1 si hay al menos dos causas de fallo.

La relación con  $F_j(t)$  viene dada por

$$\begin{aligned} \tilde{S}_j(t) &= P(T > t, C = j) = Pr(T > j | C = j)P(C = j) = [1 - P(T \leq t | C = j)]P(C = j) \\ &= P(C = j) - P(T \leq t, C = j) = P(C = j) - F_j(t). \end{aligned}$$

Por lo tanto, se comporta como una probabilidad complementaria de  $F_j(t)$ , complementaria en la probabilidad de fallar debido a la causa  $j$ ,  $P(C = j)$ . Notemos que  $S(t)$  podría descomponerse en términos de  $\tilde{S}_j(t)$  de la siguiente manera:

$$S(t) = 1 - F(t) = 1 - \sum_{j=1}^k F_j(t) = 1 - \sum_{j=1}^k P(C = j) + \sum_{j=1}^k P(T > t, C = j) = \sum_{j=1}^k \tilde{S}_j(t).$$

La expresión de  $S(t)$  como la suma de  $\tilde{S}_j(t)$  es, en efecto, diferente de la descomposición alternativa  $S(t) = \prod_{j=1}^k S_j(t)$  y muestra que  $\tilde{S}_j(t)$  y  $S_j(t)$  son diferentes.

Un estimador consistente para  $\tilde{S}_j(t)$  viene dado por

$$\widehat{\tilde{S}}_j(t) = \frac{1}{n} \sum_{i=1}^n I(Y_i > t, C_i = j) \quad j = 1, \dots, k.$$

A pesar de que estas funciones son estimables y podrían especificar el modelo de riesgos competitivos, se han utilizado poco en la literatura de riesgos competitivos (Peterson, 1976).

### 3.3 Interpretación de la función $S_j(t) = e^{-\Lambda_j(t)}$

Nos hemos encontrado con las funciones  $S_j(t)$  repetidas veces en las secciones previas. En primer lugar, las hemos encontrado en la factorización de la función de supervivencia  $S(t)$ , en la sección 2.1. Después, en la factorización de la función de verosimilitud de la sección 2.3, donde  $S_j(t)$  corresponde a la función de supervivencia obtenida de la función de riesgo acumulada  $\Lambda_j(t)$ , el riesgo acumulado cuando los tiempos de fallo de las otras causas se consideran como tiempos censurados. Las funciones  $S_j(t)$  tienen las propiedades matemáticas de la función de supervivencia, sin embargo no son funciones de supervivencia de ninguna variable aleatoria observable. Cuando los fallos debidos a otras causas son tratados como observaciones censuradas, la hipótesis de independencia entre el tiempo de fallo y el tiempo de censura posiblemente no se cumpla. Sólo cuando se supone que las distintas causas de fallo son independientes,  $1 - S_j(t)$  se interpreta totalmente como la probabilidad de fallar debido a la causa  $j$  si las otras causas de fallo fueran eliminadas (Gooley et al., 1999).

A menudo  $1 - S_j(t)$  se ha usado incorrectamente para estimar  $F_j(t)$ , debido en parte a la disponibilidad de software para obtener el estimador de Kaplan-Meier para  $S_j(t)$ :

$$KM_j(t) = \prod_{i: y_i < t} \left(1 - \frac{d_{ji}}{n_i}\right)^{\delta_{ij}}$$

donde  $d_{ij}$ ,  $n_i$  y  $\delta_{ij}$  se definen en la sección 2.4 y los fallos debido a otras causas se tratan como observaciones censuradas. Sin embargo,  $1 - KM_j(t)$  proporciona una estimación sesgada de la probabilidad acumulada de fallo del tipo  $j$ ,  $F_j(t)$  (Putter et al., 2007). Esto es intuitivamente claro, ya que  $S_j(t)$  sólo depende del riesgo de causa-específica  $\lambda_j(t)$ , mientras que  $F_j(t)$  depende de todas las causas específicas  $\hat{\lambda}_l(t)$ ,  $l \in \{1, \dots, k\}$  a través de la función de supervivencia  $S(t)$  (ver sección 2.1). Además,  $1 - KM_j(t)$  como estimador de  $1 - S_j(t)$  sobrestima la probabilidad de fallar debido a la causa  $j$ ,  $F_j(t)$ . Esto es razonable, porque si un individuo falla debido a otras causas, se trata como una observación censurada, se asume que el individuo **fallará** debido a la causa de interés  $j$  en algún momento futuro, lo que en algunas situaciones pueden ser



inviabile: si un individuo muere debido a cáncer, ciertamente no moriría (otra vez), debido a un ataque al corazón. Al censurar los individuos, se espera una mayor incidencia de fallos. En efecto, siempre existe  $t^* > 0$  tal que

$$F_j(t^*) = \int_0^{t^*} S(u)\lambda_j(u)du < \int_0^{t^*} S_j(u)\lambda_j(u) = 1 - S_j(t^*).$$

Para comprobar esto razonamos de la siguiente forma:

- Siempre existe  $l \neq j$  y  $t^* > 0$  tal que  $\Lambda_l(t^*) = \int_0^{t^*} \lambda_l(u)du > 0$ . Es decir, existe al menos otra causa de fallo con al menos un fallo. De lo contrario, no hay riesgos competitivos en nuestros datos.
- Por lo tanto,  $\Lambda(t^*) = \sum_{m=0}^k \Lambda_m(t^*) > \Lambda_j(t^*)$ .
- Dado  $g(u) = e^{-u}$  no creciente y  $\Lambda_j(u)$  no negativo,

$$S_j(t^*) = e^{-\Lambda_j(t^*)} > e^{-\Lambda(t^*)} = S(t^*)$$

cqd.

## 4. Modelos de regresión de riesgos competitivos

En un análisis de supervivencia con riesgos competitivos, son posibles dos estrategias diferentes de modelos de regresión: los modelos de los riesgos de causa-específica o los modelos de funciones de incidencia acumulada.

Cuando se modelizan los riesgos de causa-específica, cada riesgo se analiza por separado tratando los individuos que fallan por otras causas como observaciones censuradas, como se desprende de la factorización de la función de verosimilitud dada en la sección 2.3. Este enfoque es adecuado cuando interesa determinar los factores asociados al riesgo de una causa específica de fallo.

Por otra parte, las funciones de incidencia acumulada se utilizan para determinar los factores asociados a la incidencia de una causa dada. Este análisis no trata de los individuos que fallan por otras causas como las observaciones censuradas.

En las siguientes secciones, se revisan tres modelos específicos: Modelo de riesgos proporcionales de Cox (Prentice et al, 1978), el modelo aditivo de Aalen (Aalen, 1993), que especifica modelos para los riesgos de causa-específica, y el enfoque dado por Fine y Gray (1999), que se basa en las funciones de incidencia acumulada.

### 4.1 Modelos para las funciones de riesgo de causa-específica $\lambda_j(t)$

#### 4.1.1 Modelo de riesgos proporcionales de Cox

El análisis de regresión clásico de los riesgos competitivos establece un modelo de riesgos proporcionales de Cox (PH) (Prentice et al, 1978.) para cada riesgo de causa específica:

$$\lambda_j(t|Z) = \lambda_{0j} e^{\beta_j' Z} \quad j = 1, \dots, k,$$

donde  $Z$  es un vector de  $p$  covariables y  $\beta_j$  es un vector de  $p$  coeficientes de regresión para cada causa. Cada causa de fallo se analiza separadamente, tratando los individuos que fallan por otras causas como observaciones censuradas. Se asume que el efecto de las covariables actúa multiplicativamente en base a una función de riesgo desconocida  $\lambda_{0j}$ . Como en un análisis de riesgos proporcionales clásico (PH), la validez de los modelos no depende de la verdadera forma de la función de riesgo de referencia, siempre que la forma multiplicativa del modelo sea correcta. La hipótesis de PH es una fuerte hipótesis que debe comprobarse cuidadosamente para cada causa.

La estimación de los parámetros de regresión  $\beta_j$  se basa en el método de verosimilitudes parciales. Supongamos que una muestra aleatoria censurada  $(y_i, \delta_i, \delta_i c_i)$   $i = 1, \dots, n$ , tiene  $N$  tiempos distintos de fallo observados  $t_1 < \dots < t_N$  y  $n - N$  tiempos de censura (no considerados aquí). Consideremos la probabilidad de que un individuo falle por la causa  $j$  en el tiempo  $t_i$ , dado que uno de los individuos en riesgo (vivo y sin censura) falla en el tiempo  $t_i$  debido a la causa  $j$ :

$$\frac{e^{\beta_j' Z_i}}{\sum_{\ell=1}^n Y_{\ell}(t_i) e^{\beta_j' Z_{\ell}'}}$$

donde  $Y_{\ell}(t) = I(t_{\ell} \geq t)$ . La función parcial de verosimilitud se define sólo en los  $N$  tiempos de fallo, obteniéndose:

$$L(\beta_1, \dots, \beta_k) = \prod_{i=1}^n \prod_{j=1}^k \left( \frac{e^{\beta_j' Z_i}}{\sum_{\ell=1}^n Y_{\ell}(t_i) e^{\beta_j' Z_{\ell}'}} \right)^{\delta_{ij}} = \prod_{j=1}^k L_j(\beta_j)$$

donde  $\delta_{ij} = I(C_i = j)$ . El grupo de riesgo puede disminuirse por la ocurrencia de un evento por cualquier causa.

Maximizar cada factor en la ecuación anterior proporciona un estimador  $\hat{\beta}_j$  consistente y asintóticamente normal bajo condiciones apropiadas, y el vector score, la matriz de información y el estadístico de razón de verosimilitudes basados en  $L(\hat{\beta}_j)$  se comportan como si fueran deducidos de una verosimilitud ordinaria.

Dado  $\hat{\beta}_j$ , el estimador de Nelson-Aalen generalizado estima que las funciones de riesgo acumulado base relativas a cada causa-específica son:

$$\hat{\Lambda}_{0j}(t) = \sum_{i:t_i \leq t} \left( \frac{\delta_{ij}}{\sum_{\ell=1}^n Y_{\ell}(t_i) e^{\hat{\beta}_j' Z_{\ell}'}} \right) \quad j = 1, \dots, k.$$

La inferencia para las  $\beta_j$  y para las  $\Lambda_{0j}$  puede llevarse a cabo al igual que en el modelo estándar de Cox donde se considera un solo tipo de fallo. La supervivencia global y las funciones de riesgo acumulado para  $T$  dada  $Z$ , se obtienen por

$$\hat{S}(t|Z) = \exp \left\{ - \sum_{j=1}^k \hat{\Lambda}_{0j}(t) e^{\hat{\beta}_j' Z} \right\}$$

$$\hat{\Lambda}_j(t|Z) = \hat{\Lambda}_{0j}(t) e^{\hat{\beta}_j' Z} \quad j = 1, \dots, k.$$

Finalmente, la función de incidencia acumulada  $F_j(t|Z)$  puede obtenerse sustituyendo los estimadores anteriores en la expresión correspondiente de  $F_j(t|Z)$  indicada en secciones anteriores

$$\hat{F}_j(t|Z) = \int_0^t \hat{S}(u|Z) d\hat{\Lambda}_j(u|Z)$$

$$= \sum_{i:t_i \leq t} \delta_{ij} \exp \left\{ - \sum_{\ell=1}^k \hat{\Lambda}_{0\ell}(u) e^{\hat{\beta}_{\ell}' Z} \right\} \frac{e^{\hat{\beta}_j' Z}}{\sum_{r=1}^n Y_r(t_i) e^{\hat{\beta}_j' Z_r}}$$

La metodología propuesta es completamente estándar, pero se necesita cierta cautela al interpretar los modelos. Incluso cuando interesa sólo una de las causas de fallo, no es suficiente llevar a cabo un análisis simple sólo para esta causa. Es necesario modelar todas las causas de fallo con el fin de realizar una interpretación completa y adecuada del proceso de fallo.

#### 4.1.2 Modelo de riesgos aditivos de Aalen

La metodología de Cox ha sido discutida ampliamente, convirtiéndose en el análisis estándar para llevar a cabo modelos de regresión (véase Lawless (2003) por ejemplo). Sin embargo, en largos estudios de seguimiento, es natural pensar que el efecto de una covariable en el riesgo puede cambiar con el tiempo. En esta situación, los modelos con parámetros constantes a lo largo del tiempo, como el modelo de Cox, pueden ser inadecuados. La metodología propuesta por Aalen (1993, 2001) encuentra una alternativa, en la que se especifica el modelo de riesgos aditivos para cada riesgo de causa-específica involucrado:

$$\lambda_j(t|Z(t)) = \beta_{j0}(t) + Z(t)^t \beta_j(t) \quad j = 1, \dots, k$$

donde  $Z^t(t) = [Z_1(t), \dots, Z_p(t)]$  es un p-vector de covariables (posiblemente dependiente del tiempo) y  $\beta_j^t(t) = [\beta_{j1}(t), \dots, \beta_{jp}(t)]$  son funciones paramétricas desconocidas. A diferencia de la configuración de Cox, este modelo asume que las covariables actúan de manera aditiva sobre la función de riesgo base que se supone desconocida. Este efecto se evalúa a través de las funciones dependientes del tiempo  $\beta_j(t)$ , por lo que puede explorarse su variación a lo largo del tiempo.

Para obtener estimaciones de estas funciones se usan técnicas de mínimos cuadrados. Concretamente para obtener estimadores de las funciones acumuladas, es decir

$$\mathbf{B}(t) = (B_0(t), B_1(t), \dots, B_p(t)) = \left( \int_0^t \beta_0(u) du, \int_0^t \beta_1(u) du, \dots, \int_0^t \beta_p(u) du \right)$$

porque se obtienen más fácilmente, y entonces las estimaciones de  $\beta(t)$  se pueden obtener mediante las pendientes de  $\widehat{\mathbf{B}}(t)$ , o mediante el uso de un estimador suave tipo núcleo basado en ellas.

Para obtener estas estimaciones es usual formular el problema desde la metodología de los procesos de recuento. Es decir, consideremos el i-ésimo individuo, y definimos:

- Para cada causa  $j$ ,  $N_{ji}(t) = I(Y_i \leq t, C_i = j)$  es la función que determina si la j-ésima causa de fallo ya se ha producido en el tiempo  $t$ .
- Consideremos  $N_j(t) = (N_{ji}(t), i = 1, \dots, n)'$
- El indicador de riesgo para el i-ésimo individuo se define como  $Y_i(t) = I(Y_i \geq t)$ .
- Consideremos la  $n \times (p + 1)$  matriz de diseño de  $\mathbf{X}(t)$  con la i-ésima fila dada por

$$\mathbf{X}_i(t) = Y_i(t)(1, Z_{i1}(t), \dots, Z_{ip}(t)).$$

Así,  $\mathbf{X}_i(t) = (1, Z_{i1}(t), \dots, Z_{ip}(t))$  si el sujeto  $i$  está en riesgo en el tiempo  $t$ , y en cambio es un vector formado por  $p+1$  ceros si el sujeto no está en riesgo.

Los estimadores de mínimos cuadrados de  $\mathbf{B}_j(t) = (B_{j0}(t), B_{j1}(t), \dots, B_{jp}(t))$  para los parámetros de regresión para la causa  $j$  vienen dados por:

$$\hat{\mathbf{B}}_j(t) = \int_0^t \mathbf{X}^-(u) dN_j(u) \approx \sum_{y_i \leq t} \mathbf{X}^-(t_i) \mathbf{I}_j(y_i), \quad j = 1, \dots, k$$

donde  $\mathbf{X}^-(t) = (\mathbf{X}(t)' \mathbf{X}(t))^{-1} \mathbf{X}(t)'$  es una inversa generalizada de  $\mathbf{X}(t)$ , y  $\mathbf{I}_j(t)$  es el  $n \times 1$  vector con la  $i$ -ésima componente igual a 1 si el  $i$ -ésimo sujeto falla debido a la causa  $j$  en el tiempo  $t$ , y es 0 en otro caso. Notemos que el estimador  $\hat{\mathbf{B}}_j(t)$  sólo existe hasta el tiempo más corto  $t_r$  en el que  $\mathbf{X}(t_r)' \mathbf{X}(t_r)$  llega a ser una matriz singular.

La estimación de la matriz de varianza-covarianza de  $\mathbf{B}_j(t)$  es

$$\hat{\Sigma}_j(t) = \sum_{y_i \leq t} \mathbf{X}^-(y_i) \mathbf{I}_j^D(t_i) \mathbf{X}^-(y_i)'$$

donde  $\mathbf{I}_j^D(y_i)$  es una matriz diagonal, y los elementos de la diagonal son  $\mathbf{I}_j$ .

Al igual que en la sección anterior, la función de incidencia acumulada se obtiene mediante el uso de los riesgos de causa-específica y de las estimaciones de los riesgos acumulativos de referencia desde el modelo aditivo.

Los modelos aditivos son una opción conveniente cuando la hipótesis de riesgos proporcionales no se sostiene. El efecto de las covariables lo largo del tiempo puede evaluarse gráficamente, representando el exceso de riesgo acumulativo  $\hat{B}(t)$  para cada covariable y comparando las estimaciones para cada uno de los riesgos competitivos y para la supervivencia global.

Aunque estos modelos son muy flexibles y fáciles de implementar, son menos utilizados que el modelo de Cox porque la inferencia con respecto a sus términos no paramétricos no está completamente desarrollada y no está incluida en el software estadístico estándar. Ver Martinussen y Scheike (2006) para más detalles.

## 4.2 Modelos para la funciones de incidencia acumulada $F_j(t)$

Los modelos anteriores son especialmente interesantes cuando el objetivo es evaluar si un factor se asocia con el riesgo de una causa específica del fallo. Sin embargo, cuando el objetivo es comparar la incidencia observada de eventos de una causa dada entre grupos, deben usarse las funciones de incidencia acumulada.

Las estimaciones de estas funciones se pueden obtener a través de

$$\hat{F}_j(t|\mathbf{Z}) = \sum_{t_i \leq t} \hat{\lambda}_j(t_i|\mathbf{Z}) \hat{S}(t|\mathbf{Z}),$$

donde  $\hat{\lambda}_j(t|\mathbf{Z})$  son los riesgos estimados resultantes de los análisis de Cox o Aalen, y  $t_i$  los distintos tiempos de fallo. La función de supervivencia global es

$$\hat{S}(t|\mathbf{Z}) = \exp \left\{ - \sum_{j=1}^k \sum_{t_i \leq t} \hat{\lambda}_j(t_i|\mathbf{Z}) \right\}$$

El problema con este enfoque es que no estima directamente el efecto de una covariable en la función de incidencia acumulada  $F_j(t)$  dada. Aunque el efecto de las covariables en el riesgo de causa-específica  $\lambda_j(t|\mathbf{Z})$  está directamente determinado por  $\beta_j$ , el efecto sobre la función de incidencia acumulada  $F_j(t)$  combina el efecto de  $\beta_j$ , junto con el efecto global sobre  $\hat{S}(t|\mathbf{Z})$ . Además, no es posible probar los efectos significativos sobre las funciones de sub-distribución, porque algunas covariables pueden tener un efecto significativo sobre el riesgo, pero no sobre los  $F_j$ . Con el fin de realizar la selección del modelo y obtener estimaciones de los efectos de las covariables en las funciones de incidencia acumulada, se proponen los modelos basados directamente en las funciones de sub-distribución.

#### 4.2.1 El modelo de Fine y Gray

Fine y Gray (1999) consideran una nueva función, la función de sub-riesgo  $\gamma_j(t)$  derivada de la función de sub-distribución:

$$\begin{aligned} \gamma_j(t|\mathbf{Z}) &= \lim_{\Delta t \rightarrow 0} \frac{\Pr(T < t + \Delta t, C = j | \mathbf{Z}, \{T \geq t \text{ or } (T < t \text{ and } C \neq j)\})}{\Delta t} \\ &= \frac{f_j(t|\mathbf{Z})}{1 - F_j(t|\mathbf{Z})} \quad j = 1, \dots, k. \end{aligned}$$

Este sería el riesgo obtenido de  $F_j$  si se tratara de una distribución apropiada. La expresión condicional incluye dos casos diferentes:

- i) el evento no ha ocurrido en el tiempo  $t$ ,
- ii) el evento se ha producido por una causa diferente antes de  $t$ .

Así, el conjunto de riesgo en el tiempo  $t$  está formado por dos tipos de individuos, que corresponden a dos escenarios diferentes. A diferencia de los análisis basados en los riesgos de causa-específica, un paciente que falla por otras causas no sería eliminado del conjunto de riesgos en el momento en que su fallo es observado. La función de sub-distribución se expresa en términos de los sub-riesgos como

$$F_j(t|\mathbf{x}) = 1 - \exp \left( - \int_0^t \gamma_j(t|\mathbf{x}) \right).$$

Fine y Gray propusieron adaptar la función de sub-riesgo usando un modelo de Cox, es decir

$$\gamma_j(t|\mathbf{x}) = \gamma_{0j}(t)e^{\beta_j' \mathbf{x}}, \quad j = 1, \dots, k,$$

donde las covariables actúan linealmente sobre una transformación logarítmica de una complementaria de la función de incidencia acumulada. Cuando la censura está ausente o siempre es observable, Fine y Gray (1999) mostraron que el método de verosimilitud parcial es válido para la estimación.

Si hay  $N$  fallos en los tiempos  $t_1 < t_2 < \dots < t_N$ , la verosimilitud parcial se definió como

$$\bar{L}(\beta_j) = \prod_{i=1}^N \left( \frac{e^{\beta_j' \mathbf{Z}_i}}{\sum_{\ell \in \mathcal{R}_i} w_{i\ell} e^{\beta_j' \mathbf{Z}_\ell}} \right)$$

Ahora el conjunto de riesgo para la causa  $j$  en el momento  $t_i$  es  $\tilde{\mathcal{R}}_i = \{l: t_l \geq t_i \text{ o } (t_l \leq t_i \text{ y } C \neq j)\}$ , donde los sujetos que experimentan una causa alternativa permanecen en el grupo de riesgo. El peso  $w_{i\ell}$  dado a cada individuo es  $\tilde{G}(t_i) / \tilde{G}(\min(t_i, t_\ell))$ , donde  $\tilde{G}$  es la función de supervivencia para la distribución de la censura. Un individuo que satisface  $t_\ell \geq t_i$  tiene un peso de 1.

#### 4.2.2 Otros modelos basados en $F_j$

Se han propuesto otros modelos en la misma línea de trabajo de Fine y Gray. Fine (2001) consideró un modelo para la siguiente transformación de las funciones de incidencia acumulada

$$g(F_j(t|\mathbf{Z})) = \alpha_j(t) - \beta_j' \mathbf{Z},$$

donde  $g(\cdot)$  es una función diferenciable conocida, y  $\alpha_j(t)$  es la probabilidad de fallo de referencia cuando  $\mathbf{Z} = \mathbf{0}$ , que no está especificada, es invertible y estrictamente creciente en  $t$ . Para la estimación se usan técnicas de mínimos cuadrados.

## 5. Análisis estadístico mediante el uso del paquete “cmprsk” de R

Para el análisis estadístico de los modelos de riesgos competitivos es útil el uso del programa R, concretamente hablamos aquí del paquete de funciones *cmprsk*. Este paquete ha sido desarrollado por el Dr. Robert Gray y contiene los instrumentos necesarios para un análisis completo que tenga en cuenta los riesgos competitivos. Así, se pueden obtener gráficos de probabilidad para el evento de interés y un valor de p basado en la prueba de Gray, que es un test *log-rank* modificado para verificar la situación de riesgos competitivos.

El modelo aporta la posibilidad de verificar la proporcionalidad de los riesgos, y pueden incluirse términos para coeficientes dependientes del tiempo. El código no permite acomodar un truncamiento por la izquierda ni datos agrupados. El truncamiento por la izquierda sería útil para el análisis de eventos múltiples/recurrentes por paciente o para el análisis de cohortes.

El paquete ‘cmprsk’ se describe como un paquete de R para la estimación, análisis de regresión y modelización de funciones de sub-distribución en riesgos competitivos.

### 5.1. Documentación de funciones de ‘cmprsk’

#### 5.1.1 La función *crr*

Se usa para el análisis de regresión basado en las funciones de sub-distribución de los modelos en riesgos competitivos.

#### Sintaxis:

```
crr(ftime, fstatus, cov1, cov2, tf, cengroup, failcode=1, cencode=0, subset, na.action=na.omit, gtol=1e-06, maxiter=10, init)
```

#### Argumentos empleados en la sintaxis:

**ftime:** vector de tiempos de fallo/censura.

**fstatus:** vector con un código único para cada tipo de error y un código separado para las observaciones censuradas.

**cov1:** matriz (Nobs x ncovs) de las covariables fijas (ya sea cov1, cov2, o ambas cosas).

**cov2:** matriz de covariables que se multiplicará por funciones de tiempo. Estas covariables también aparecerán en cov1 para proporcionar un efecto de apoyo en los riesgos además de un tipo de interacción.

**tf:** funciones del tiempo. Una función que toma un vector de tiempo como un argumento y devuelve una matriz cuya j-ésima columna es el valor de la función de tiempo correspondiente a la columna j de cov2 evaluada en el vector de tiempo de entrada. En el momento tk, el modelo incluye el término  $\text{cov2} [ , j ] * \text{tf} (\text{tk}) [ , j ]$  como covariable.



**cengroup**: vector con valores diferentes para cada grupo con una distribución distinta de censura (la distribución de censura se estima por separado dentro de estos grupos). Todos los datos en un grupo, si faltan.

**failcode**: código de fstatus que indica el tipo de fallo de interés.

**cencode**: código de fstatus que denota observaciones censuradas.

**subset**: un vector lógico que especifica un subconjunto de casos para incluir en el análisis.

**na.action**: función que especifica la acción a llevar a cabo para aquellos casos faltantes de ftime, fstatus, cov1, cov2, cengroup, o subset.

**gtol**: iteración que se detiene cuando una función del gradiente es <gtol.

**maxiter**: número máximo de iteraciones en el algoritmo de Newton.

**init**: valores iniciales de los parámetros de regresión (por defecto = todo 0).

### Detalles

Se adapta el modelo de regresión de "riesgos de subdistribución proporcionales" descrito en Fine y Gray (1999). Este modelo evalúa directamente el efecto de las covariables sobre la subdistribución de un tipo particular de fallo en un entorno de riesgos competitivos. El método aplicado aquí se describe como la ecuación ponderada de estimación.

El modelo básico supone que la subdistribución con covariables  $z$  es constante. Esto se puede generalizar mediante la inclusión de las interacciones de  $z$  con funciones de tiempo para permitir que la magnitud cambie con el tiempo de seguimiento a través de cov2 y los argumentos tfs. Por ejemplo, si  $z$  es un vector de valores de covarianza y uft es un vector que contiene los tiempos de fallo únicos para los fallos del tipo de interés (en orden ascendente), entonces los coeficientes  $a$ ,  $b$  y  $c$  en el modelo de ecuación cuadrática (en tiempo)  $az + bzt + z t^2$  pueden especificarse como cov1= $z$ , cov2=cbind( $z,z$ ), tf=function(uft) cbind (uft, uft\*uft).

Esta función utiliza una estimación de la función de supervivencia de la distribución de censura para volver a considerar las contribuciones a los grupos de riesgo de los fallos por causas competitivas.

En una generalización de la metodología teórica, la distribución de la censura se puede estimar por separado dentro de los estratos definidos por el argumento *cengroup*. Si la distribución de la censura es diferente dentro de los grupos definidos por covariables en el modelo, entonces la validez del método requiere el uso de estimaciones separadas de la distribución de censura dentro de esos grupos.

Los residuos devueltos son análogos a los residuos de Schoenfeld en modelos de supervivencia ordinarios. Representando la  $j$ -ésima columna de "res" frente al vector de tiempos únicos de fallo, se puede comprobar la falta de ajuste en el tiempo para la correspondiente covariable (columna de cov1).

## Valor

Devuelve una lista de clase *crr*, con componentes

\$ **Coef**: los coeficientes de regresión estimados.

\$ **Loglik**: logaritmo de la pseudo-verosimilitud evaluada en coef.

\$ **score**: derivadas del logaritmo de la pseudo-verosimilitud evaluada en coef.

\$ **Inf**: segunda derivada del logaritmo de la pseudo-verosimilitud.

\$ **var**: matriz de varianza-covarianza estimada de coef.

\$ **res**: matriz de residuos dando la contribución a cada puntuación (columnas) en cada tiempo de fallo (filas) único.

\$ **uftime**: vector de tiempos de fallo únicos.

\$ **bfitj**: saltos en la estimación de tipo Breslow del riesgo de la sub-distribución subyacente acumulativa (utilizado por `predict.crr()`).

\$ **tfs**: la matriz tfs (salida de `tf()`, si se utiliza).

\$ **converged**: TRUE si el algoritmo iterativo converge.

\$ **call**: llamada a la función `crr`.

\$ **n**: El número de observaciones utilizadas en el ajuste del modelo.

\$ **n.missing**: El número de observaciones eliminadas de los datos de entrada, debido a los valores perdidos.

\$ **loglik.null**: El valor del logaritmo de la pseudo-verosimilitud cuando todos los coeficientes son 0.

## Ejemplos

Exponemos el siguiente ejemplo con R:

### Código

```
# simulated data to test
set.seed(10)
ftime <- rexp(200)
fstatus <- sample(0:2,200,replace=TRUE)
cov <- matrix(runif(600),nrow=200)
dimnames(cov)[[2]] <- c('x1','x2','x3')
print(z <- crr(ftime,fstatus,cov))
summary(z)
z.p <- predict(z,rbind(c(.1,.5,.8),c(.1,.5,.2)))
plot(z.p,lty=1,color=2:3)
crr(ftime,fstatus,cov,failcode=2)
# quadratic in time for first cov
crr(ftime,fstatus,cov,cbind(cov[,1],cov[,1]),function(Uft) cbind(Uft,Uft^2))
#additional examples in test.R
```

Al escribir el código obtenemos:

```

> #additional examples in test.R
> #simulated data to test
> set.seed(10)
> f.time<-rexp(200)
> fstatus<-sample(0:2,200,replace=TRUE)
> cov<-matrix(runif(600), nrow=200)
> dimnames(cov)[[2]]<-c('x1','x2','x3')
> print(z<-crr(f.time,fstatus,cov))

```

convergence: TRUE

coefficients:

	x1	x2	x3
	-0.6306	0.2024	0.6384

standard errors:

[1] 0.4043 0.3782 0.4105

two-sided p-values:

	x1	x2	x3
	0.12	0.59	0.12

```
> summary(z)
```

Competing Risks Regression

Call:

```
crr(f.time = f.time, fstatus = fstatus, cov1 = cov)
```

	coef	exp(coef)	se(coef)	z	p-value
x1	-0.631	0.532	0.404	-1.560	0.12
x2	0.202	1.224	0.378	0.535	0.59
x3	0.638	1.893	0.411	1.555	0.12

	exp(coef)	exp(-coef)	2.5%	97.5%
x1	0.532	1.879	0.241	1.18
x2	1.224	0.817	0.583	2.57
x3	1.893	0.528	0.847	4.23

Num. cases = 200

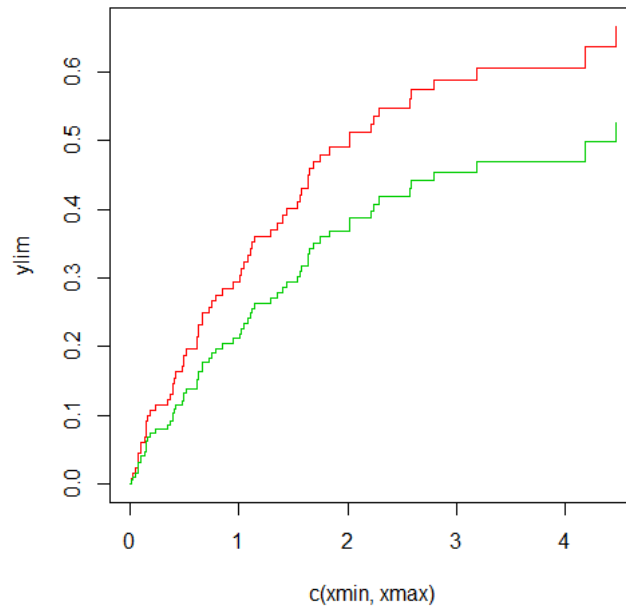
Pseudo Log-likelihood = -314

Pseudo likelihood ratio test = 5.36 on 3 df,

```

> z.p<-predict(z,rbind(c(.1,.5,.8),c(.1,.5,.2)))
> plot(z.p,lty=1,color=2:3)

```



```
> crr(ftime,fstatus,cov,failcode=2)
```

```
convergence: TRUE
```

```
coefficients:
```

x1	x2	x3
0.01017	-0.17020	-0.58140

```
standard errors:
```

```
[1] 0.4418 0.4799 0.4217
```

```
two-sided p-values:
```

x1	x2	x3
0.98	0.72	0.17

```
> #quadratic in time for first cov
```

```
> crr(ftime,fstatus,cov,cbind(cov[,1],cov[,1]),function(Uft) cbind(Uft,Uft^2))
```

```
convergence: TRUE
```

```
coefficients:
```

```

      x1          x2
-0.7257      0.1976
      x3      cbind(cov[, 1], cov[, 1])1*Uft
0.6392      0.3692
cbind(cov[, 1], cov[, 1])2*tf2
-0.1459
standard errors:
[1] 0.8403 0.3785 0.4116 1.1810 0.3142
two-sided p-values:
      x1          x2
0.39          0.60
      x3      cbind(cov[, 1], cov[, 1])1*Uft
0.12          0.75
cbind(cov[, 1], cov[, 1])2*tf2
0.64
> #additional examples in test.R

```

### 5.1.2 La función *cuminc* (Análisis de la incidencia acumulada)

Esta función sirve para estimar las funciones de incidencia acumulada de los datos de los riesgos competitivos y prueba la igualdad entre todos los grupos.

#### Sintaxis:

```
cuminc(ftime, fstatus, group, strata, rho=0, cencode=0, subset, na.action=na.omit)
```

#### Argumentos empleados en la sintaxis:

**ftime:** variable del tiempo de fallo.

**fstatus:** variable con códigos distintos para diferentes causas de fallo y también un código distinto para observaciones censuradas.

**group:** las estimaciones que se calculan dentro de los grupos dados por los distintos valores de esta variable. Las pruebas comparan estos grupos. Si este argumento está ausente se considera un solo grupo (no hay prueba estadísticas).

**strata:** variable de estratificación. No tiene ningún efecto sobre las estimaciones. Las pruebas serán estratificadas en esta variable. (Todos los datos en un estrato, si falta este argumento).

**rho:** Potencia de la función de peso utilizada en las pruebas.

**cencode:** valor de la variable fstatus que indica si el tiempo de fallo está censurado.

**subset:** un vector lógico que especifica un subconjunto de casos para incluir en el análisis.

**na.action**: una función que especifica la acción que se realizará si algún caso está incompleto.

### **Valor**

Una lista con los componentes que dan la estimación de la subdistribución para cada causa en cada grupo, y un componente “Tests” que da el test estadístico y p-valor para comparar la subdistribución para cada causa entre los grupos (si el número de grupos es > 1). Los componentes que dan las estimaciones tienen nombres que son una combinación del nombre del grupo y del código de la causa. Estos componentes son también listas, con componentes.

**time**: los tiempos donde se calculan los estimadores.

**est**: las funciones de sub-distribución estimadas. Estas son funciones escalonadas.

**var**: la varianza estimada de los estimadores, que son estimaciones asintóticas (Aalen, 1978).

### **Ejemplos**

Dan los siguientes resultados:

```
> set.seed (2)
> ss<-rexp(100)
> gg<-factor(sample(1:3,100,replace=TRUE),1:3,c('a','b','c'))
> cc<-sample(0:2,100,replace=TRUE)
> strt<-sample(1:2,100,replace=TRUE)
> print(xx<-cuminc(ss,cc,gg,strt))
```

Tests:

	stat	pv	df
1	0.1110522	0.9459873	2
2	0.1344095	0.9350037	2

Estimates and Variances:

\$est

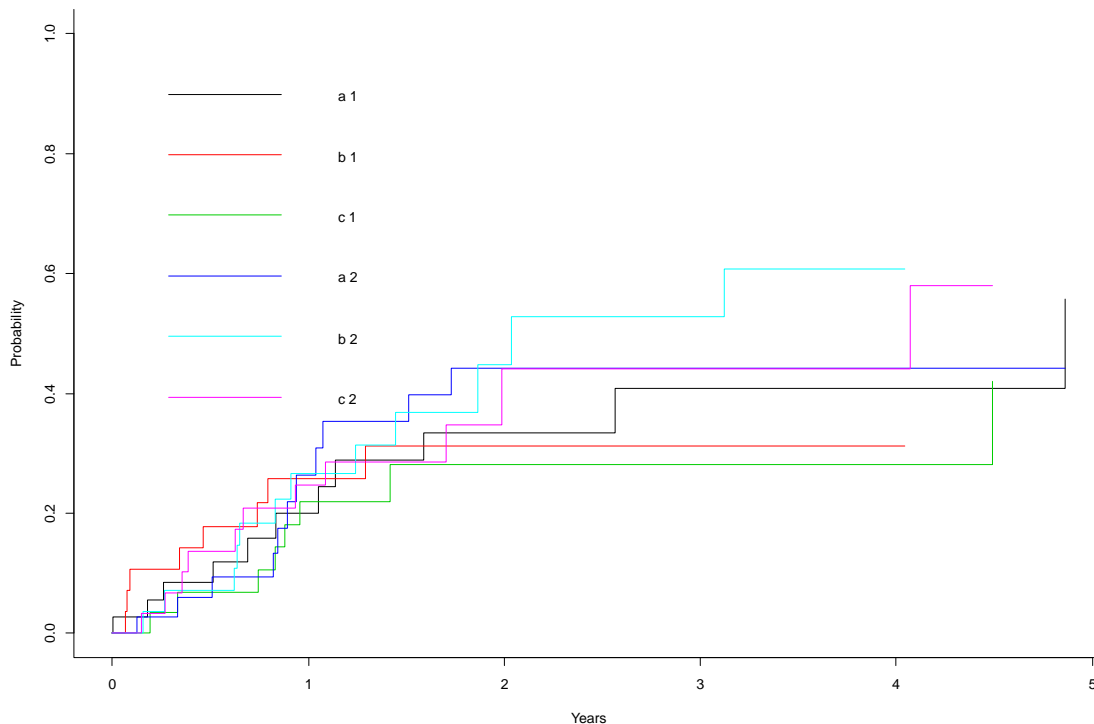
		1	2	3	4
a	1	0.1997374	0.3337599	0.4082169	0.4082169
b	1	0.2582589	0.3116745	0.3116745	0.3116745
c	1	0.2194354	0.2813451	0.2813451	0.2813451
a	2	0.2641725	0.4428692	0.4428692	0.4428692
b	2	0.2669364	0.4479557	0.5280790	0.6082023
c	2	0.2471892	0.4400615	0.4400615	0.4400615

\$var

		1	2	3	4
a	1	0.005842447	0.009576915	0.013764577	0.013764577

b	1	0.007508487	0.009432554	0.009432554	0.009432554
c	1	0.006644662	0.009550507	0.009550507	0.009550507
a	2	0.008127032	0.011168223	0.011168223	0.011168223
b	2	0.008030980	0.014868553	0.016688860	0.016481796
c	2	0.006941955	0.016574813	0.016574813	0.016574813

> `plot(xx,lty=1,color=1:6)`



> `#see also test.R,test.out`

### 5.1.3 La función *plot.cuminc*

Es una función **se describe como** un método de dibujo para la función `cuminc`.

#### Sintaxis

```
## S3 method for class 'cuminc'
plot(x, main=" ", curvlab, ylim=c(0, 1), xlim, wh=2,xlab="Years", ylab="Probability",
lty=1:length(x), color=1, lwd=par('lwd'),...)
```

#### Argumentos empleados en la sintaxis:

**x**: una lista, con cada componente que representa una curva en el gráfico. Cada componente de `x` en sí es una lista cuyo primer componente proporciona los valores de `x` y el segundo componente pintará los valores de `y`. Aunque está escrito para las curvas

de incidencia acumulada, en principio, puede ser utilizado para cualquier conjunto de líneas.

**main:** el título principal del gráfico.

**curvlab:** Etiquetas de curva para el gráfico. El valor por defecto es (x), o si es un valor perdido, 1: nc, donde nc es el número de curvas en x.

**ylim:** rango del eje y.

**xlim:** rango del eje x (por defecto es 0 para el mayor tiempo en cualquiera de las curvas).

**wh:** En un vector de longitud 2, las coordenadas tienen que estar en la parte superior derecha de la leyenda, de lo contrario, la leyenda se coloca en la esquina superior derecha del gráfico).

**xlab:** etiqueta del eje X.

**ylab:** etiqueta del eje Y.

**lty:** vector de tipos de línea. Por defecto 1: nc (nc es el número de curvas en x). Para obtener un color pantalla, lty = 1, color = 1: nc, puede ser más apropiado. Si la longitud (lty) <nc, entonces se utiliza para todos lty [1].

**color:** vector de colores. Si la longitud (color) <nc, entonces el color [1] se utiliza para todos.

**lwd:** vector del ancho de línea. Si la longitud (lwd) <nc, entonces lwd [1] se utiliza para todos.

... argumentos adicionales que han pasado a la llamada inicial de la función plot.

### **Valor**

No se devuelve ningún valor.

### **5.1.4 La función *plot.predict.crr***

Esta función se describe como un método gráfico para la función predict.crr.

### **Sintaxis**

```
## S3 method for class 'predict.crr' plot(x, lty=1:(ncol(x)-1), color=1, ylim=c(0, max(x[, -1])), xmin=0, xmax=max(x[, 1]), ...)
```

### **Argumentos empleados en la sintaxis:**

**x:** Salida de predict.crr.

**lty:** vector de tipos de líneas. Si la longitud es <# curvas, entonces lty[1] se usa para todos.

**color:** vector de los colores de línea. Si la longitud es <# curvas, entonces color[1] se utiliza para todos.

**ylim:** rango del eje y (vector de longitud dos).

**xmin:** límite inferior del eje x (a menudo 0, por defecto).

**xmax:** límite superior del eje x



... Otros argumentos para dibujar.

### **Efectos:**

Esta función dibuja las funciones de sub-distribución estimadas por `predict.crr`, por defecto utiliza un tipo de línea diferente para cada curva.

### **5.1.5 La función *predict.crr***

Esta función es un método de predicción para `crr`. **Estima las funciones de sub-distribución de la salida de `crr`.**

### **Sintaxis:**

```
## S3 method for class 'crr' predict(object, cov1, cov2, ...)
```

### **Argumentos empleados en la sintaxis:**

**object:** salida de `crr`.

**cov1, cov2:** cada fila de `cov1` y `cov2` es un conjunto de valores de covarianza donde la subdistribución debe ser estimada. Las columnas de `cov1` y `cov2` deben estar en el mismo orden como en la llamada original a `crr`. Cada uno debe darse si está presente en la llamada original a `crr`.

### **Detalles**

Calcula  $1 - \exp(-B(t))$ , donde  $B(t)$  es el riesgo de la sub-distribución acumulativa estimada obtenido para los valores de covariable especificados, y a partir de la estimación de tipo Breslow del riesgo base y los coeficientes de regresión estimados.

### **Valor**

Devuelve una matriz con los tiempos de fallo distintos en la primera columna, y las otras columnas dan la función de sub-distribución estimada correspondiente a las combinaciones de covariable en la filas de `cov1` y `cov2`, en cada tiempo de fallo (el valor de la estimación que aumenta en el tiempo de fallo).

### **5.1.6 La función *print.crr***

Esta función es un método de impresión para objetos `crr`.

### **Sintaxis**

```
## S3 method for class 'crr'  
print(x, ...)
```

### **Argumentos empleados en la sintaxis:**

**x:** objeto `crr` (salida de `crr()`)

...: argumentos adicionales para imprimir ()

### **Detalles**

Imprime el estado de convergencia, los coeficientes estimados, los errores estándar estimados, y los p-valores para el test de que los coeficientes individuales sean iguales a 0.

### **5.1.7 La función *print.cuminc***

Esta función es un método de impresión para objetos de la clase *cuminc* (salida de *cuminc*()).

### **Sintaxis**

```
## S3 method for class 'cuminc'  
print(x, ntp=4, maxtime, ...)
```

### **Argumentos empleados en la sintaxis:**

**x**: un objeto de la clase *cuminc*.

**ntp**: número de puntos temporales donde se imprimen las estimaciones.

**maxtime**: el punto de tiempo máximo en el que se imprimen los valores. El valor predeterminado es el máximo tiempo en las curvas en *x*.

... Argumentos adicionales para imprimir ().

### **Detalles**

Imprime los tests estadísticos y p-valores (si están presentes en *x*), y para cada curva de incidencia acumulada estimada, imprime su valor y la varianza estimada en un vector de tiempos. Los tiempos se eligen entre 0 y *maxtime* utilizando la función de *pretty* ().

### **5.1.8 La función *summary.crr***

Esta función genera e imprime resúmenes de la salida de *crr*.

### **Sintaxis**

```
## S3 method for class 'crr'  
summary(object, conf.int = 0.95, digits = max(options())$digits - 5, 2), ...)
```

```
## S3 method for class 'summary.crr'  
print(x, digits = max(options())$digits - 4, 3), ...)
```

### **Argumentos empleados en la sintaxis:**

**object**: Un objeto de la clase *crr* (salida de la función *crr*).

**conf.int**: el nivel para un intervalo de confianza bilateral para los coeficientes. El valor predeterminado es 0,95.

**digits**: En `summary.crr`, `digits` determina el número de dígitos significativos en los p-valores. En `print.summary.crr`, `digits` establece los valores de la opción `digits` para imprimir la salida.

...: Se incluye por compatibilidad con las funciones genéricas. No se utiliza actualmente.

**x**: Un objeto de clase `summary.crr` (salida del método `summary` para `crr`).

### **Valor**

`summary.crr` devuelve una lista de clase `summary.crr`, que contiene componentes.

**call**: La llamada a la función `crr`.

**converged**: TRUE si el algoritmo iterativo converge.

**n**: El número de observaciones usadas en el ajuste del modelo.

**n.missing**: El número de observaciones eliminadas por `crr` de los datos de entrada debido a la falta de valores.

**loglik**: El valor negativo de la función objetivo.

**coef**: Una matriz que devuelve los coeficientes estimados, los cocientes de riesgos, errores estándar, z-scores, y p-valores.

**conf.int**: Una matriz donde figura la relación de riesgos estimados, las relaciones inversas de riesgo y los límites de confianza inferiores y superiores de las relaciones de riesgo.

**logtest**: El doble de la diferencia en los valores de la log pseudo-verosimilitud.

### **5.1.9 La función *timepoints***

Esta función encuentra valores en puntos de tiempo específicos de curvas especificadas así como en todos los extremos de funciones escalonadas.

### **Sintaxis**

`timepoints(w, times)`

### **Argumentos empleados en la sintaxis:**

**w**: una lista que contiene las estimaciones, con puntos de todos los extremos de la función escalonada. (Por lo general creada por `cuminc`.) Cada componente de la lista contiene la estimación para un grupo diferente. Cada componente tiene los componentes de los tiempos dados, la estimación de la función y varianzas (véase `cuminc`).

**times**: vector de tiempos donde son necesarias las estimaciones.

### **Valor**

Una lista con los componentes

**\$est**: una matriz de estimaciones de las subdistribuciones con una fila para cada componente en  $w$  y una columna para cada tiempo.

**\$var**: una matriz que devuelve las varianzas correspondientes.

## 6. Referencias

- Aalen, O. O. (1993). Further results on the non-parametric linear regression model in survival analysis. *Stat Med*, **12**(17), 1569-1588.
- Cox, D. y Oakes, D. (1984). *Analysis of survival data*. London: Chapman-Hall.
- Crowder, M.J. (2001). *Classical competing risks*. Chapman & Hall/CRC: Boca raton.
- Fine, J. y Gray, R. (1999). A proportional hazards model for the subdistribution of a competing risk. *Journal of the American Statistical Association*, **94**(446), 496-509.
- Fine, J. P. (2001). Regression modeling of competing crude failure probabilities. *Biostatistics*, **2**(1), 85-97.
- Gooley, T. A., Leisenring, W., Crowley, J. y Storer B.E. (1999). Estimation of failure probabilities in the presence of competing risks: New presentations of old estimators. *Statistics in Medicine*, **18**: 695-706.
- Kalbfleisch J.D. y Prentice R.L. (2002). *The Statistical Analysis of Failure Time Data*. Wiley Series in Probability and Statistics. John Wiley & Sons, Inc.
- Lawless, J. (2003). *Statistical Models and Methods for Lifetime Data*. Wiley Series in Probability and Statistics. John Wiley & Sons, Inc.
- Lindqvist, B.H. y Langseth, H. (2006). Competing risks for repairable systems: A data study. *Journal of Statistical Planning and Inference* **136**(5), 1687-1700.
- Matadamas Segura, M. (2010). Inferencia para modelos de supervivencia de un solo evento y extensiones para modelos de riesgos competitivos.
- Martinussen, T. y Scheike, T. H. (2006). *Dynamic regression models for survival data*. Statistics for Biology and Health. Springer, New York.
- Peterson, A. V. (1976). Bounds for a joint distribution function with fixed sub-distributions functions: Applications to competing risks. *Proceedings of the National Academy of Sciences of the USA*, **73**, 11-13.
- Porta, N., Gómez, G., Calle, M.L. y Malats, N. (2007). Competing risks methods. Technical Report DR 2007/14
- Prentice, R.L., Kalbfleisch, J.D., Peterson, A.V., Flournoy, N., Farewell, V.T. y Breslow, N.E. (1978). The analysis of failure times in the presence of competing risks, *Biometrics*, **34**: 541-554.
- Tsiatis, A. (1975). A nonidentifiability aspect of the problem of competing risks. *Proceedings of National Academy of Sciences USA*, **72**: 20-22.