



Universidad Nacional Autónoma de México

FACULTAD DE CIENCIAS

# RIESGOS COMPETITIVOS

PROYECTO ANÁLISIS DE SUPERVIVENCIA

*Modelos de Supervivencia y Series de Tiempo*

Autores:

Cuéllar Chávez Eduardo de Jesús

García Tapia Jesús Eduardo

Miranda Meraz Areli Gissell

Ramirez Maciel José Antonio

Saldaña Morales Ricardo

Enero 2022

# Contents

<b>1</b>	<b>Introducción</b>	<b>2</b>
1.1	Notas históricas . . . . .	2
1.2	Definición de riesgos competitivos . . . . .	2
1.3	Uso del método kaplan-meier en la presencia de riesgos competitivos . . . . .	2
<b>2</b>	<b>Definiciones de riesgos competitivos</b>	<b>3</b>
2.1	Reconocimientos de riesgos competientes . . . . .	3
2.1.1	Enfoques prácticos . . . . .	3
2.1.2	Criterios de valoración comunes en la investigación médica . . . . .	3
2.2	Dos definiciones matemáticas . . . . .	4
2.2.1	Riesgos competitivos como variable aleatoria bivariada . . . . .	4
2.2.2	Riesgos competitivos como tiempos de falla latente . . . . .	4
2.2.3	Riesgos competitivos como variable aleatoria bivariada . . . . .	4
2.2.4	Riesgos competitivos como tiempos de falla latente . . . . .	5
2.3	El estimador de Nelson-Aalen . . . . .	5
<b>3</b>	<b>Modelado multiestado de riesgos competitivos</b>	<b>7</b>
3.1	El modelo multiestado de riesgos competitivos . . . . .	7
<b>4</b>	<b>Modelos de riesgos proporcionales de causa específica</b>	<b>9</b>
4.1	Estimación . . . . .	9
4.1.1	Estimación de los coeficientes de regresión . . . . .	9
<b>5</b>	<b>Modelo de riesgos de subdistribución proporcional</b>	<b>12</b>
5.1	El proceso de subdistribución . . . . .	12
5.2	Estimación . . . . .	14
<b>6</b>	<b>Modelado en presencia de riesgos competitivos</b>	<b>19</b>
6.1	Modelando el peligro de la función de incidencia acumulada . . . . .	19
6.1.1	Detalles teóricos . . . . .	19
6.1.2	Estimación basada en modelos del CIF . . . . .	20
<b>7</b>	<b>Análisis de riesgos de subdistribución proporcional de todas las funciones de incidencia acumulada</b>	<b>20</b>
<b>8</b>	<b>Bibliografía</b>	<b>21</b>

## 1 Introducción

### 1.1 Notas históricas

A mediados del siglo XVII John Graunt, un comerciante de mercería de oficio, empezó a recopilar y tabular la información incluida en las Cartas de Mortalidad publicadas semanalmente en ese tiempo en Londres. Este trabajo, probablemente realizado en colaboración con Sir William Petty, apareció en 1662 y contenía un resumen demográfico de las causas de muerte en Inglaterra y Gales. Como resultado, John Graunt fue elegido miembro de la Royal Society con recomendación del rey Carlos II, hazaña nada despreciable teniendo en cuenta la estructura jerárquica de la sociedad existente en ese momento. En 1693, el famoso astrónomo Edmund Halley desarrolló el concepto de tablas de vida en un formato no muy diferente al que se usa hoy en día en análisis supervivencia. Sus datos se basaron en el registro de nacimientos y defunciones para la ciudad de Breslau (ahora Wrocław, en el suroeste de Polonia).

En 1760 Daniel Bernoulli aplicó el método de Halley para demostrar la Ventajas de la inoculación de la viruela. Calculó el aumento de la función de superviviente de Halley si se eliminara la viruela como causa de la muerte. De esta manera, Bernoulli fundó la teoría de riesgos competitivos. Se puede encontrar un resumen de su trabajo sobre riesgos competitivos en David y Moeschberger (1978). A partir del siglo XVIII, se desarrollaron las matemáticas, se cambiaron las notaciones y se ramificó la estadística como una ciencia por derecho propio. Los grandes descubrimientos del siglo XIX en la física y la biología condujo a una vista determinista del universo. Sin embargo, a medida que crecieron las herramientas de medición cada vez más se hizo evidente que algunos inexplicables, factores aleatorios estaban en el trabajo. En el siglo XX, debido en parte a la necesidad de explicar la variación aleatoria y en parte al extraordinario desarrollo de las capacidades informáticas, la teoría de las estadísticas experimentó un progreso significativo. Basado en el trabajo de Bernoulli, la teoría de los riesgos competitivos, denominada en ciencias actuariales como múltiples decrementos, también se ha desarrollado.

### 1.2 Definición de riesgos competitivos

El análisis de datos de "tiempo para un evento", los cuales consisten de pares de observaciones para cada individuo (el primero es la longitud de tiempo durante el cual ningún evento fue observado; el segundo es un indicador de si el final de dicho tiempo corresponde a un evento o solo al final de la observación), se suele referir como análisis de supervivencia, porque la teoría se desarrolló utilizando la muerte como evento de interés. También se le llama análisis de datos incompletos. El término incompleto se refiere a registros donde un evento no ha sido observado, pero está obligado a tener lugar si se sigue por el tiempo suficiente. Este tipo de observación se denomina censurada. El evento no necesita ser la muerte, puede ser cualquier evento que ocurre con el tiempo, como la recaída de la enfermedad, la reincidencia o el alta hospitalaria. Las técnicas de supervivencia están bien desarrolladas e implementadas en los principales programas estadísticos. Sin embargo, hay algunas situaciones en las que puede no ser adecuado aplicar las normas habituales de métodos de supervivencia para el análisis de tiempo hasta el evento. Una de esas situaciones es donde los riesgos competitivos están presentes. La situación de los riesgos competitivos puede definirse de varias maneras diferentes. En general, una situación de riesgos competitivos surge cuando un individuo puede experimentar más de un tipo de evento y la ocurrencia de un tipo de evento impide la ocurrencia de otros tipos de eventos.

### 1.3 Uso del método kaplan-meier en la presencia de riesgos competitivos

En presencia de riesgos competitivos, los métodos habituales de supervivencia deben aplicarse con precaución y hay que ser conscientes de las consecuencias de su uso. El método Kaplan-Meier es el

más común, así como la técnica más controvertida en el marco de riesgos competitivos. Es un método para estimar probabilidades de supervivencia (Kaplan y Meier, 1958) en diferentes momentos, como hemos visto en clase. Es relativamente fácil de aplicar e interpretar y se puede representar visualmente. Su amplia disponibilidad en el software estadístico hace que su uso sea atractivo. Cuando existen riesgos competitivos, las estimaciones Kaplan-Meier (indicadas por KM) no pueden interpretarse como probabilidades. Su complemento  $(1 - \text{KM})$  se puede interpretar como la probabilidad de un evento de interés en un mundo ideal donde los otros tipos de eventos no existen. Sin embargo, este concepto no es útil en la práctica. Kalbfleisch y Prentice (1980) sugirieron un enfoque que representaba los riesgos competitivos. Este método está etiquetado como la función de incidencia acumulada (CIF). Utilizando esta técnica, la probabilidad de que suceda cualquier evento se divide en las probabilidades para cada tipo de evento.

## 2 Definiciones de riesgos competitivos

### 2.1 Reconocimientos de riesgos competientes

#### 2.1.1 Enfoques prácticos

La situación en la que están presentes los riesgos competitivos ha sido expresada de diferentes maneras por diferentes autores. Dado que muchos de estos las descripciones arrojan luz sobre diferentes aspectos de la situación de los riesgos competitivos, varios se presentan aquí para ilustrar mejor el punto. Kalbfleisch y Prentice (2002) describieron los riesgos competitivos como la situación en la que un individuo puede experimentar más de un tipo de evento. Gelman (1990) y Caplan (1994) lo explicaban como el fracaso en lograr la independencia entre el tiempo a un evento y el mecanismo de censura. Y finalmente, Gooley (1999) definió el concepto de riesgos competitivos como la situación donde un tipo de evento “impide la ocurrencia de otro evento bajo investigación o altera fundamentalmente la probabilidad de ocurrencia de este otro evento”. Esto refleja fielmente la situación con la que se enfrenta el investigador clínico o el estadístico aplicado: la posibilidad de muchos tipos de fallas que compiten entre sí a ser observado. Suponga que se realiza un estudio de un nuevo tratamiento para la enfermedad cardíaca. llevado a cabo en una población de mayor edad. Dada la avanzada edad de la grupo de estudio es posible, y tal vez probable, que un paciente pueda morir por una causa distinta a la insuficiencia cardíaca. Una muerte debido a algún otro causa constituye un tipo de evento de riesgo competitivo, porque impide la observación del evento de interés – muerte por enfermedad cardíaca.

#### 2.1.2 Criterios de valoración comunes en la investigación médica

Esta sección comienza con una ilustración de puntos finales que son comúnmente encontrados en estudios de cáncer. Explicaciones más generales y ejemplos de otros campos de la investigación médica también son introducidos. En la investigación del cáncer, los pacientes se someten a uno o más de los tres principales tipos de tratamiento: cirugía, quimioterapia y radiación terapia. Supongamos que el tratamiento parece exitoso y todos se eliminó la evidencia de la enfermedad. Un punto final común de interés en estudios de cáncer es el tiempo hasta el regreso de la enfermedad (recaída) después el aparente éxito inicial del tratamiento. La recaída puede ser en el sitio de la enfermedad inicial, en cuyo caso el punto final se llama local recaída, o en un sitio diferente, llamado recaída a distancia o metástasis. La quimioterapia, como tratamiento sistémico, afecta a todo el organismo mientras que tanto la cirugía como la radioterapia son tratamientos dirigidos hacia el sitio específico de la enfermedad. Por lo tanto, en los estudios de radiación o cirugía, el investigador puede estar más interesado en el tiempo a la recaída local que en el tiempo a la metástasis o muerte. En este caso, es deseable identificar características que se asocian con recidiva local. Sin embargo, un paciente puede desarrollar una enfermedad a distancia y mueren antes de que se observe una recaída local. En este caso el la observación de enfermedades distantes dificulta la observación de enfermedades locales.

enfermedad. Además, la aparición de enfermedad local después de la enfermedad puede no ser de mucho interés ya que el tratamiento de la distancia la enfermedad puede alterar las posibilidades de recurrencia de la enfermedad local. En términos más generales, el término "recaída" se refiere al retorno o recurrencia de cualquier afección o enfermedad potencialmente crónica después de un mejora inicial. Por ejemplo, en los estudios sobre el abandono del hábito de fumar, La recaída se refiere a la reanudación de la conducta de fumar anterior. Los estudios psiquiátricos de pacientes con trastorno bipolar podrían considerar un episodio repetido de manía para ser una recaída, mientras que un clínico ensayo en el que participaron pacientes con bronquitis crónica podría definir recaída es la reaparición de los síntomas. Pulmonar tuberculosis podría reaparecer dentro del pulmón (recaída local) o en sitios extrapulmonares (otras recaídas). Del mismo modo, una recaída de el virus del herpes simple podría experimentarse como una lesión en la piel (recaída local) o dentro del sistema nervioso central (otra recaída). Por tanto, el concepto de recaída (local, a distancia u otra) es un punto final significativo en otras áreas médicas, además del cáncer, donde la enfermedad o condición bajo estudio puede reaparecer en diferentes sitios dentro del cuerpo sin ser exhaustivo.

## 2.2 Dos definiciones matemáticas

Hay dos enfoques matemáticos para hacer frente a la competencia riesgos, y ambos han recibido atención en la literatura. En esta sección las dos definiciones matemáticas de riesgos competitivos se introducen, así como los conceptos principales: subdistribución, subdensidad, subpeligro, peligro de causa específica y el peligro de la subdistribución.

### 2.2.1 Riesgos competitivos como variable aleatoria bivariada

En ausencia de riesgos competitivos, los datos de supervivencia suelen ser presentados como una variable aleatoria bivariada o par  $(TC)$ . La variable de censura,  $C$ , es 1 si se observó el evento de interés, y es 0 si la observación fue censurada. Cuando  $C = 1$  el primer miembro del par,  $T$ , es el momento en que ocurrió el evento y cuando  $C = 0$ ,  $T$  es el momento en que se censuró la observación. Esta definición puede extenderse a la situación de riesgos competitivos donde  $p \geq 2$  tipos de fallas o eventos son posibles. los datos son nuevamente representado como un par  $TC$ , y el indicador de censura  $C$  volverá a definirse como 0 si se censura la observación. Sin embargo, en el caso de que la observación no sea censurada,  $C$  asumirá el valor  $i$ , donde  $i$  es el tipo de la primera falla/evento observado ( $i = 1, 2, \dots, p$ ). Si  $C = i$ , entonces  $T$  es el momento en que el caso tipo  $i$  ocurrió; de lo contrario, es el momento de la censura (Kalbfleischy Prentice, 1980; Crowder, 2001).

### 2.2.2 Riesgos competitivos como tiempos de falla latente

La segunda definición matemática del resultado de riesgos competitivos se escribe en términos de tiempos de eventos latentes o no observados  $T_1, T_2, \dots, T_p$ , uno para cada uno de los tipos de eventos  $p$ , en la competencia situación de riesgos sólo se observa el primer evento, por lo que la variable de tiempo,  $T$ , se define como la primera de estos hipotéticos tiempos no observados. Matemáticamente, entonces,  $T = \min\{T_1, T_2, \dots, T_p\}$ . La variable de censura,  $C$ , se define como  $C = 0$  si la observación es censurada, y  $C = i$   $i = 1, 2, \dots, p$ , de lo contrario. En el ensayo de tamoxifeno,  $T_1$  podría ser el momento de la recaída local,  $T_2$  tiempo hasta la recaída a distancia,  $T_3$  tiempo hasta otra malignidad y  $T_4$  hora hasta la muerte sin ninguno de estos eventos. El tiempo observado es  $T = \min\{T_1, T_2, T_3, T_4\}$ , y  $C$  es 1 cuando el primer evento es una recaída local, 2 si el primer evento es una recaída a distancia, 3 si se observó malignidad, 4 para muerte sin eventos y cero en caso contrario.

### 2.2.3 Riesgos competitivos como variable aleatoria bivariada

El CIF, o subdistribución, para un evento de tipo  $i$  ( $i = 1, 2, \dots, p$ ) se define como la probabilidad conjunta

$$F_i(t) = P(T \leq t, C = i)$$

En otras palabras, el CIF es la probabilidad de que un evento de tipo  $i$  ocurre en o antes del tiempo  $t$ . La función de distribución general es la probabilidad de que un evento de cualquier tipo ocurre en o antes del tiempo  $t$ .

#### 2.2.4 Riesgos competitivos como tiempos de falla latente

Dado que el enfoque de variables latentes se basa en una colección de  $p$  tiempos latentes, comencemos definiendo el sobreviviente conjunto multivariante función

$$S(t_1, t_2, \dots, t_p) = P(T_1 > t_1, T_2 > t_2, \dots, T_p > t_p).$$

Entonces la subdensidad para el tipo de evento  $i$  es

$$f_i(t) = \left( -\frac{(t_1, t_2, \dots, t_p)}{i} \right)_{t_1=t_2=\dots=t_p=t}$$

La función de supervivencia marginal para el tipo de evento  $i$  es

$$S_i(t) = S(t_1 = 0, t_2 = 0, \dots, t_i = t, \dots, t_p = 0)$$

y la función de subdistribución se define como sigue:

$$F_j(t) = \int_0^t f_j(s) ds.$$

### 2.3 El estimador de Nelson-Aalen

Con las curvas de Kaplan-Meier, aproximamos la función de supervivencia, expresando la probabilidad de que el evento de interés (por ejemplo, el evento de muerte) no ocurra en un cierto tiempo  $t$ . Un enfoque alternativo para visualizar la información agregada de un conjunto de datos centrado en la supervivencia implica el uso de la función de riesgo, que se puede interpretar como la probabilidad de que el sujeto experimente el evento de interés dentro de un pequeño intervalo de tiempo, asumiendo que el sujeto ha sobrevivido hasta el comienzo de dicho intervalo. Desafortunadamente, no podemos transformar la estimación de Kaplan-Meier de la función de supervivencia en la función de riesgo. Sin embargo, podemos usar otro estimador no paramétrico de la función de riesgo acumulativo: el estimador de Nelson-Aalen. En resumen, se utiliza para estimar el número acumulado de eventos esperados dentro de un período de tiempo determinado. Y la razón por la que es acumulativo es que la suma de las estimaciones es mucho más estable que las estimaciones puntuales. El estimador de Nelson-Aalen se puede calcular como:

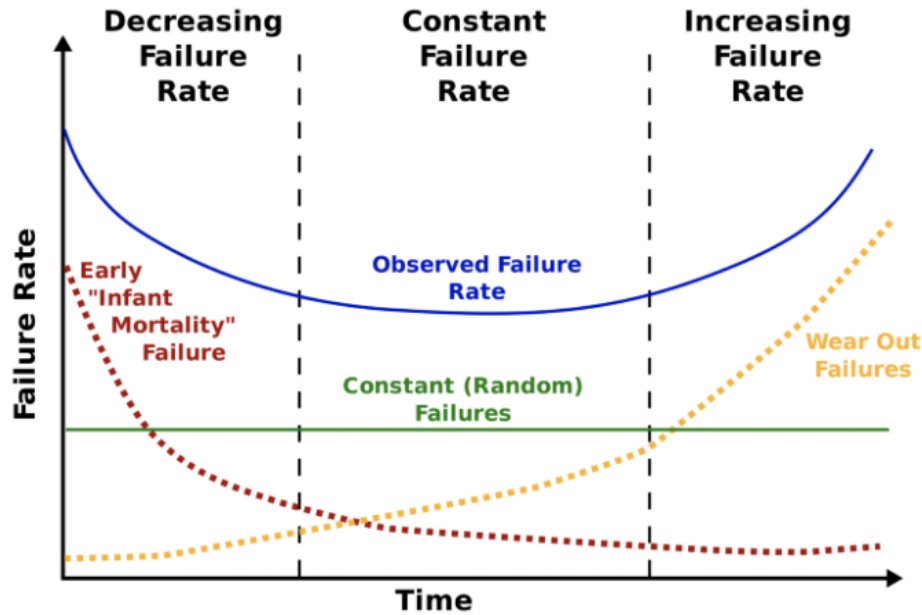
$$\hat{H}(t) = \sum_{t_i \leq t} \frac{d_i}{n_i}$$

donde  $d_i$  representa el número de eventos de interés en el tiempo  $t$ , mientras que  $n_i$  es el número de observaciones en riesgo. Todos estos términos son naturalmente similares a los de la fórmula del estimador de Kaplan-Meier. El estimador de Nelson-Aalen, o más generalmente la visualización de la función de riesgo a lo largo del tiempo, no es un enfoque muy popular para el análisis de supervivencia. Esto se debe a que, en comparación con la función de supervivencia, la explicación de las curvas no es tan simple e intuitiva. Sin embargo, la función de riesgo es de gran importancia para los enfoques más avanzados del análisis de supervivencia, por ejemplo, la regresión de Cox. Por eso es importante entender el concepto y trataré de proporcionar algunas ideas al respecto. Podemos decir que la función de riesgo acumulada:

- Mide la cantidad total de riesgo que se ha acumulado hasta un cierto punto de tiempo  $t$ .
- Proporciona el número de veces que esperaríamos matemáticamente que ocurra el evento de interés durante un cierto período si solo los eventos fueran repetibles. Esto puede ser un poco confuso, así que para hacer la declaración un poco más simple (aunque no tan realista), puede pensar en la función de riesgo acumulativo como el número esperado de muertes de un individuo hasta el tiempo  $t$ , si el individuo pudiera resucitar. después de cada muerte sin reiniciar el tiempo. Como dije, no es tan realista, pero esto también podría aplicarse a fallas de máquinas, etc.

El último concepto que puede ser útil para obtener cierta intuición sobre la función de riesgo acumulativo es la curva de la bañera, o más bien sus componentes. La curva representa los ciclos de vida de muchos productos electrónicos de consumo. La tasa de riesgo de la curva de la bañera se crea combinando:

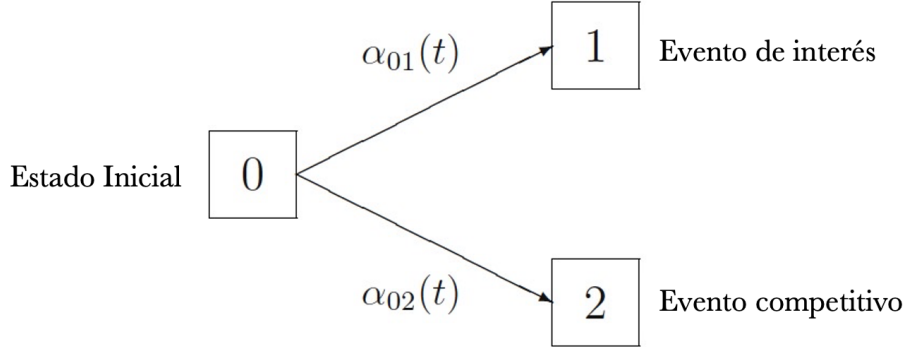
- La tasa de fallas tempranas de "mortalidad infantil" cuando el producto se introduce por primera vez,
- La tasa de fallas aleatorias con una tasa de fallas constante durante la vida útil diseñada del producto,
- La tasa de fallas por "desgaste" a medida que el producto excede su vida útil esperada.



Si bien la imagen de arriba representa la tasa de riesgo (¡no la acumulativa!), la forma de la curva del estimador de Nelson-Aalen nos da una idea de cómo cambia la tasa de riesgo con el tiempo.

### 3 Modelado multiestado de riesgos competitivos

#### 3.1 El modelo multiestado de riesgos competitivos



Modelo multiestado de riesgos competitivos con peligros de causa específica  $\alpha_{0j}(t), j = 1, 2$ .

Inicialmente cada individuo está en el estado 0 en el origen del tiempo. El individuo permanece en este estado hasta que ocurra algún primer evento. Por lo general, hay un evento de interés, modelado por transiciones al estado 1, y todos los demás primeros tipos de eventos son incluidos en el estado 2 de eventos competitivos.

Por ejemplo en el caso de los datos de estancia hospitalaria, la mortalidad hospitalaria modelada por el estado 1 puede ser de especial interés. El estado 2 de eventos competitivos tiene entonces la interpretación "descarga viva". Sin embargo, este estado de eventos competitivos no distingue más si el paciente regresa a casa, es readmitido en un hospital diferente o ingresa a un asilo de ancianos. El proceso de riesgos competitivos  $(X_t)_{t \geq 0}$  de la imagen anterior denota el estado en el que se encuentra un individuo para cada punto en el tiempo  $X_t \in \{0, 1, 2\}$ . Todo individuo comienza en el estado inicial 0 en el tiempo origen 0.

$$P(X_0 = 0) = 1$$

Un individuo permanece en el estado 0, es decir,  $X_t = 0$  mientras no haya ocurrido ni el evento competitivo 1 ni el evento competitivo 2. El individuo pasa al estado 1 si ocurre el evento de interés. Asimismo, el individuo pasa al estado 2 si el otro evento competitivo ocurre primero. El proceso de riesgos competitivos sale del estado inicial 0 en el tiempo  $T$

$$T := \inf\{t > 0 \mid X_t \neq 0\}.$$

$T$  es denominada tiempo de supervivencia o tiempo de falla; se puede considerar como un tiempo de espera en el estado inicial 0. En muchas aplicaciones, cada individuo dejará el estado inicial en algún momento, es decir,  $P(T \in (0, \infty)) = 1$ . Por lo tanto, se piensa a  $T$  como un tiempo de evento de valor real no negativo.

Recordemos que los procesos multiestado son continuos por la derecha. El proceso de riesgos competitivos está en el estado 1 o en el estado 2 en el momento  $T$ . El tipo del primer evento, llamado causa de falla, es

$$X_T \in \{1, 2\},$$

es el estado en el que entra el proceso en el momento  $T$ .



La observación del proceso de riesgos competitivos  $(X_t)_{t \geq 0}$  en general estará sujeta a censura por la derecha y/o truncamiento por la izquierda, si la observación del proceso está sujeta a un tiempo de censura por la derecha  $C$  únicamente, los datos observados serán  $(T \wedge C, 1(T \leq C) \cdot X_T)$  hay que tener en cuenta el indicador de estado,

$$1(T \leq C) \cdot X_T \in \{0, 2, 1\},$$

es igual a 0, si la observación fue censada. Encontramos que desde la perspectiva de la codificación de software puede ser ventajoso codificar eventos de censura por un valor diferente de 0; a menudo usamos 'cens' como código de censura. En caso de un evento de censura, la observación se detuvo mientras el individuo aún estaba en el estado inicial 0. De lo contrario, se observaron el tiempo de falla y la causa de falla, y el indicador de estado es igual a  $X_t$ . Si además tenemos un tiempo de truncamiento por la izquierda  $L$ , los datos observados serán  $([L, T \wedge C], 1(T \leq C) \cdot X_T)$ .

Las cantidades clave en los riesgos competitivos son los peligros de causa específica  $\alpha_{0j}(t), j = 1, 2$ ,

$$\alpha_{0j}(t)dt := P(T \in dt, X_T = j \mid T \geq t), j = 1, 2.$$

Donde escribimos a  $dt$  tanto para la longitud del intervalo infinitesimal  $[t, t + dt)$  como para el intervalo mismo. También escribimos  $A_{0j}(t)$  para los peligros acumulativos específicos de la causa

$$A_{0j}(t) := \int_0^t \alpha_{0j}(u)du, j = 1, 2.$$

Los peligros específicos de la causa se pueden considerar como fuerzas momentáneas de transición, moviéndose a lo largo de las flechas en la imagen anteriores. Es crucial para cualquier análisis de riesgos competitivos que ambos peligros específicos de causa determinen completamente el comportamiento estocástico del proceso de riesgos competitivos. Los  $\alpha_{0j}s$  resumen el peligro por todas las causas:

$$\begin{aligned} \alpha_{0\cdot}(t)dt &:= p(T \in dt \mid T \geq t) \\ &= (\alpha_{01}(t) + \alpha_{02}(t))dt. \end{aligned}$$

Este resultado es consecuencia de la usual suma de probabilidades. También tenemos

$$A_{0\cdot}(t) := \int_0^t \alpha_{0\cdot}(u)du = A_{01}(t) + A_{02}(t),$$

para el riesgo acumulativo por todas las causas.

La función de supervivencia del tiempo de espera  $T$  en el estado inicial 0 es:

$$P(T > t) = \exp(-\int_0^t \alpha_{0\cdot}(u)du).$$

La función de supervivencia  $P(T > t)$  a menudo se denota como  $S(t)$ . Es una función de ambos  $\alpha_{0j}S$ , porque  $\alpha_{0\cdot}(t) = \alpha_{01}(t) + \alpha_{02}(t)$ .

A menudo, el interés en los riesgos competitivos se centra en la función de incidencia acumulada, es decir, la proporción esperada de personas que experimentan un determinado evento competitivo en el transcurso del tiempo:

$$P(T \leq t, X_T = j) = \int_0^t P(T > u-) \alpha_{0j}(u)du, j = 1, 2,$$

donde escribimos  $P(T > u-)$  para el valor de la función de supervivencia justo antes a  $u$ . Observamos que la función de incidencia acumulada es una función tanto de  $\alpha_{0j}s$  como de  $P(T > u-)$ .

Hay una interpretación intuitiva del lado derecho de la ecuación anteriores, que vuelve a aparecer cuando consideramos la estimación no paramétrica. Uno integra o, en términos generales, suma

sobre 'probabilidades infinitesimales' de hacer la transición  $0 \rightarrow j$  exactamente en el momento  $u$ .  $P(T > u)$  es la probabilidad de estar todavía en el estado inicial 0 antes de  $u$ ; esta probabilidad se multiplica con la 'probabilidad de transición infinitesimal'  $\alpha_{oj}(u)du$  para hacer que la transición  $0 \rightarrow j$  en el tiempo  $u$  esté condicionada a que aún se encuentre en el estado inicial antes de  $u$ . El lado derecho de la ecuación es la solución de un producto integrado. Más precisamente  $P(T \leq t, X_t = 1)$  es la entrada (1,2) de la solución de

$$\prod_{u \in (0, t]} \left\{ \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + d \begin{pmatrix} -A_{0\cdot}(u) & A_{01}(u) & A_{02}(u) \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \right\};$$

La entrada (1,3) es  $P(T \leq t, X_T = 2)$ , y la función de supervivencia  $P(T > t)$  está en la entrada (1,1). Las entradas en las filas 2 y 3 son como las filas 2 y 3 de la matriz identidad de 3x3, lo que refleja que los riesgos en competencia no modelan eventos después del primer evento. Nos encontramos de nuevo con esta estructura al estimar las funciones de incidencia acumulada.

Finalmente notamos que la función de incidencia acumulada para el estado de evento 1 de interés,  $P(T \leq t, X_T = 1)$ , y para el estado de evento competidor 2,  $P(T \leq t, X_T = 2)$ , suman el función de distribución de todas las causas,

$$P(T \leq t, X_T = 1) + P(T \leq t, X_T = 2) = P(T \leq t)$$

y que las funciones de incidencia acumulada se acercan a  $P(X_T = j)$ ,  $j = 1, 2$  a medida que aumenta el tiempo  $t$ ,

$$\lim_{t \rightarrow \infty} P(T \leq t, X_T = j) = P(X_T = j), j = 1, 2,$$

suponiendo que  $T$  es un tiempo finito.

## 4 Modelos de riesgos proporcionales de causa específica

### 4.1 Estimación

Primero consideramos la estimación de los coeficientes de regresión. A continuación, consideramos la estimación basada en modelos de los riesgos de línea de base específicos de la causa acumulada y, finalmente, se introduce la predicción de las funciones de incidencia acumulada.

#### 4.1.1 Estimación de los coeficientes de regresión

Recordando que tenemos procesos de conteo individuales por causas específicas  $N_{0j,i}(t)$  que cuentan si hemos observado un evento tipo  $j$  para individuos  $i$  durante el intervalo de tiempo  $[0, t]$ . Recuerde también que el individuo en riesgo proceso es  $Y_{0,i}(t)$ . Los respectivos procesos agregados sobre todos los individuos  $i, i = 1, \dots, n$  se han definido como  $N_{0,j}(t)$  y  $Y_0(t)$ , respectivamente. Finalmente,  $N_0(t) = N_{01}(t) + N_{02}(t)$  es el número de transiciones observadas fuera del estado inicial 0 durante el intervalo de tiempo  $[0, t]$ . Definimos el conjunto de riesgos ponderados:

$$S_{oj}^{(0)}(\beta, t) := \sum_{i=1}^n \exp(\beta \cdot Z_{0j,i}) \cdot Y_{0j,i}(t) \quad (1)$$

Explicamos el significado de este conjunto de riesgos ponderados en un momento. Tenga en cuenta que la notación  $S_{0j}^{(0)}(\beta, t)$  no debe confundirse con una función de supervivencia  $P(T > t)$ , que a veces se denota como  $S(t)$ . También tenga en cuenta que  $S_{0j}^{(0)}$  realmente depende del tipo de transición  $0 \rightarrow j$  a través de las covariables específicas de la causa  $Z_{0j,i}$ .

La estimación de  $\beta$  se basa en la veromilitud parcial:

$$L(\beta) = \prod_t \prod_{i=1}^n \prod_{j=1}^2 \left( \frac{\exp(\beta \cdot Z_{0j,i})}{S_{0j}^{(0)}(\beta, t)} \right)^{\Delta N_{0j,i}(t)} \quad (2)$$

Donde el primer producto es sobre todos los tiempos  $t$ , donde se observó un evento de tipo 1 o 2 para algún individuo. La interpretación es que  $\exp(\beta \cdot Z_{0j,i})/S_{0j}^{(0)}(\beta, t)$  es la probabilidad de que sea exactamente el individuo  $i$  quien falle del evento tipo  $j$  dado que un evento de tipo  $j$  se observa en el tiempo  $t$ : se deduce de la definición de los riesgos específicos de la causa que la probabilidad de que observemos un evento tipo  $j$  en el intervalo de tiempo muy pequeño  $dt$ , dado que conocemos todos los eventos anteriores de cualquiera tipo y censura de eventos, es

$$P(dN_{0j}(t) = 1 | Past) = \left( \sum_{i=1}^n Y_{0,i}(t) \cdot \alpha_{0j,i}(t, Z_i) \right) dt \quad (3)$$

suponiendo que dos individuos no experimentan un evento al mismo tiempo. En (3), hemos escrito brevemente 'Past' para eventos anteriores de cualquier tipo y eventos de censura. Usando la Definición (1), la cantidad (3) es igual a  $\alpha_{0j,0}(t) S_{0j}^{(0)}(\beta, t) \cdot dt$ . Análogamente, la respectiva probabilidad condicional de que observemos un evento de tipo  $j$  en el intervalo de tiempo muy pequeño  $dt$  precisamente para el individuo  $i$  es  $Y_{0,i}(t) \alpha_{0j,0}(t) \cdot \exp(\beta \cdot Z_{0j,i}) dt$ . De donde:

$$P(\text{individuo } i \text{ con fallo observado por } j \text{ en } t | \Delta N_{0j}(t) = 1, past) = \frac{Y_{0,i}(t) \cdot \exp(\beta \cdot Z_{0j,i})}{S_{0,j}^{(0)}(\beta, t)} \quad (4)$$

Tenga en cuenta que usamos un argumento relacionado sobre la proporción de un riesgo individual y una suma apropiada de riesgos. El factor principal  $Y_{0,i}(t)$  en el numerador es eliminado de la probabilidad parcial  $L(\beta)$  en (2) como un exponente  $\Delta N_{0j,i}(t) = 0$  implica  $Y_{0,i}(t) = 1$ .  $L()$  es una verosimilitud parcial, porque la línea base los riesgos  $\alpha_{0j,0}(t)$  se han cancelado en (4) y en (2). En su artículo, Cox (1972) sugirió que la inferencia estadística para  $\beta$  podría basarse en maximizar  $L(\beta)$ , porque las funciones de riesgo de línea de base se han dejado completamente no especificadas y, por lo tanto, los intervalos de tiempo entre los tiempos de eventos observados no debe contener ninguna información sobre  $\beta$ .

La cuestión de qué tipo de verosimilitud muestra  $L(\beta)$  ha sido considerado en detalle en la literatura posterior al artículo de Cox. De hecho, la teoría martingala muestra que  $L(\beta)$  disfruta de una teoría de muestra grande que comparte muchos puntos de la teoría estándar de máxima verosimilitud, y que podemos usar para (aproximado) inferencia en la práctica. Gill (1984) da una cuenta muy accesible de este problema.

Consideramos propiedades de muestras grandes de  $\hat{\beta}$ . Escribamos  $\beta^{(0)}$  para el verdadero vector de parámetros y entonces  $\sqrt{n}(\hat{\beta} - \beta^{(0)})$  se aproxima a una distribución multinormal con media cero (Andersen y Gill, 1982; Andersen y Borgan, 1985). Para escribir un estimador de la matriz de covarianza, necesitamos para introducir alguna notación adicional. Tenemos entonces:

$$S_{(0j)}^{(1)}(\beta, t) := \sum_{i=1}^n Z_{0j,i} \cdot \exp(\beta \cdot Z_{0j,i}) \cdot Y_{0,i}(t), \quad j=1,2 \quad (5)$$

y

$$E_{0,j}(\beta, t) = \frac{S_{0j}^{(1)}(\beta, t)}{S_{0j}^{(0)}(\beta, t)} \quad (6)$$

Si  $\beta$  es un vector de  $1 \times q$  y  $Z_{0j,i}$  son vectores de  $q \times 1$ , entonces ambos  $S_{0j}^{(1)}(\beta, t)$  y  $E_{0,j}(\beta, t)$  son vectores de  $q \times 1$  podemos ver:

$$S_{(0j)}^{(2)}(\beta, t) := \sum_{i=1}^n Z_{0j,i} Z_{0j,i}^T \cdot \exp(\beta \cdot Z_{0j,i}) \cdot Y_{0,i}(t), \quad j=1,2 \quad (7)$$

como una matrix de  $qxq$  y

$$V_{0,j}(\beta, t) = \frac{S_{0j}^{(2)}(\beta, t)}{S_{0j}^{(0)}(\beta, t)} - E_{0,j}(\beta, t)E_{0,j}(\beta, t)^T, \quad j=1,2 \quad (8)$$

también como una matriz de  $qxq$ . Entonces se puede observar que la matriz de segundo orden derivadas parciales del logaritmo de verosimilitud parcial  $\ln(L(\beta))$  es el negativo de:

$$J_\tau(\beta) = \sum_{j=1}^2 \sum_{T_i \wedge C_i \leq \tau} V_{0j}(\beta, T_i \wedge C_i) \Delta N_{0j}(T_i \wedge C_i) \quad (9)$$

donde  $\tau$  se toma como el mayor tiempo de evento observado. Recuerde también que  $T_i \wedge C_i$  es el mínimo del tiempo de evento  $T_i$  del individuo  $i$  y el tiempo de censura  $C_i$ . Nosotros estimamos la matriz de covarianza de  $(\hat{\beta} - \beta^{(0)})$  por  $(J_\tau(\hat{\beta}))^{-1}$ . El resultado se usa para calcular estimaciones de varianza, intervalos de confianza y p-values para los coeficientes de regresión. También se sigue que la distribución asintótica del estadístico de la prueba de Wald

$$(\hat{\beta} - \beta^{(0)}) J_\tau(\hat{\beta}) (\hat{\beta} - \beta^{(0)})^T$$

para la hipótesis  $\beta = \beta^{(0)}$  es una ji-cuadrada con  $q$  grados de libertad. La estadística de prueba de puntaje es la estadística de prueba de rango logarítmico habitual en el caso de estándar de datos de supervivencia de criterio único, una sola covariable categórica y  $\beta^{(0)} = 0$ . Una discusión de los méritos relativos de estas pruebas se puede encontrar en Therneau y Grambsch (2000, Sección 3.4) y Hosmer et al. (2008, Sección 3.3), entre otros. Hosmer et al. resume que estas pruebas suelen ser bastante similares; Si no están de acuerdo, las conclusiones deben basarse en la prueba de razón de verosimilitud.

Vale la pena señalar que la probabilidad parcial  $L(\beta)$  se divide en dos factores  $L(\beta_{01})$  y  $L(\beta_{02})$ ,

$$L(\beta_{0j}) = \prod_t \prod_{i=1}^n \left( \frac{\exp(\beta_{0j} \cdot Z_i)}{\sum_{l=i}^n \exp(\beta_{0j} \cdot Z_l)} \right)^{\Delta N_{0j;t}(t)} \quad (10)$$

dependiendo del respectivo  $\beta_{0j}$  solo si asumimos dos modelos completamente separados para los riesgos específicos de la causa; es decir,  $\beta = (\beta_{01}, \beta_{02})$  sin efecto común sobre estos riesgos. Obviamente, la estimación de máxima verosimilitud puede obtenerse maximizando por separado  $\hat{\beta}$ , maximizando a su vez  $L(\beta_{01})$  y  $L(\beta_{02})$ . Como consecuencia adicional,  $q = p + p$  y  $pp$  bloques en  $(J_\tau(\beta))^{-1}$  correspondientes a  $\beta_{0j}$  son los que se obtendrían analizando  $L(\beta_{0j})$  solamente.

### Estimador de Breslow de los riesgos acumulativos específicos de la causa y predicción

Sea  $\beta$  el estimador que resulta de maximizar  $L(\beta)$ . Un tipo Nelson-Aalen estimador de los riesgos de línea de base acumulativos específicos de la causa es:

$$\hat{A}_{0j;0}(t) := \sum_{T_i \wedge C_i \leq t} \frac{\Delta N_{0j}(T_{ii})}{S_{0j}^{(0)}(\hat{\beta}, T_{ii})}, \quad j = 1, 2 \quad (11)$$

Para motivar este estimador, nótese que, siguiendo (3), hemos visto que  $P(dN_{0j}(t) = 1 | \text{Pasado}) = \alpha_{0j;0}(t) \cdot S_{0j}^{(0)}(\beta, t) \cdot dt$ . En ausencia de covariables, es decir, para una muestra homogénea, esta cantidad se convierte en  $\alpha_{0j}(t) \cdot Y_0(t)$ , lo que motiva los incrementos  $\Delta N_{0j}(t)/Y_0(t)$  del estimador de Nelson-Aalen.

El estimador de tipo Nelson-Aalen (11) a menudo se denomina estimador de Breslow (Breslow, 1972). En análisis de riesgos proporcionales estándar (punto final único) datos de supervivencia, el riesgo de línea de base por todas las causas con frecuencia no se considera, ya que se cancela de la verosimilitud parcial. Esto es conveniente en que el analista de datos no necesita considerar el riesgo de línea de base para estudiar el impacto de las covariables en la supervivencia. Sin embargo, esto también puede verse desfavorable, ya que tienta a uno a descuidar el riesgo de referencia, que determina la función de supervivencia junto con  $\beta \cdot Z_i$ . En todo caso, mirando los riesgos de referencia específicos de la causa son aún más importantes en un análisis de competencia de riesgos, como coeficientes de regresión causa-específicos de una magnitud similar pueden tener efectos muy diferentes en las funciones de incidencia acumulada, dependiendo en la magnitud de los riesgos de referencia específicos de la causa.

Ahora consideramos la predicción bajo el siguiente modelo para algún individuo (futuro)  $\tilde{i}$  con covariable individual  $Z_{\tilde{i}=z}$ , que podemos reescribir como covariable individual de causa específica  $Z_{0j;\tilde{i}=z_{0j}}$ . A continuación, eliminamos el índice  $\tilde{i}$  para conveniencia notacional. Los riesgos acumulativos individuales específicos de la causa pueden ser predecidos como:

$$\hat{A}_{0j}(t; z) = \hat{A}_{0j;0}(t) \cdot \exp(\hat{\beta}_{0j}), j = 1, 2 \quad (12)$$

y el peligro acumulativo individual por todas las causas se puede predecir como:

$$\hat{A}_0(t; z) = \hat{A}_{01}(t; z) + \hat{A}_{02}(t; z) \quad (13)$$

Para llegar a un predictor de la función de supervivencia individual por todas las causas, sustituimos los incrementos del estimador habitual de Nelson-Aalen por los incrementos  $\Delta\hat{A}_0(t; z)$ :

$$\hat{P}(T > t|z) = \prod_{T_i \wedge C_i \leq t} (1 - \Delta\hat{A}_0(T_i \wedge C_i; z)) \quad (14)$$

Con incrementos:

$$\Delta\hat{A}_0(t; z) = \sum_{j=1}^2 \frac{\Delta N_{0j}(T_i \wedge C_i; z)}{S_0^{(0)}(\hat{\beta}, T_i \wedge C_i)} \cdot \exp(\hat{\beta} \cdot z_{0j}) \quad (15)$$

Recordando el estimador de la función de incidencia acumulada llegamos a un predictor individual de esta probabilidad sustituyendo el estimador de Kaplan-Meier por (14) y el incremento de la causa específica del estimador de Nelson-Aalen por  $\Delta\hat{A}_{0j}(t; z)$ :

$$\hat{P}(T \leq t, X_T = j|z) = \sum_{T_i \wedge C_i \leq t} \hat{P}(T > (T_i \wedge C_i) - |z) \cdot \frac{\Delta N_{0j}(T_i \wedge C_i)}{S_{0j}^{(0)}(\hat{\beta}, T_i \wedge C_i)} \cdot \exp(\hat{\beta} \cdot z_{0j}) \text{ para } j=1,2 \quad (16)$$

Las propiedades de muestras grandes, como las discutidas anteriormente para  $\beta$ , también están disponibles para el estimador de línea de base acumulativa (11) y las cantidades pronosticadas (12)–(16).

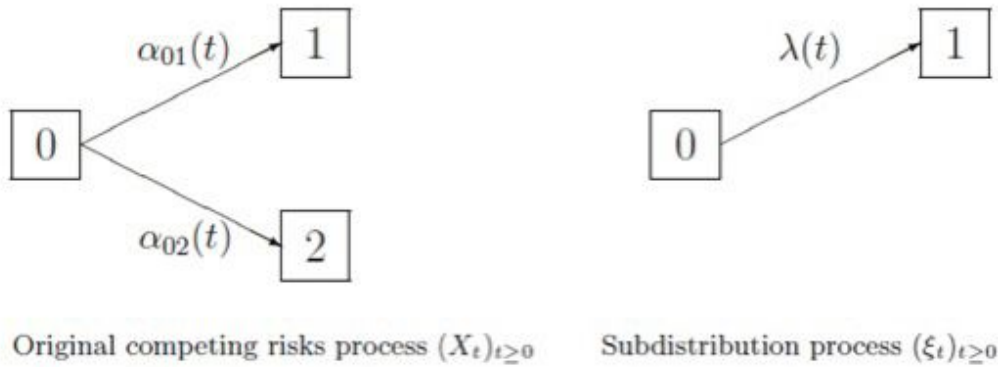
## 5 Modelo de riesgos de subdistribución proporcional

### 5.1 El proceso de subdistribución

El análisis de los datos de riesgos competitivos requiere analizar todos los factores específicos de la causa de riesgos. Como se explica en la introducción, el objetivo de un análisis de subdistribución de riesgos es proporcionar un análisis único de una cantidad, el riesgo de subdistribución, que permite la interpretación directa de los resultados en términos de una función de incidencia acumulada, tal como podríamos interpretar el análisis único de un riesgo por todas las causas en la sección anterior

en términos de la distribución del tiempo de espera en el estado inicial. En la breve introducción, definimos el riesgo de subdistribución  $\lambda(t)$  bastante técnico, requiriendo que la función de incidencia acumulada  $P(T \leq t, X_T = 1)$  para el tipo de evento 1 es igual a  $\exp(\int_0^t \lambda(u) du)$ . Esto imita la fórmula habitual de función de supervivencia menos uno, pero con el riesgo de subdistribución reemplazando el riesgo de todas las causas. En esta sección, nos acercamos a la subdistribución marco desde un punto de vista de proceso/estado múltiple, que creemos que es más intuitivo. También tiene la ventaja de mostrar claramente la conexión con el modelo original de riesgos competitivos y es fácilmente generalizable a otros modelos multiestado con estados absorbentes en competencia. Considere el proceso original de riesgos en competencia  $(X_t)_{t \geq 0}$  de la figura. El tiempo de evento T es el tiempo de espera de  $(X_t)_{t \geq 0}$  en el estado inicial 0 (es decir, hasta que ocurra cualquier primer evento). La idea de el marco de subdistribución es considerar un tiempo de subdistribución  $\theta$  hasta ocurrencia del evento 1 de interés.

#### 124 5 Proportional hazards models



Esto debe tener en cuenta el hecho de que la ocurrencia de un primer evento puede ser del tipo competitivo 2. La manera correcta de hacer esto es deteniendo el proceso original de riesgos competitivos justo antes de T, si el proceso pasa al estado de evento competitivo 2. Por lo tanto, la subdistribución proceso  $(\xi_t)_{t \geq 0}$  de la figura se define como:

$$\xi(t) := 1(X(t) \neq 2) \cdot X(t) + 1(X(t) = 2) \cdot X(T-) \quad (17)$$

En palabras,  $X_t$  y  $\xi_t$  son iguales, siempre que el proceso original no se mueva al estado de evento competitivo 2. Si  $(X_t)_{t \geq 0}$  pasa al estado 2 en el tiempo T, los individuos 'se estancan' en el estado inicial del proceso de subdistribución:  $\xi_t$  permanece igual a 0 para todo  $t \geq T$ , si  $X_T = 2$ . En consecuencia, la interpretación de estar en los respectivos estados iniciales 0 cambia:  $X(t) = 0$  tiene la interpretación 'ningún evento por el tiempo t', y  $\xi(t) = 0$  significa 'ningún evento tipo 1 por tiempo t'. A partir de esto, ahora podemos definir el tiempo hasta la ocurrencia de el evento 1 de interés como:

$$\theta := \inf\{t > 0 | \xi_t \neq 0\} \quad (18)$$

El conjunto del lado derecho de (18) puede estar vacío. Si  $X_T = 2$ ,  $\xi_t$  es igual a 0 para todos los tiempos t y no hay tiempo más pequeño tal que el proceso de subdistribución es distinto de 0. En tal caso,  $\theta$  es el mínimo de un conjunto vacío, que (como es habitual) se define como infinito:  $\theta = \inf$ , si  $X_T = 2$ . La interpretación de esto es que no hay un tiempo finito en el que entra el proceso de subdistribución el estado de evento 1, si  $X_T = 2$ . Sin embargo, si el proceso original de riesgos competitivos entra en el estado 1, también lo hace el proceso de subdistribución, y tenemos  $T = \theta$ . En consecuencia, la función de distribución del fallo de subdistribución el tiempo  $\theta$  es igual a la función de incidencia acumulada para el evento tipo 1,

$$P(\beta \leq t) = P(T \leq t, X_t = 1) \text{ para todo } t \in [0, \inf) \quad (19)$$

y  $P(\theta = \inf) = P(X_T = 2)$  La subdistribución de riesgo ahora esta definida como el riesgo 'adjunto a  $\theta$ ',

$$\lambda(t)dt := P(\theta \in dt | \theta \geq t) \quad (20)$$

Tenga en cuenta que esta definición es paralela a la definición del peligro por todas las causas en pero con  $\theta$  en lugar de T. Como consecuencia,

$$P(\theta \leq t) = 1 - \exp\left(-\int_0^t \lambda(u)du\right) \quad (21)$$

Debido a las Ecuaciones (19) y (21), el marco de subdistribución restablece una correspondencia uno a uno entre el riesgo de subdistribución y la función de incidencia acumulada, que de otro modo es una función involucrada de ambos peligros de causa específica  $\alpha_{01}(t)$  y  $\alpha_{02}(t)$ :

$$1 - \exp\left(-\int_0^t \lambda(u)du\right) = \int_0^t \exp\left(-\int_0^u \alpha_{01}(u) + \alpha_{02}(v)dv\right) \alpha_{01}(u)du \quad (22)$$

La idea del modelo de riesgos de subdistribución proporcional sugerido por Fine y Gray (1999) y discutido a continuación es ajustar un modelo de Cox a la subdistribución de riesgo, que tiene una interpretación de probabilidad directa en términos de la función de incidencia acumulada  $P(T \leq t, X_T = 1)$  para el evento tipo 1 debido a Ecuaciones (19) y (21). Sin embargo, ajustar el modelo en la práctica tiene que tratar con ciertas dificultades técnicas que implica que  $\theta$  tenga masa en el infinito. Esto es se discute más adelante.

La ecuación 22 implica que:

$$\alpha_{01}(t) = \left(1 + \frac{P(t \leq t, X_T = 2)}{P(T > t)}\right) \cdot \lambda(t) \quad (23)$$

La ecuación (23) tiene varias consecuencias importantes:  $\lambda(t)$  se pondera hacia abajo en comparación a  $\alpha_{01}(t)$ , asegurando que  $P(X_T = 1) = \lim_{t \rightarrow \infty} \exp\left(-\int_0^t \lambda(u)du\right) = 1 - P(X_T = 2)$ . Como la ponderación depende del tiempo, suponiendo una relación proporcional modelo de riesgos para  $\alpha_{01}(t)$  excluye que  $\lambda(t)$  siga una relación proporcional modelo de riesgos de subdistribución, y viceversa. Nos ocupamos de las consecuencias de tales modelos potencialmente mal especificados en los análisis de datos. La ponderación también depende de la ocurrencia de los eventos competitivos, lo que refleja que una función de incidencia acumulada depende de todas las causas específicas de riesgos. (Latouche (2004) sugirió una buena presentación gráfica de este hecho).

Además, podemos calcular el riesgo de subdistribución basado en el conocimiento del proceso original, como cabría esperar de nuestra derivación del proceso de subdistribución. Por lo tanto, deberíamos ser capaces de hacer una conjetura decente sobre el resultados de un análisis de riesgos de subdistribución, si previamente hemos analizado todos riesgos específicos de la causa, incluidos los riesgos básicos. Esto también se aborda en los análisis de datos. Notamos que Gray (1988) primero sugirió considerar la falla de subdistribución veces al analizar datos de riesgos competitivos censurados por la derecha. Gray llamó a la tiempo de falla de subdistribución "impropia", porque su distribución tiene masa en el infinito. También señaló que el enfoque del proceso de conteo (Andersen et al., 1993) en principio permite tales variables aleatorias. A continuación, en la Sección 5.3.1, considerar ajustar un modelo de riesgos de subdistribución proporcional, junto con determinadas dificultades técnicas relacionadas con el fallo de la subdistribución en ese momento "impropio". También notamos que definir un tiempo de evento como infinito, si el evento de interés no ocurre, es natural desde un punto de modelado probabilístico de vista.

## 5.2 Estimación

Estimación y predicción para el modelo de riesgos de subdistribución proporcional es, en espíritu, similar a un modelo estándar de Cox. Explicamos el proceso de subdistribución, sin embargo, plantea

un problema en el sentido de que el riesgo asociado a un conjunto puede ser desconocido. Las soluciones se han desarrollado principalmente para datos censurados por la derecha y requieren un modelado explícito de la distribución de censura. En la práctica, típicamente se asume una censura aleatoria. Describimos los enfoques disponibles que asumen la censura administrativa o abordan el problema de la incógnita en conjuntos de riesgo utilizando probabilidad inversa de ponderación de censura o imputación múltiple.

**Proceso de conteo de eventos de subdistribución observados, proceso de riesgo de subdistribución y dificultades técnicas**

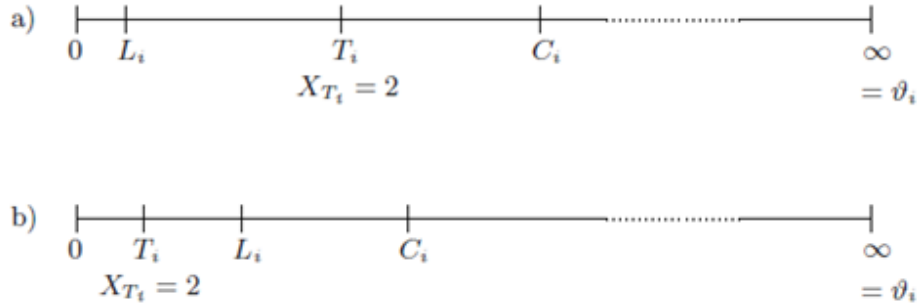
En principio, el modelo de riesgos de subdistribución proporcional es "solo" un modelo proporcional del modelo de riesgos bajo una estructura de datos, podemos invocar las respectivas técnicas de estimación, simplificadas en la medida en que solo tienen un riesgos de subdistribución en lugar de dos riesgos de causa específica. Definimos el proceso de conteo individual:

$$N_{i*}(t) := 1(\theta_i \wedge C_i, L_i < \theta_i \leq C_i) \quad (24)$$

donde  $N_{01;i}(t)$  y  $N_{01}(t)$  son los procesos de conteo originales de tipo 1.

La dificultad técnica del marco de subdistribución, aludida en el apartado anterior, surge a través de los procesos en riesgo  $Y_{0;i}(t)$ . Los lectores pueden comprobar fácilmente que estos son conocidos por los individuos en estudio y sin tipo 1, tipo 2 ni evento de censura todavía, para individuos con un evento tipo 1 observado y para individuos, que han sido censurados por la derecha. Sin embargo,  $Y_{0;i}(t)$  es desconocido para las siguientes personas.

1. Individuos que han sido observados para experimentar un evento de tipo 2; esto es ilustrado en la Figura.
2. Individuos que nunca ingresaron al estudio, porque su tiempo de falla fue menor que o igual a su tiempo de truncamiento a la izquierda, y que experimentó un evento tipo 2 ; esto se ilustra en la Figura.



Discutimos la Figura, primero. Para tales individuos, sabemos tanto  $T_i$  como que  $X_{T_i} = 2$ . Como consecuencia, sabemos que  $\theta_i = \inf \theta_i \wedge C_i = C_i$ .

Sin embargo, la observación puede haberse detenido en  $T_i$ ; por ejemplo, si el individuo deja de existir después, porque el punto final competitivo 2 es la muerte. Pero incluso si el el estado competitivo 2 no es fatal en la vida real, el protocolo de estudio puede haber planeado recopilar datos solo hasta  $T_i$  (independientemente del tipo de punto final) o hasta  $C_i$ , lo que venga primero. Como consecuencia, en general, no sabemos  $C_i$ , que es un evento futuro con respecto a  $T_i$  en la Figura a). Por lo tanto,  $Y_{0;i}(t)$  es desconocido para tal individuo para  $t \in (T_i, C_i]$ . A continuación, discutimos la Figura b): el individuo ilustrado en esta figura nunca entra en el estudio, porque  $T_i$  es menor que  $L_i$ . Esto no sería problemático, si  $X_{T_i}$  fuera igual a 1. Aquí, sin embargo, el individuo experimenta el tipo de evento competitivo (es decir,  $X_{T_i} = 2$ ) y tenemos nuevamente que  $\theta_i = \inf$ . Como consecuencia,  $\theta_i > L_i$ , es decir, deberíamos haber visto a este individuo en términos del proceso de



subdistribución, y debería haber estado en riesgo (en el sentido de  $Y_{0;i}(t)$ ) en el intervalo de tiempo  $(L_i, C_i)$ . El problema aquí es doble: porque el individuo no ha ingresado al estudio en la "vida real", por lo general ni siquiera estamos conscientes del individuo ni conocemos a  $L_i$  y  $C_i$ . A continuación, consideramos los datos de riesgos competitivos censurados aleatoriamente por la derecha. Brevemente comentamos sobre el truncamiento a la izquierda. Observamos que el supuesto de censura aleatoria es más restrictivo que la suposición de censura independiente, que normalmente se hace al modelar riesgos de causa específica. En particular, la censura puede depender de las covariables incluidas en el modelo, si los riesgos específicos de la causa son analizados. Tal dependencia potencial no se tiene en cuenta cuando se supone censura aleatoria. Se desconoce el motivo para suponer una censura aleatoria. procesos de subdistribución en riesgo. Como mostramos a continuación, esta dificultad técnica se aborda estudiando directamente la distribución de la censura. En principio, como se ha comentado por Fine y Gray (1999) y Ruan y Gray (2008), este enfoque también puede extenderse a la dependencia del modelo del mecanismo de censura en las covariables.

**Datos de riesgos competitivos censurados administrativamente por la derecha: Conocidos conjunto de riesgos de subdistribución**

Como (24) es en general desconocida, la idea de Fine and Gray era reemplazar  $Y_{0;i}(t)$  por un conjunto de riesgo "estimado".

$$Y_{0;i}^*(t) = 1(t \leq T_i \wedge C_i) + 1(T_i < t < C_i, X_{T_i} = 2) \quad (25)$$

en ausencia de truncamiento a la izquierda. El primer sumando en el lado derecho de (25) es conocido, el segundo sumando generalmente es desconocido. Una excepción, donde se conoce  $1(T_i < t \leq C_i, X_{T_i} = 2)$ , es censura administrativa.

En este caso, la censura se debe únicamente a la terminación administrativa del estudio, tal que se conoce el tiempo potencial de censura de un individuo para cada individuo por adelantado. Como consecuencia, la dificultad técnica discutida anteriormente en el contexto de la figura a) desaparece y podemos usar el software estándar de Cox.

Debido a que el proceso de subdistribución en riesgo es conocido en tal situación, ha sido denominado "censura de datos completos" en Fine y Gray (1999). Este mecanismo de censura también se conoce como censura Tipo I (progresiva, generalizada).

**Datos de riesgos competitivos censurados aleatoriamente por la derecha: "estimados" conjunto de riesgos de subdistribución**

Como (25) es en general desconocida, la idea de Fine and Gray era reemplazar  $Y_{0;i}(t)$  por un conjunto de riesgo "estimado".

$$\hat{Y}_{0;i}(t) := 1(C_i \geq T_i \wedge t) \cdot \frac{\hat{G}(t-)}{(\{T_i \wedge \hat{C}_i \wedge t\}-)} \cdot (1(t \leq T_i) + 1(T_i < t, X_{T_i} = 2)) \quad (26)$$

donde  $\hat{G}(t)$  es el estimador de Kaplan-Meier de la función de supervivencia de censura  $P(C > t)$  usando el "indicador de evento de censura"  $1(T \leq C)$ . Usando  $\hat{G}(t)$  como en (26) es una aplicación de probabilidad inversa de censura ponderación (Robins y Rotnitzky, 1992); la ponderación se aplica a  $1(t \leq T_i) + 1(T_i < t, X_{T_i} = 2)$ , que es el riesgo de subdistribución fijado en el caso de datos completos. La idea detrás del estimador (26) es esta:

Primero, es computable a partir de los datos observables. El término  $1(C_i \geq T_i \wedge t)$  es igual a uno, si tenemos conocimiento del estado vital del individuo  $i$  justo antes de  $t$ :  $1(C_i \geq T_i \wedge t) = 1$  para individuos sin tipo 1, tipo 2, ni evento de censura en  $[0, t)$  (si  $t \leq T_i$ ) y para individuos con un tipo 1 o tipo 2 observado evento en  $[0, t)$  (si  $T_i < t$ ). Para tal individuo, el resto de (26) puede calcularse a partir de los datos observables. Tenga en cuenta que el último factor en el lado derecho de (26) es igual a 1, si y solo si la subdistribución del individuo  $i$ .

Segundo,  $Y_{0;i}(t)$  es asintóticamente imparcial en el sentido de que su expectativa es igual a  $E(Y_{0;i}(t))$  asintóticamente. Para ver esto, note que (26) equivale a

$$1(t \leq T_i \wedge C_i) + 1(C_i \geq T_i) \frac{\hat{G}(t-)}{\hat{G}(T_i-)} 1(T_i < t, X_{T_i} = 2) \quad (27)$$

Debido a la representación de  $Y_{0;i}(t)$  en (24), solo necesitamos considerar el segundo sumando en (27). Ahora considere la expectativa de este término condicional a  $T_i = s, X_{T_i} = j$ , es decir,

$$\begin{aligned} & E(1(C_i \geq) \frac{\hat{G}(t-)}{\hat{G}(s-)} 1(s < t, j = 2) | T_i = s, X_{T_i} = j) \\ &= 1(s < t, j = 2) \cdot E(1(C_i \geq) s \frac{\hat{G}(t-t)}{\hat{G}(s-)} | T_i = s, X_{T_i} = j) \\ &= 1(s < t, j = 2) \cdot P(C_i \geq s) \cdot E(\frac{\hat{G}(t-t)}{\hat{G}(s-)} | T_i = s, X_{T_i} = j) \end{aligned} \quad (28)$$

Tenga en cuenta que el término  $\hat{G}(t)/\hat{G}(s)$  es igual al estimador de Kaplan-Meier de  $P(C \geq t | C > s)$ , asumiendo que no ocurre censura en  $s$  cuando  $T_i = s$ .  $\hat{G}(t)/\hat{G}(s)$  tiene la forma de un estimador habitual de Kaplan-Meier de una función de supervivencia de censura, pero tomando el tiempo  $s$  como el nuevo origen del tiempo. Por lo tanto, el  $i$ -ésimo individuo hace que ya no contribuya a  $\hat{G}(t)/\hat{G}(s)$ .

Usando eso, el estimador de Kaplan-Meier converge en distribución hacia la verdadera función de supervivencia de censura y que la esperanza de una variable aleatoria es la esperanza de la esperanza condicional de esa variable aleatoria, encontramos que, asintóticamente, la esperanza del segundo sumando en (27) es igual a:

$$G(t-)p(T < t, X_T = 2)$$

que también es igual a la esperanza del segundo sumando de la representación de  $Y_{0;i}^*(t)$  en (26). La estimación de los coeficientes de regresión a partir de una subdistribución proporcional de modelo de riesgos basado en una probabilidad parcial con conjuntos de riesgos "estimados" es implementado en el paquete R `cmprsk`. El paquete también proporciona un estimador de la matriz de covarianza de los coeficientes de regresión estimados. Sin embargo, el estimador de covarianza tiene una forma más complicada que  $(J_\tau(\hat{\beta}))^{-1}$  de la Sección 5.2.1 sobre la estimación a partir de modelos de riesgos proporcionales de causa específica. Esto se deriva del hecho de que la estimación del riesgo de subdistribución el conjunto también contribuye a la covarianza. Una derivación del estimador de covarianza está más allá del nivel técnico de este libro. Se remite a los lectores a Fine and Gray (1999), donde el estimador se deriva explícitamente mediante un proceso empírico (van der Vaart y Wellner, 1996). También notamos que Geskus (2011) ha utilizado recientemente argumentos de martingala en este contexto.

### **Datos de riesgos competitivos censurados aleatoriamente por la derecha: Múltiples imputación de tiempos de censura faltantes**

Como se explicó anteriormente, ajustar un modelo de riesgos de subdistribución proporcional es sencillo, si el tiempos potenciales de censura para individuos con un evento competitivo observado son conocidos. En tal caso, se puede usar el software estándar de Cox como, por ejemplo, el proporcionado por el paquete de supervivencia. En ausencia de tal conocimiento, el paquete `cmprsk` permite ajustar un modelo de riesgos de subdistribución proporcional.

Una desventaja de tener que usar `cmprsk` es que el paquete ofrece menos funcionalidad que la `survival`. Por ejemplo, `survival` también permite a los usuarios especificar estratos con peligros de línea de base específicos de estratos potencialmente diferentes, incluyen debilidades (es decir, efectos) o maximizar una probabilidad parcial penalizada (cf. Therneau y Gramsch (2000)). Ninguno de estos análisis es factible con `cmprsk`. Ruan y Gray (2008) por lo tanto sugirieron un enfoque de imputación que permite un análisis de riesgos de subdistribución utilizando supervivencia, incluso si

se desconocen los posibles tiempos de censura. La idea es recuperar esta información faltante de un estimador de la censura de la función de supervivencia  $P(C > t)$  en múltiples pasos de imputación. Brevemente, el procedimiento de imputación múltiple funciona de la siguiente manera:

1. El objetivo es imputar tiempos de censura a partir de una distribución de censura determinada. Uno calcula el estimador de Kaplan-Meier ( $\hat{t}$ ) de la función de supervivencia censurada  $P(C > t)$  usando el "indicador de evento de censura"  $11(T \leq C)$ . Esto proporciona una distribución adecuada de los tiempos de censura observados, a menos que el tiempo de evento más grande no sea un evento de censura. En este caso, uno agrega otro tiempo de censura a los datos, que es más grande que el más grande tiempo del evento, antes de calcular ( $\hat{t}$ ).
2. Imputación: Considere a todos los individuos en los datos para quienes un competidor se ha observado un evento de tipo 2. Deja que sea un individuo fracasado al tiempo  $T_i$ ,  $T_i \geq C_i$  y  $X_{T_i} = 2$ . El potencial tiempo futuro de censura  $C_i$  no es conocido. Para el individuo  $i$ , se imputa un tiempo de censura sorteando al azar de la distribución condicional  $1 - \hat{P}(C > t | C > T_i) = 1 - \hat{G}(t) / G(T_i)$ ,  $t > T_i$ .
3. Imputación múltiple: Repetir el paso anterior  $k$  veces. Esto da como resultado  $k$  conjuntos de datos, en los que el individuo  $i$  es censurado en su respectiva censura imputada.
4. Múltiples análisis: El análisis, normalmente ejecutando coxph, se realiza para cada uno de los  $k$  conjuntos de datos.
5. Estimaciones agrupadas: denotemos por  $\hat{\gamma}^{(l)}$  a la estimación de  $\gamma$  obtenida de el  $i$ -ésimo análisis,  $i = 1, \dots, k$ . También escriba  $(J_\tau(\gamma^{(l)}))^{-1}$  para la estimación matriz de covarianza del análisis  $l$ -ésimo. La estimación de la coeficientes de regresión basados en imputación múltiple entonces es:

$$\hat{\gamma} = \sum_{l=1}^k \hat{\gamma}^{(l)} / k$$

y el estimador de la matriz de covarianza es (Schanfer, 1997)

$$\sum_{l=1}^K (J_\tau(\hat{\gamma}^{(l)}))^{-1} / k + \frac{k+1}{k} \sum_{l=1}^k (\hat{\gamma}^{(l)} - \hat{\gamma})^T (\hat{\gamma}^{(l)} - \hat{\gamma}) / (k-1)$$

donde la primera parte es la varianza dentro de la imputación estimada, y la segunda parte proviene de la varianza entre imputaciones.

Ruan y Gray (2008) también consideraron usar un estimador de la censura función de supervivencia basada en muestras de arranque para dar mejor cuenta de la incertidumbre en la estimación de  $P(C > t)$ . En este enfoque, el estimador de Kaplan-Meier de la función de supervivencia de censura se calcula a partir de una muestra de arranque dibujado con reemplazo de los datos originales. Una nueva muestra de arranque de extrae para cada paso de imputación, y la muestra de arranque solo se usa para estimar  $P(C > t)$ . Tanto para datos reales como simulados, Ruan y Gray encontraron resultados similares para el procedimiento de imputación múltiple simple y el que adicionalmente utiliza el arranque.

**Prediction** La predicción de funciones de incidencia acumulada bajo una subdistribución proporcional la suposición de peligros funciona de manera análoga a la Sección anterior. El paso clave es obtener un estimador tipo Breslow de la subdistribución acumulada de riesgo de referencia  $\Lambda_0(t) = \int_0^t \lambda_0(u) du$ . Considera el caso simple de datos de riesgos competitivos censurados administrativamente por la derecha primero. En este caso, todos los tiempos de subdistribución censurados son conocidos, y un estimador está dado por:

$$\hat{\Lambda}_0(t) := \sum_{\theta \wedge C_i \leq t} \frac{\Delta N_{01}(\theta_i \wedge C_i)}{\sum_{i=1^n} \exp(\hat{\gamma} \cdot Z_i) \cdot Y_{0;i}^*(\theta_i \wedge C_i)} \quad (29)$$

donde  $\hat{\gamma}$  es el coeficiente de regresión estimado de la subdistribución proporcional modelo de riesgos. Un predictor de la función de incidencia acumulada para algún (futuro) individuo  $\mathbb{A}$  con covariable  $Z_{\mathbb{A}} = z$  es:

$$\hat{P}(T, X_T = 1|z) = 1 - \prod_{\theta_i \wedge C_i \leq t} (1 - \Delta \hat{\Lambda}_0(\theta_i \wedge C_{i;z})) \quad (30)$$

Donde  $\hat{\Lambda}_0(t; z) = \hat{\Lambda}_0(t) \exp(\hat{\gamma}z)$ . Tenga en cuenta que el producto en la última parte sólo puede calcularse sobre tiempos  $\theta_i \wedge C_i$  con  $\theta_i \leq C_i$ .

Si la censura no es enteramente administrativa, podemos proceder como antes. Nosotros ya sea sustituir  $Y_{0;i}^*(t)$  por (26) o usamos imputación múltiple. Con el fin de aproximar la distribución asintótica de las cantidades predichas, uno puede considerar la posibilidad de arrancar los datos.

## 6 Modelado en presencia de riesgos competitivos

### 6.1 Modelando el peligro de la función de incidencia acumulada

#### 6.1.1 Detalles teóricos

Esta sección presenta el trabajo de Fine y Gray (1999) sobre modelado de la peligrosidad del CIF. La metodología tiene muchas puntos en común con la prueba de k-muestra introducida por Gray (1988) y el modelo de riesgos proporcionales de Cox. Aunque la teoría se desarrolla gradualmente (a partir del caso de no censura), se presenta el caso general de datos incompletos aquí. Este es el tipo de datos que solemos encontrar: hay personas físicas con un evento de interés registrado, otras con tipo de evento de riesgo competitivo y otros que están censurados. El modelo se basa en

$$\gamma(t, x) = \gamma_0(t) e^{\beta x}, \quad (31)$$

donde  $\gamma$  es el riesgo de la subdistribución,  $\gamma_0$  es la línea de base riesgo de la subdistribución,  $x$  es el vector de las covariables y  $\beta$  es el vector de los coeficientes. La forma de la verosimilitud parcial es similar a la utilizada en el Cox modelo de riesgos proporcionales. Escrito para una sola covariable,  $x$ , la verosimilitud parcial viene dada por la fórmula

$$L(\beta) = \prod_{j=1}^r \frac{\exp(\beta x_j)}{\sum_{i \in R_j} w_{ji} \exp(\beta x_i)} \quad (32)$$

El producto se toma  $r$  todos los puntos de tiempo  $(t_1 < t_2 < \dots, t_r)$ , donde se observó un evento de interés. El conjunto de riesgos  $R_j$  se define de otra manera se suman formas y pesos  $w_{ji}$ . El conjunto de riesgo está formado por aquellos que no experimentó un evento en el tiempo  $t$  y de los que experimentaron un evento de riesgo competitivo en el tiempo  $t$ :

$$R_j(t) = \{i; T_i \leq t \text{ (} T_i \geq t \text{ y el sujeto experimentó un evento de riesgo competitivo)}\} \quad (33)$$

Así, aquellos que vivieron otro tipo de eventos permanecen en el riesgo establecido todo el tiempo. Los pesos se definen como

$$w_{ji} = \frac{\hat{G}(t_j)}{\hat{G}(\min(t_j, t_i))}, \quad (34)$$

donde  $\hat{G}$  es la estimación de Kaplan-Meier de la función de supervivencia de la distribución de la censura. La distribución de la censura está definida por  $(T_i, C_i)$ , donde  $T_i$  es el momento del primer evento y  $C_i$  es 1 si no se observó un evento y 0 si se observó algún tipo de evento. En cada punto de tiempo donde ocurrió un evento de interés (indexado por  $j$ ) el conjunto de individuos en riesgo (indexado por  $i$ ) incluye a aquellos que no tuvo ningún tipo de evento en el tiempo  $t_j$  así como aquellos que tuvo un evento de riesgos competitivos antes de  $t_j$ . El peso es 1 para el primero y menor o igual a 1 para el segundo. De esta forma aquellos las personas que experimentan un evento de riesgo competitivo no participan plenamente en la verosimilitud parcial: cuanto más lejos esté el punto temporal ( $t_j$ ) es desde el momento en que ocurrió el evento de riesgo competitivo, el menor el peso.

Aplicando la derivada al logaritmo de la verosimilitud parcial dada en 10 da el estadístico de puntuación, que tiene la forma

$$U(\beta) = \sum_{j=1}^r \left\{ x_j - \frac{\sum_{i \in R_j} w_{ji} x_i \exp(x_i \beta)}{\sum_{i \in R_j} w_{ji} \exp(x_i \beta)} \right\}$$

La estimación de  $\beta$ , denotada por  $\hat{\beta}$ , es el valor que maximiza la función  $U(\beta)$ .  $\hat{\beta}$  no tiene forma cerrada. La ecuación 13 puede escribirse para un conjunto de covariables y puede extenderse para incluir covariables dependientes del tiempo.

### 6.1.2 Estimación basada en modelos del CIF

El CIF predicho se puede calcular para un cierto valor de las covariaciones usando la formula

$$f(t) = 1 - \exp(-H(t))$$

Donde  $H(t)$  es el riesgo acumulativo de la subdistribución.  $\hat{H}(t)$  se puede calcular usando un estimador tipo Breslow:

$$\hat{H}(t; x : 0, \hat{\beta}) = \sum_{t_j \leq t} \left\{ \frac{\exp(x_0 \hat{\beta})}{\sum_{i \in R_j} w_{ji} \exp(x_i \hat{\beta})} \right\}$$

La suma externa es sobre todos los puntos de tiempo para los que se observó un evento de interés.

## 7 Análisis de riesgos de subdistribución proporcional de todas las funciones de incidencia acumulada

El hecho de que el análisis de subdistribución permita una interpretación de probabilidad directa lo ha hecho popular en las aplicaciones. Esto es particularmente relevante, porque en los análisis de riesgos competitivos aplicados basados en los peligros habituales de causa específica, a menudo falta un análisis del peligro competitivo de causa específica.

Una limitación del enfoque de riesgos de subdistribución en comparación con los riesgos de causa específica es que solo estos últimos determinan completamente el comportamiento estocástico de todo el proceso de riesgos competitivos. El riesgo de subdistribución solo especifica la función de incidencia acumulada de interés, pero no las funciones de incidencia acumulada competidoras. Es por eso que decimos que la forma en que las funciones de incidencia acumulada evolucionan a lo largo del tiempo se comprende mejor desde una perspectiva de peligros de causa específica.

Los hechos antes mencionados han llevado a que el modelado de riesgos de subdistribución proporcional se aplique a ambas funciones de incidencia acumulada en las aplicaciones. Conceptualmente, este enfoque presenta problemas si se supone que ambos riesgos de subdistribución siguen un modelo

de riesgos de subdistribución proporcional.

Denotemos como  $\Lambda_0(t)$  al riesgo de la base de la subdistribución acumulada para el evento de interés con el coeficiente de regresión  $\gamma$  del modelo  $\lambda_i(t; Z_i) = \lambda_0(t) \cdot \exp(\gamma \cdot Z_i)$ ,  $i = 1, \dots, n$ , y  $\hat{\Lambda}_0(t)$ ,  $\hat{\gamma}$  para las cantidades respectivas adjuntas a la función de incidencia acumulada competidora, implica que

$$1 - \exp(-\exp(\gamma Z_i) \Lambda_0(\infty)) = \exp(-\exp(\hat{\gamma} Z_i) \hat{\Lambda}_0(\infty)),$$

asumiendo modelos de riesgos de subdistribución proporcional para ambas funciones de incidencia acumulada. Esto implica que el coeficiente de regresión  $\gamma$  del modelo original y el límite de todos los riesgos de subdistribución de la base, es decir, tanto  $P(X_T = 1)$  como  $P(X_T = 2)$  en la situación de la base, determinan el coeficiente de regresión del modelo competidor. Sin embargo, esta restricción no se tiene en cuenta cuando se ejecuta un análisis de riesgos de subdistribución proporcional dos veces, una para el evento de interés y otra para el evento competitivo. Los lectores pueden verificar que este enfoque puede conducir a estimaciones inconsistentes basadas en modelos para todas las funciones de incidencia acumulada usando `predict.crr` también para el evento competitivo en un análisis de datos simulados.

Por lo tanto, es natural interpretar dichos análisis en términos de un efecto promediado en el tiempo, si se llevan a cabo análisis de riesgos de subdistribución proporcional para todas las funciones de incidencia acumulada.

## 8 Bibliografía

- 1.- Pintilie, M. (2006). Competing risks: a practical perspective (Vol. 58). John Wiley Sons.
- 2.- Beyersmann, J., Allignol, A., Schumacher, M. (2011). Competing risks and multistate models with R. Springer Science Business Media.
- 3.- Geskus, R. B. (2019). Data analysis with competing risks and intermediate states. Chapman and Hall/CRC.