

REVISIÓN

Modelos marginales y condicionales en el análisis de supervivencia multivariante

M.A. Barceló

Grup de Recerca en Estadística, Economia Aplicada i Salut (GRECS). Departament d'Economia. Universitat de Girona.
Centre de Recerca en Economia de la Salut (CRES). Universitat Pompeu Fabra. Barcelona.

Correspondencia: Prof. Maria Antònia Barceló. Departament de Economia. Universitat de Girona. Campus de Montilivi. 17071 Girona.
Correo electrónico: antonia.barcelo@udg.es

Recibido: 27 de febrero de 2002.

Aceptado: 15 de julio de 2002.

(Marginal and conditional models in multivariate survival analysis)

Resumen

El hecho de que los individuos puedan experimentar varios sucesos (fallos) durante el período de observación (análisis de supervivencia multivariante) hace que sea necesaria la aplicación de una metodología diferente de la empleada en el análisis de supervivencia estándar univariante. Al analizar medidas repetidas, el principal problema del modelo de Cox se debe a que las observaciones no son independientes, lo cual llevará al incumplimiento de la hipótesis de riesgos proporcionales. Con el propósito de solucionar este problema, así como otros que puedan darse debido a este tipo de datos, aparece el modelo de Andersen-Gill (AG), que no es más que una generalización del modelo de Cox. Sin embargo, la hipótesis clave de este modelo (hipótesis de incrementos independientes) supone que las múltiples observaciones de un mismo individuo son independientes, condicionados a las variables explicativas. En la práctica, esta limitación puede conducir a estimadores sesgados e ineficientes, sobrestimándose normalmente la precisión de los mismos ya que, por lo común, las observaciones de un mismo individuo suelen estar positivamente correlacionadas. Con el fin de hallar una solución a esto, aparecen los modelos marginales y los modelos condicionales. La principal diferencia entre estos modelos es el trato que recibe la dependencia. Mientras que los primeros estiman el modelo ignorando la dependencia entre observaciones, corrigiéndola posteriormente mediante estimadores *jackknife*, *bootstrap* o «sándwich», los segundos la estiman especificando explícitamente la distribución de probabilidad de la misma e incorporándola en el modelo. Nuestro objetivo en este artículo consiste en ilustrar estas dos aproximaciones utilizando una base de datos de infecciones nosocomiales.

Palabras clave: Análisis de supervivencia multivariante. Enfoque marginal. Enfoque condicional. Modelos de fragilidad.

Abstract

The fact that individuals can present more than one event (fail) in the observation period (multivariate survival analysis) demands a different methodology from that used in univariate survival analysis. The main problem of the Cox model with multivariate data is that the observations are not independent implying, among others, the violation of the proportionality hypothesis. The Andersen-Gill approximation to the Cox model (AG) overcomes in part this problem. The underlying hypotheses in the AG model, however, are very restrictive, in particular that of independent increments. Under this hypothesis the multiple observations of an individual are independent, although conditioned on the explanatory variables. In practice, this limitation could let us to biased and inefficient estimators that are usually overestimated because, in general, the observations of an individual use to be positively correlated. As a solution to this problem, marginal and conditional models turn up. The main difference between these models is the treatment that dependence receives. Whereas the first one estimates the model ignoring the dependence between recurrences and adjusts this at the end with jackknife, bootstrap or «sandwich» estimates; the second one estimates the dependence specifying explicitly the probability distribution of the dependence and incorporating this into the model. Our objective in this paper is to illustrate these two approaches using nosocomial infection data.

Key words: Multivariate survival data. Marginal models. Conditional models. Frailty models.

Introducción

En los estudios de supervivencia se está interesado por el tiempo que transcurre desde un origen (nacimiento, inicio de tratamiento, etc.) hasta que ocurre un suceso (muerte, mejora, etc.)¹. En numerosas ocasiones este suceso no es único y los sujetos pueden experimentar varios sucesos (o técnicamente fallos) que pueden ser de distinto tipo o bien tratarse de recurrencias del mismo suceso de interés. Debido a esto, los individuos con más sucesos estarán «sobrerrepresentados»² y, por otra parte, los efectos de las variables explicativas (efectos «tratamiento») podrán ser distintos entre el primer suceso y los siguientes. Esto se traducirá en una posible dependencia entre los tiempos de recurrencia, consecuencia de dos fenómenos diferentes: la existencia de heterogeneidad individual y dependencia «serial». En la primera, los tiempos de supervivencia dependen unos de otros porque los individuos comparten una o más covariables omitidas. Esta heterogeneidad individual también se denomina fragilidad, dependencia entre individuos o falso «contagio». Por lo que respecta a la dependencia «serial», que equivaldría a la autocorrelación en series temporales, la dependencia se produce en el mismo individuo, entre sus observaciones repetidas. Como consecuencia de que el individuo presenta una conducta dinámica *per se*, la probabilidad de que se produzca un suceso no es independiente de que se produzca uno previo. Esta dependencia también se denomina dependencia dentro de un mismo individuo o «contagio» verdadero.

Con el fin de ilustrar el trabajo utilizaremos una cohorte prospectiva formada por todos los pacientes admitidos un mínimo de 24 h en la Unidad de Cuidados Intensivos (UCI) del Hospital Universitario Dr. Josep Trueta de Girona, durante el período comprendido entre el 15 de marzo y el 15 de junio de 1999³. Se trata de un hospital público de nivel terciario, con 12 camas en la UCI, que cubre prácticamente toda la población de la provincia de Girona. Se pretende analizar los factores que determinan la incidencia de las infecciones nosocomiales en la UCI de dicho hospital. Las infecciones nosocomiales vienen definidas como las causadas por microorganismos adquiridos en el hospital, que afectan a los pacientes por un proceso diferente del de la infección, y que no estaban presentes en la admisión del paciente, ni tan siquiera en período de incubación. Entre los posibles factores de riesgo para la aparición de dichas infecciones distinguiremos los asociados a los pacientes (factores de riesgo intrínsecos, como su edad o sexo) y los relacionados con la UCI (factores de riesgo extrínsecos, como tratamiento antibiótico). El número total de sujetos incluidos en el estudio fue de 61 pacientes (42 varones y 19 mujeres), con un total

de 40 infecciones nosocomiales. La variable dependiente (episodio de infección) consiste en el tiempo transcurrido entre el ingreso en la UCI hasta la aparición de la infección (originada por un microorganismo, bacteria u hongo), así como el tiempo entre el principio y el final de una infección. Obviamente los pacientes podrán presentar más de un episodio de infección. Sólo se analizaron infecciones nosocomiales originadas en la UCI, incluyéndose en la definición las infecciones comunitarias, las procedentes de otros hospitales y las procedentes de otras áreas hospitalarias. Para una explicación más detallada el lector interesado puede consultar en Barceló y Saez³.

El objetivo principal de este artículo es presentar una revisión de la aproximación marginal y condicional en el análisis de supervivencia multivariante. Adicionalmente, y como ilustración, se presenta una aplicación a datos reales que permite valorar la utilidad de tales aproximaciones.

Métodos

Con el propósito de solucionar los problemas ocasionados por la existencia de más de un tiempo de supervivencia por individuo, el modelo de Cox de riesgos proporcionales fue generalizado y situado en el contexto mucho más general de los modelos de intensidad multiplicativos para procesos contadores multivariantes, a través de la aproximación de Andersen y Gill (modelo AG)⁴⁻⁸. En la aproximación AG cada sujeto es tratado como un proceso contador con sucesos múltiples y con «incrementos» (es decir, los tiempos entre acontecimientos sucesivos) independientes, dada la «historia» de todas las variables observables hasta el tiempo de presentación de los sucesos. Es decir, con el modelo AG el riesgo de una determinada recurrencia para un individuo sigue la hipótesis de riesgos proporcionales y no está afectado por la presentación en dicho individuo de sucesos anteriores. Dicha hipótesis, llamada de «incrementos independientes», es la hipótesis clave del modelo AG. En caso de existir alguna interrelación entre sucesos se modelizaría mediante una o más variables explicativas dependientes del tiempo⁹.

Desde un punto de vista práctico, la aproximación AG representa a cada individuo por un conjunto de observaciones s_{ij} , t_{ij} , d_{ij} , x_{ij} , k_{ij} , $j = 1, \dots, n_i$ siendo $(s_{ij}, t_{ij}]$ el intervalo de riesgo, abierto a la izquierda y cerrado a la derecha; $d_{ij} = 1$ si el individuo tuvo un suceso en el instante t_{ij} y 0 en otro caso; x_{ij} es un vector de variables explicativas en el intervalo, y k_{ij} denota un estrato al que puede pertenecer el sujeto durante el intervalo. Dependiendo de la escala de tiempo, la primera observación puede o no empezar en cero.

Aclaremos lo anterior mediante un sencillo ejemplo. Sean dos individuos, un varón que padeció dos episodios de infección (INF) y una mujer que no presentó ninguno (tabla 1 y fig. 1a). Los individuos se dividen en distintos registros que se corresponden con diver-

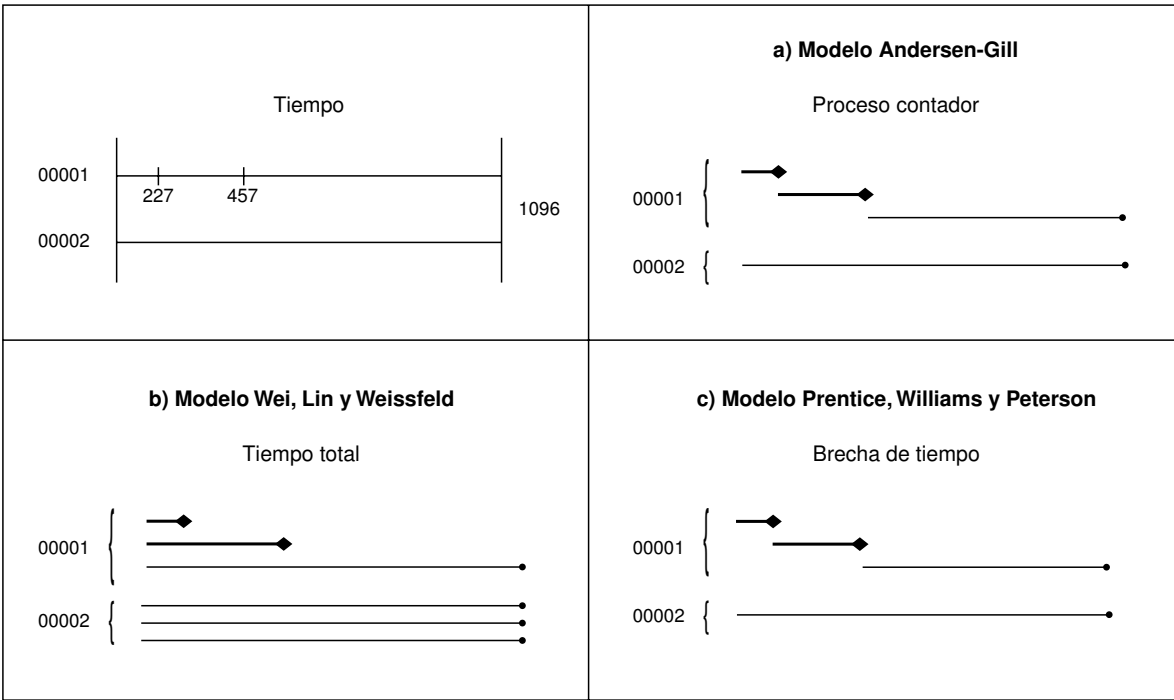
sas bandas temporales (intervalos de observación) y cada registro es tratado como si fuesen individuos diferentes. Las bandas se construyen en función de si el individuo presenta o no un episodio de infección y de posibles cambios en las variables explicativas, no ne-

Tabla 1. Estructura de los datos según las diferentes aproximaciones

Modelo	Individuo	T. entrada	T. salida	Estatus	Enum	Sexo	Ventilación mecánica
Andersen-Gill	00001	0	227	1		Varón	Sí
	00001	227	457	1		Varón	Sí
	00001	457	1.096	0		Varón	No
	00002	0	1.096	0		Mujer	Sí
Modelo de Wei, Lin y Weissfeld	00001		227	1	1	Varón	Sí
	00001		457	1	2	Varón	Sí
	00001		1.096	0	3	Varón	Sí
	00002		1.096	0	1	Mujer	No
	00002		1.096	0	2	Mujer	No
	00002		1.096	0	3	Mujer	No
Prentice, Williams y Peterson	00001	0	227	1	1	Varón	Sí
	00001	227	457	1	2	Varón	Sí
	00001	457	1.096	0	3	Varón	Sí
	00002	0	1.096	0	1	Mujer	No

T. entrada (s_i) y T. salida (t_i): días transcurridos desde el 1 de enero de 1994; estatus (d_i): 0, sin infección, 1, infección; enum: número de recurrencia.

Figura 1. Estructura de los datos a comparar según el modelo de análisis.



cesitando ser adyacentes. En este diseño las variables explicativas son fijas dentro de cada episodio, pero no así dentro de cada individuo (tal como suponen los modelos estándar, como el modelo de Cox de riesgos proporcionales). Nótese que es el riesgo de cada suceso concreto para un individuo determinado el que cumple la hipótesis usual de riesgos proporcionales y que el modelo supone que este riesgo no está afectado por sucesos anteriores del mismo individuo.

La aproximación AG es sencilla, pero sus hipótesis subyacentes son muy restrictivas y pueden llegar a ser irrealizables. En concreto, no es razonable suponer, tal como hemos visto, que las bandas pertenecientes a un mismo individuo no estén correlacionadas. En la práctica esta limitación puede conducir a estimadores sesgados e ineficientes. Además, puede sobrestimarse la precisión de los estimadores puesto que, habitualmente, las observaciones de un mismo individuo suelen estar correlacionadas positivamente.

Como posible solución, a finales de los años ochenta empezaron a aparecer los modelos marginales y condicionales en el análisis de datos de supervivencia multivariante. Sin embargo, la aplicación de estos modelos aun hoy día no está muy extendida debido, principalmente, a una falta de sistematización en la bibliografía especializada. Es por ello por lo que consideramos especialmente importante describir estos modelos no sólo desde un punto de vista teórico¹⁰, sino también ilustrarlos sobre la base de datos utilizada anteriormente.

En los modelos marginales la dependencia de las observaciones puede entenderse como un «estorbo», como algo que hay que eliminar o, en todo caso, cuyos efectos se deben controlar en el análisis, pero no como objetivo en sí. Por el contrario, los modelos condicionales consideran que la dependencia puede ser también de interés por sí misma. A grandes rasgos, los modelos marginales estiman el modelo ignorando la posible dependencia entre las observaciones, aunque corrigen posteriormente la varianza «ingenua» mediante estimadores *jackknife*, *bootstrap* o «sándwich». Los modelos condicionales estiman la dependencia (o la heterogeneidad) especificando explícitamente su distribución de probabilidad e incorporándola en el modelo.

Modelos marginales

La primera aproximación marginal que comentaremos es la estimación robusta de la matriz de varianzas y covarianzas. Se parte del hecho de que la aproximación AG proporciona estimadores consistentes de los parámetros asociados a las variables explicativas. Aun así, si existe dependencia, el método proporciona varianzas sesgadas. Se trataría, por tanto, de estimar de forma robusta la matriz de covarianzas, es decir, de

corregirla de forma adecuada utilizando, por ejemplo, un estimador *jackknife*⁹. La idea es sencilla, se trata de ir eliminando a un individuo de la muestra cada vez y reestimar el modelo resultante. Los estimadores finales se obtienen ponderando los obtenidos tras cada una de estas eliminaciones. El problema de esta aproximación marginal estriba en que se basa en la consistencia de las estimaciones. Sin embargo, las estimaciones de los parámetros serán consistentes siempre y cuando el modelo marginal esté correctamente especificado.

Otro modelo marginal es el modelo de Wei, Lin y Weissfeld (WLW)¹¹, en el que cada recurrencia del suceso (o tipo distinto de suceso) se modeliza como un estrato diferente, utilizándose los datos dentro de cada estrato de modo marginal. El resultado es que (todos) los individuos aparecen en todos los estratos. Es decir, el modelo WLW no es más que un modelo estratificado, siendo cada estrato una presentación del suceso de interés. Los estimadores finales son una media ponderada de los estimadores en cada estrato. Como vemos, se trata de un modelo marginal con respecto al conjunto de riesgo por cuanto el sujeto está en riesgo desde el inicio del estudio y, por tanto, no depende de sucesos previos.

Es decir, para cada una de las k recurrencias que pueda padecer un individuo i ($k = 1, 2, \dots, K$), la función de riesgo puede expresarse como:

$$\lambda_{ik}(t) = \lambda_{0k}(t) e^{\beta_k' X_{ik}(t)} \quad t \geq 0 \quad [1]$$

siendo $\lambda_{0k}(t)$ ($k = 1, \dots, K$) una función de riesgo subyacente no especificada, distinta para cada recurrencia, y $\beta_k = (\beta_{1k}, \dots, \beta_{pk})'$ un vector de orden $p \times 1$ de parámetros desconocidos de la regresión.

En el modelo WLW los datos se organizan de forma diferente de la que veíamos en el modelo AG (tabla 1 y fig. 1b, en la que se ha supuesto que el máximo número de episodios de infección entre todos los individuos analizados es de tres). Nótese que cada intervalo de tiempo empieza en cero. Se trata de un modelo marginal con respecto al conjunto de riesgo por cuanto el sujeto está en riesgo desde el inicio del estudio y, por tanto, no depende de sucesos previos.

Las limitaciones del modelo WLW pueden ser superadas, en parte, utilizando la aproximación de Prentice, Williams y Peterson (PWP)¹². Se trata de un modelo marginal respecto a la estimación de los parámetros, pero condicional en relación con la construcción del conjunto de individuos en riesgo. El modelo PWP permite que el riesgo basal varíe entre diferentes recurrencias. Se trata, por tanto, de un modelo de riesgos proporcionales con estratos dependientes del tiempo, en el que la dependencia entre recurrencias se controla estratificando por el número previo de presentaciones del suceso de interés. Así, a dife-

rencia del método WLW, en el que todos los individuos están en riesgo en todo momento, el modelo PWP incluye en el conjunto de riesgo para la recurrencia k únicamente a los individuos que han experimentado $k-1$ recurrencias.

Prentice, Williams y Peterson¹² proponen dos modelos semiparamétricos de riesgos proporcionales:

1. Tiempo desde el inicio del estudio (de tiempo total). Se incluye la función de riesgo subyacente como una función del tiempo desde el inicio del estudio

$$\lambda_{ik}(t) = \lambda_{0k}(t) e^{\beta'_{ik} X_{ik}(t)} \quad [2]$$

2. Tiempo desde la recurrencia inmediatamente precedente (de brecha de tiempo). La función de riesgo subyacente es una función del tiempo desde la recurrencia inmediatamente precedente al tiempo de fallo:

$$\lambda_{ik}(t) = \lambda_{0k}(t - t_{n(i)}) e^{\beta'_{ik} X_{ik}(t)} \quad [3]$$

La particularidad de ambos modelos reside en que se estratifica en función del número de recurrencias previas.

Siguiendo con el ejemplo de antes, tanto para el modelo AG como para PWP los datos se organizarían tal como se muestra en la tabla 1 y la figura 1c.

Modelos condicionales

En algunos casos no sólo nos interesa el efecto marginal de las covariables sobre la función de riesgo de los tiempos de fallo de los individuos, sino también cómo los sucesos anteriores pueden influir en el riesgo de desarrollar futuros fallos. Para responder a estas cuestiones necesitamos recurrir a los modelos condicionales.

En primer lugar, la hipótesis de incrementos independientes en el modelo AG puede relajarse incluyendo en el modelo variables dependientes del tiempo, como el número de recurrencias previas, a fin de capturar la posible estructura de dependencia entre las recurrencias. En concreto, se podría trabajar con el modelo autorregresivo o markoviano. Este modelo será plenamente apropiado cuando la dependencia entre observaciones sea, principalmente, «serial». Sin embargo, y aunque es un modelo muy general, supone que la dependencia entre recurrencias decae exponencialmente y que la dependencia es idéntica para todos los individuos¹⁰.

Parte de las limitaciones anteriores pueden solucionarse mediante los modelos de fragilidad. Se trata de introducir una (modelos de fragilidad univariante) o más (modelos de fragilidad multivariante) variables aleatorias comunes a todos los individuos con el objetivo de permitir que los tiempos de supervivencia multiva-

riantes, es decir, las observaciones repetidas de un mismo individuo, estén correlacionadas. Se trata, en definitiva, de que este efecto multiplicativo en la función de intensidad (o de riesgo) capture dos fuentes de variación distintas pero relacionadas¹³, con el propósito de modelar la dependencia entre los diferentes tiempos de supervivencia entre recurrencias. Una de las fuentes de variación surge de covariables comunes a todos los individuos, no observadas, las cuales, cuando se omiten, generan dependencia entre sucesos. La otra fuente de variación proviene de covariables individuales no observadas y, por tanto, no incluidas en el estudio, bien por circunstancias prácticas, bien porque no se asocian a ningún factor de riesgo de la supervivencia del individuo. Los efectos de esta fuente serían totalmente equivalentes a la dependencia serial citada en la introducción de este artículo.

Posibles alternativas para la distribución de probabilidad de la variable aleatoria fragilidad son la distribución gamma, la distribución positiva estable y la distribución gaussiana inversa, entre otras. El modelo de fragilidad gamma es el más adecuado en nuestro caso, debido, en primer lugar, al tipo de datos de que disponemos. En concreto, bajo este tipo de distribución la fuerza de la asociación entre recurrencias es estacionaria, es decir, no depende del tiempo. Este tipo de asociación es la que más se produce en el mundo real. La ventaja más importante de la distribución gamma, sin embargo, se encuentra en la facilidad analítica que permite.

En particular, el modelo AG de fragilidad^{14,15} puede expresarse de la siguiente manera en su caso más simple:

$$\lambda_i(t) = \omega_i Y_i(t) \alpha_0(t) e^{\beta' X_i(t)} \quad [4]$$

siendo $N = (N_i; i = 1, \dots, n)$ un proceso contador multivariante que cuenta el número de sucesos que han ocurrido en cada uno de los n individuos en el tiempo t . Asociado a este proceso contador, existe un proceso de intensidad λ con componentes λ_i . El proceso Y con componentes Y_i es un proceso no negativo, observable y predecible, es decir, su valor en cualquier momento t es conocido justo antes de t , indicando si un individuo i se observa en el conjunto a riesgo de experimentar un determinado suceso en el tiempo t ; α_0 denota ahora una función de riesgo basal desconocida, y ω_i (fragilidad) es un conjunto de variables aleatorias independientes e idénticamente distribuidas según una distribución de probabilidad paramétrica determinada, en este caso la distribución gamma, con media nula y varianza común. La variabilidad (medida por su varianza) de la fragilidad determina el grado de heterogeneidad. De hecho, cuando la varianza es igual a cero el modelo se reduce al modelo AG estándar.

El modelo descrito es un modelo muy simple, ya que asume que la fragilidad es constante a lo largo del tiem-

po, es decir, la fragilidad reflejaría en este caso diferencias individuales presentes al comienzo del estudio. Asimismo, este modelo supone que la fragilidad actúa proporcionalmente sobre la función basal común a todos los individuos. Al imponer esto se asume que todos los individuos siguen un patrón uniforme.

El problema es que la estimación de los parámetros del modelo utilizando directamente la máxima verosimilitud no es posible cuando la función de riesgo basal no está parametrizada, como en nuestro caso. Por otra parte, los modelos de fragilidad univariantes presentan una serie de limitaciones; entre ellas cabe destacar las siguientes: suponen que la dependencia entre observaciones es siempre positiva, no permiten la interacción entre covariables observables y la fragilidad, y más importante aún, sólo permiten la presencia de un único efecto aleatorio, siendo además éste constante a lo largo del tiempo. Con el propósito de solucionar estos problemas, Barceló y Saez¹⁶ proponen un método de especificación y estimación original basado en la verosimilitud extendida de Nelder y Lee¹⁷, así como en el método de las variables instrumentales, y estudian las ventajas analíticas de dicho método, en particular, en relación con alternativas de especificación (modelos de fragilidad multivariantes) y de estimación (algoritmo EM). De una manera muy resumida, introducen un nuevo paso en el algoritmo EM, modelizando la varianza de la fragilidad directamente. Mediante el algoritmo así modificado (algoritmo EMB), se estima no sólo el modelo de supervivencia, sino también la varianza de la fragilidad de forma simultánea. Mediante esta modificación solucionan parte de las limitaciones citadas anteriormente; en particular, permiten que la fragilidad no sea constante entre individuos. Una explicación más detallada puede encontrarse en Barceló y Saez¹⁶.

Una alternativa a los modelos de fragilidad la proporcionan los modelos de Cox penalizados¹⁸. En resumen, se trata de maximizar una función de verosimilitud parcial penalizada. El término penalizador captura la variabilidad local presente en la densidad conjunta de los datos.

Los modelos que se han comentado no están anidados, por lo que las usuales medidas de bondad de ajuste no pueden utilizarse. Como alternativa se suele emplear el criterio de información de Akaike (AIC).

El software estadístico que permite ajustar todos los modelos vistos y que, por otra parte, se ha utilizado en la aplicación práctica, es el S-PLUS 2000¹⁹. En concreto, se debe utilizar la instrucción *coxph*, que permite considerar diferentes tipos de truncamiento y censura. La estimación robusta de la varianza se obtiene mediante la instrucción *cluster*. Desafortunadamente, hay pocos programas que permitan la estimación de modelos marginales y/o condicionales en el contexto multivariante. Además del S-PLUS 2000, el lector interesado puede recurrir a otros programas como R (versión libre y gra-

tuita de S-PLUS), SAS (versión 6 o posteriores, instrucciones *proc phreg*, aunque no permite la estimación robusta de la varianza) y, únicamente para el modelo AG estándar, STATA (versión 5 o posteriores, instrucciones *stset* y *stcox*).

Resultados

Los resultados de la estimación del modelo AG estándar aparecen en las primeras columnas de la tabla 2. En esta tabla se presentan las tasas de riesgo asociadas a toda una serie de variables explicativas y sus respectivos intervalos de confianza (IC). Por último, se ofrece una serie de medidas de bondad de ajuste del modelo. Como puede apreciarse al observar la tabla, pocas variables resultaron estar asociadas de forma significativa con el tiempo de presentación de una infección nosocomial. En concreto, ser mujer (0,258; IC del 95%, 0,067-0,997), haberse sometido a cirugía urgente (0,295; IC del 95%, 0,092-0,949), haber permanecido intubado más de 6 días mediante traqueotomía (0,182; IC del 95%, 0,041-0,814) y haber recibido tratamiento antibiótico durante más de 7 días (0,179; IC del 95%, 0,034-0,947) reducían el riesgo de tales infecciones.

Los resultados de la aplicación práctica del modelo AG con estimación robusta de la matriz de varianzas y covarianzas serían los que hemos visto anteriormente al hablar del modelo AG, ya que esta primera aproximación marginal sólo afecta a la matriz de varianzas y covarianzas, no a la construcción de los conjuntos a riesgo. Si nos fijamos en la tabla 2 y comparamos el modelo AG estándar y el modelo AG robusto, apreciamos una mayor precisión en los intervalos de confianza proporcionados por este último modelo, siendo este cambio muy significativo en variables como ventilación mecánica. Por el contrario, otras variables tales como cirugía urgente (0,295; IC del 95%, 0,083-1,048) y tratamiento antibiótico durante más de 7 días (0,179; IC del 95%, 0,031-1,047) llegan incluso a perder su significación estadística.

En relación con los resultados del modelo PWP con estimación robusta de la varianza (tabla 3), con una importante excepción, presencia de catéter venoso central durante más de tres días (1,065; IC del 95%, 1,007-1,775), el signo de los estimadores de los parámetros coincide en los tres modelos vistos (modelo AG estándar, modelo AG robusto y modelo PWP), y no ocurre lo mismo con la magnitud ni con la significación de los parámetros. Nótese ahora que variables como la edad (1,009; IC del 95%, 0,985-1,033), la existencia previa de infecciones comunitarias (4,56; IC del 95%, 1,435-14,477), estar inestable en UCI (54,47; IC del 95%, 0,786-3.776,406), estar inestable en coma o shock (219,97; IC del 95%, 3,187-15.182,609), la presencia

Tabla 2. Resultados de la estimación del modelo Andersen-Gill (AG) estándar y del modelo AG con estimación robusta de la matriz de variancias y covariancias

	AG estándar			AG robusto		
	Tasa de riesgo	IC del 95%		Tasa de riesgo	IC del 95%	
Factores de riesgo intrínsecos						
Género (varón)	0,258* *	0,067	0,997	0,258* *	0,075	0,882
Edad (años)	1,006	0,979	1,030	1,006	0,984	1,028
Infecciones previas (no)						
Comunitarias	2,106	0,544	8,150	2,106	0,693	6,394
Otras infecciones	1,168	0,299	4,560	1,168	0,408	3,345
CDC (estable)						
Inestable en UCI	11,406	5,00e-034	2,60e + 035	11,406	0,243	534,508
Inestable en coma o shock	35,952	1,60e-033	8,10e + 035	35,952* *	0,943	1.370,330
Cirugía urgente (no)	0,295* *	0,092	0,949	0,295*	0,083	1,048
Factores de riesgo extrínsecos						
Localización (resto de camas)						
Cama 4	1,892	0,358	10,000	1,892	0,578	6,193
Cama 5, 10, 11	0,779	0,238	2,560	0,779	0,356	1,708
Ventilación mecánica (no)						
≤ 3 días	0,013	5,13e-066	3,45e + 061	0,013* *	0,001	0,556
4-10 días	85,201	7,64e-064	9,50e + 066	85,201* *	1,726	4.205,852
>10 días	72,932	6,59e-064	8,08e + 066	72,932* *	1,524	3.490,499
Catéter venoso central (≤ 3 días)	0,261	2,50e-066	2,72e + 064	0,261	0,012	5,784
Catéter arterial (no)						
> 0 días	1,268	0,389	4,140	1,268	0,404	3,978
Traqueotomía (no)						
≤ 6 días	0,402	0,100	1,610	0,402	0,126	1,278
> 6 días	0,182* *	0,041	0,814	0,182* *	0,038	0,858
Sonda urinaria (no)						
≤ 4 días	0,677	0,009	52,200	0,678	0,025	18,031
5-12 días	1,552	0,034	71,800	1,552	0,219	10,991
> 12 días	0,822	0,017	38,700	0,822	0,064	10,594
Sonda nasogástrica (no)						
≤ 9 días	1,860	0,497	6,960	1,860	0,429	8,068
> 9 días	0,503	0,157	1,620	0,503	0,199	1,270
Tratamiento antibiótico (no)						
≤ 7 días	0,549	0,155	1,940	0,549	0,185	1,629
> 7 días	0,179* *	0,034	0,947	0,179*	0,031	1,047
Dosis de antibiótico-DDD	1,042	0,978	1,110	1,042	0,938	1,157
Deviance (grados de libertad)		204,45 (76)			205,45 (76)	
AIC		254,45			254,45	

Categorías de referencia entre paréntesis. * < 0,10; ** < 0,05.

IC: intervalo de confianza; UCI: unidad de cuidados intensivos; AIC: criterio de información de Akaike; DDD: dosis diaria definida; CDC: Center for Disease Control.

de ventilación mecánica de cuatro a 10 días (61,93; IC del 95%, 1,325-2.894,463), la presencia de ventilación mecánica durante más de 10 días (129,53; IC del 95%, 4,005-4.189,427) y la presencia de catéter venoso central durante más de tres días (1,07; IC del 95%, 1,007-1,775) aumentan el riesgo de padecer un episodio de infección nosocomial. Contrariamente, la presencia de sondas nasogástricas durante más de 9 días (0,13; IC del 95%, 0,023-0,748) y el tratamiento antibiótico durante más de 7 días (0,04; IC del 95%, 0,004-0,385) lo reducen. Variables tales como género y traqueotomía

pierden en este modelo su significación. Resumiendo, no sólo los factores de riesgo intrínsecos (como la gravedad de la enfermedad), sino también los extrínsecos (procedimientos terapéuticos agresivos utilizados en la UCI) llevaron a un mayor riesgo de infección nosocomial durante la estancia del paciente en la UCI.

Los resultados del modelo AG de fragilidad gamma estimado mediante el algoritmo EMB, así como los del modelo de Cox penalizado, también se encuentran en la tabla 3. En el modelo AG de fragilidad gamma estimado mediante el algoritmo EMB, observamos que

Barceló MA. Modelos marginales y condicionales en el análisis de supervivencia multivariante

Tabla 3. Resultados de la estimación del modelo de Prentice, Williams y Peterson (PWP) del modelo de Cox penalizado del modelo EMB. PWP y EMB con estimación robusta de la varianza

	PWP			Modelo Cox penalizado			EMB		
	Tasa de riesgo	IC del 95%		Tasa de riesgo	IC del 95%		Tasa de riesgo	IC del 95%	
Factores de riesgo intrínsecos									
Género (varón)	0,631	0,222	1,788	0,631	0,133	2,999	0,578	0,221	1,512
Edad (años)	1,009*	0,985	1,033	1,009*	0,981	1,037	1,009*	0,990	1,028
Infecciones previas (no)									
Comunitarias	4,560**	1,435	14,477	4,558*	0,954	21,772	6,923**	2,578	18,594
Otras infecciones	0,560	0,201	1,557	0,560	0,118	2,648	0,418*	0,162	1,077
CDC (estable)									
Inestable en UCI	54,468*	0,786	3.776,406	8.025,270**	1.740	36.968,770	49,012**	1,111	2.162,135
Inestable en coma o shock	219,967**	3,187	15.182,609	NA	NA	NA	257,040**	5,997	11.017,03
Cirugía urgente (no)	0,668	0,258	1,731	0,668	0,221	2,021	0,579	0,268	1,251
Factores de riesgo extrínsecos									
Localización (resto de camas)									
Cama 4	2,255	0,545	9,340	2,255	0,346	14,707	2,247	0,621	8,135
Cama 5, 10, 11	0,946	0,381	2,348	0,946	0,257	3,478	0,784	0,369	1,667
Ventilación mecánica (no)									
≤ 3 días	0,011*	0,000	1,175	NA	NA	NA	0,006**	0,000	0,314
4-10 días	61,929**	1,325	2.894,463	64,406**	1,35	3.065,471	59,981**	2,105	1.709,126
>10 días	129,534**	4,005	4.189,427	134,714**	3,06	5.934,619	98,713**	4,260	2.287,525
Catéter venoso central (≤ 3 días)	1,065*	1,007	1,775	NA	NA	NA	1,082*	1,011	1,822
Catéter arterial (no)									
> 0 días	1,833	0,579	5,805	1,833	0,495	6,782	1,629	0,638	4,162
Traqueotomía (no)									
≤ 6 días	0,231	0,028	1,924	0,231	0,030	1,768	0,243*	0,044	1,355
> 6 días	0,351	0,072	1,705	0,351	0,057	2,171	0,303*	0,075	1,224
Sonda urinaria (no)									
≤ 4 días	2,453	0,018	334,750	2,453	0,024	251,814	3,922	0,066	231,811
5-12 días	6,355	0,306	132,032	6,355	0,137	294,936	7,942*	0,630	100,190
> 12 días	1,566	0,079	31,014	1,566	0,035	70,518	1,998	0,159	25,104
Sonda nasogástrica (no)									
≤ 9 días	1,066	0,310	3,660	1,066	0,285	3,990	1,063	0,377	3,000
> 9 días	0,133**	0,023	0,748	0,133**	0,027	0,649	0,129**	0,030	0,550
Tratamiento antibiótico (no)									
≤ 7 días	0,415	0,112	1,539	0,415	0,093	1,858	0,329**	0,122	0,883
> 7 días	0,038**	0,004	0,385	0,038**	0,005	0,327	0,028**	0,004	0,216
Dosis de antibiótico-DDD	1,035	0,948	1,129	1,035	0,972	1,101	1,037	0,982	1,096
Deviance (grados de libertad)		130,43 (76)			130,42 (76)			112,41(76)	
AIC		180,43			180,42			162,41	
Varianza fragilidad					5e-007			1,125	
Modelo para la varianza									
Ventilación mecánica (no)							β	s.e. (β)	
≤ 3 días	—	—	—	—	—	—	0,051	0,036	
4-10 días	—	—	—	—	—	—	0,056	0,036	
>10 días	—	—	—	—	—	—	0,086	0,032	
Traqueotomía (no)									
≤ 6 días	—	—	—	—	—	—	0,008	0,026	
> 6 días	—	—	—	—	—	—	-0,006	0,023	
Sonda urinaria (no)									
≤ 4 días	—	—	—	—	—	—	0,038	0,045	
5-12 días	—	—	—	—	—	—	-0,011	0,042	
> 12 días	—	—	—	—	—	—	-0,019	0,046	
Sonda nasogástrica (no)									
≤ 9 días	—	—	—	—	—	—	-0,001	0,021	
> 9 días	—	—	—	—	—	—	-0,006	0,021	

Categorías de referencia entre paréntesis. * < 0,10; ** < 0,05. IC: intervalo de confianza; UCI: unidad de cuidados intensivos; NA: no aplicable; AIC: criterio de información de Akaike; DDD: dosis diaria definida; CDC: Center for Disease Control.

variables tales como la edad (1,009; IC del 95%, 0,990-1,028), existencia previa de infecciones comunitarias (6,923; IC del 95%, 2,578-18,594), estar inestable en UCI (49,012; IC del 95%, 1,111-2.162,135), estar inestable en coma o shock (257,040; IC del 95%, 5,997-11.017,03), la presencia de ventilación mecánica en todas sus categorías y la presencia de catéter venoso central durante más de tres días (1,082; IC del 95%, 1,011-1,822) continúan aumentando el riesgo de padecer infección nosocomial de forma estadísticamente significativa. Por el contrario, la presencia de sonda nasogástrica durante más de 9 días (0,129; IC del 95%, 0,03-0,55) y el tratamiento antibiótico en todas sus categorías lo reducen, también de forma estadísticamente significativa. Por otra parte, se observa que variables que no eran significativas en el modelo PWP, como la existencia previa de otras infecciones, traqueotomía y presencia de sonda urinaria de 5 a 12 días, pasan a serlo ahora. Estos resultados cambian un poco al observar el modelo de Cox penalizado. Mientras que variables tales como la edad (1,009; IC del 90%: 0,981-1,037), infecciones previas comunitarias (4,558; IC del 90%, 0,954-21,772), estar inestable en UCI (8,025,27; IC del 95%, 1.740-36.968,77), presencia de ventilación mecánica de cuatro a 10 días (64,406; IC del 95%, 1,35-3.065,471), presencia de ventilación mecánica durante más de 10 días (134,714; IC del 95%, 3,06-5.934,619), presencia de sonda nasogástrica durante más de 9 días (0,133; IC del 95%, 0,027-0,649) y tratamiento antibiótico durante más de 7 días (0,038; IC del 95%, 0,005-0,327) continúan manteniendo su significación, otras como traqueotomía, existencia de otras infecciones previas, presencia de sonda urinaria de 5 a 12 días y tratamiento antibiótico durante 7 días o menos la pierden. Por otra parte, el modelo de Cox penalizado parece que es incapaz de estimar los parámetros en algunas de las variables explicativas presentes en el modelo (inestable en coma o shock, ventilación mecánica inferior o igual a tres días y catéter venoso central), lo cual no ocurre en el modelo EMB. Asimismo, el modelo EMB es el que proporciona intervalos de confianza para los estimadores de los parámetros más reducidos.

Adicionalmente, y a diferencia de los demás modelos, el modelo EMB permite observar los factores explicativos de la fragilidad. Como se ve, la varianza de la fragilidad depende únicamente de factores extrínsecos tales como ventilación mecánica, traqueotomía, sonda urinaria y sonda nasogástrica.

Discusión

Como ya se ha dicho, la problemática planteada por la existencia de datos de supervivencia multivariantes

hace necesaria la generalización del modelo de Cox de riesgos proporcionales, concretándose ésta en el modelo AG. Este modelo permite trabajar con datos multivariantes; sin embargo, no contempla la posibilidad que las observaciones pertenecientes a un mismo individuo estén correlacionadas, ni la presencia de heterogeneidad individual. Con el propósito de solucionar esto, citamos los principales modelos marginales (modelo AG con estimación robusta de la varianza, modelo WLW y modelo PWP) y condicionales (modelo autorregresivo o markoviano, modelos de fragilidad y modelo de Cox penalizado).

Aun cuando los modelos marginales son capaces de corregir los principales efectos de la dependencia entre observaciones (siempre suponiendo que las estimaciones de los parámetros son consistentes), son incapaces de distinguir cómo los sucesos anteriores pueden influir en el riesgo de desarrollar futuros fallos. Esto último es tratado mediante los modelos condicionales.

El método WLW presenta importantes desventajas. Cook y Lawless²⁰ señalan que el modelo WLW es válido únicamente bajo la posiblemente muy restrictiva hipótesis de censura independiente (restrictiva como consecuencia de las recurrencias). La desventaja más importante es que el método WLW supone que en el instante t todos los individuos (observados en t) constituyen el conjunto de riesgo para la correspondiente recurrencia, con independencia del número de recurrencias previas que tuviese el individuo^{20,21}. Esto no representaría ningún problema si se tratase de diferentes sucesos de un tipo totalmente distinto, pero si se trata de recurrencias del mismo suceso podría ocurrir que algún hecho emprendido con el fin de reducir el número de tales recurrencias redujese el número de las mismas a partir de su implementación (lo mismo ocurriría en el caso de *harvesting*). Como consecuencia aumentaría la correlación entre los estimadores de los parámetros. Otro problema es que las estimaciones de los parámetros serán consistentes siempre y cuando la función de riesgo esté correctamente especificada.

La principal desventaja del modelo PWP radica en la manera de construir los conjuntos a riesgo, ya que lleva a una «pérdida de aleatorización»²², por cuanto los individuos con más riesgo van saliendo del estudio. Esta pérdida será mayor conforme vayamos moviéndonos a estratos superiores (y cuantos más estratos se consideren). En estos casos, la consistencia de las estimaciones depende de la inclusión de todas las variables relevantes. Por otra parte, tanto el modelo AG como el modelo PWP son sensibles a los errores de especificación de la estructura de dependencia entre los tiempos de recurrencia.

Aunque los métodos marginales cumplen (marginamente) la hipótesis de riesgos proporcionales, suponen explícita (PWP) o implícitamente (WLW) que la función de riesgo basal varía de una recurrencia a otra²².

Como consecuencia, son extremadamente sensibles al cumplimiento (en realidad) de la hipótesis de riesgos proporcionales⁹. Finalmente, y respecto a las aproximaciones condicionales, los métodos marginales no son totalmente eficientes^{23,24}. La ineficiencia aumentará de forma proporcional al número de estratos (recurrencias) e inversamente proporcional al número de observaciones en cada estrato. En analogía a los modelos lineales es razonable suponer que la ineficiencia también aumentará conforme lo haga la correlación entre recurrencias. Es por tanto de esperar que WLW sea menos eficiente que PWP.

Por lo que hace referencia a los modelos condicionales, sólo comentaremos aquí el problema en los modelos de Cox penalizados. Como argumentábamos anteriormente, la variabilidad que pretenden capturar estos modelos recoge, en realidad, dos fenómenos diferentes, fragilidad y correlación serial. Por otra parte, presentan el inconveniente de que el parámetro de suavizado debe escogerse utilizando técnicas no del todo validadas en el campo de supervivencia.

Situándonos en nuestro ejemplo práctico, de entre los tres modelos marginales presentados en el artículo, el modelo PWP es posiblemente el mejor debido al tipo de datos de que disponemos. El hecho de que, tras experimentar una primera infección, aumente el riesgo

de aparición de nuevas infecciones hace que el modelo PWP con estimación robusta de la varianza sea especialmente apropiado, ya que sería posible estratificar los datos para cada episodio de infección.

Por lo que respecta a las estimaciones de los modelos condicionales mencionados en el artículo, el algoritmo EMB de estimación de un modelo AG de fragilidad gamma es el que presenta mejores resultados. Por un lado, es el más eficiente al presentar intervalos de confianza más reducidos, debido quizá a que el algoritmo EMB es capaz de capturar no sólo la heterogeneidad individual, sino también la dependencia serial. Por otro lado, de entre todos los modelos estudiados, es el que proporciona un mejor ajuste en términos del criterio de Akaike. Finalmente, es el único que permite estimar los efectos explicativos de la fragilidad.

Agradecimientos

Agradezco al Dr. Marc Saez su gran ayuda en la elaboración de este trabajo. Agradezco, asimismo, al Dr. Santiago Pérez Hoyos y a dos evaluadores anónimos sus sugerencias. Este trabajo fue parcialmente financiado por el proyecto 07/50/1998 de la Agència d'Avaluació de Tecnologia Mèdica (AATM) del Servei Català de la Salut, Generalitat de

Bibliografía

1. Cox DR. Regression models and life tables. *JRSS, Series B*, 1972;34:187-220.
2. McCall BP. Testing the proportional hazards assumption in the presence of unmeasured heterogeneity. *J Applied Econometrics* 1994;9:321-34.
3. Barceló MA, Saez M. Factors determining the incidence of nosocomial infections in an Intensive Care Unit of a public hospital in Girona, Spain. Mimeo, 2001.
4. Aalen OO. Non-parametric inference for a family of counting processes. *Ann Stat* 1978;6:701-26.
5. Andersen PK, Gill RD. Cox's regression model for counting processes: a large sample study. *Ann Stat* 1982;10:1100-20.
6. Andersen PK, Borgan O. Counting process models for life history data: a review. *Scan J Stat* 1985;12:97-158.
7. Fleming TR, Harrington DP. Counting processes and survival analysis. New York: John Wiley & Sons, Inc., 1991.
8. Andersen PK, Borgan O, Gill RD, Keiding N. Statistical models based on counting processes. New York: Springer-Verlag, 1993.
9. Therneau TM, Hamilton SA. rhDNase as an example of recurrent event analysis. *Stat Med* 1997;16:2029-47.
10. Barceló MA. Modelos de fragilidad en el análisis multivariante de supervivencia [tesis doctoral]. Girona: Universitat de Girona, 2001.
11. Wei LJ, Lin DY, Weissfeld L. Regression analysis of multivariate incomplete failure time data by modeling marginal distributions. *JASA* 1989;84:1065-73.
12. Prentice RL, Williams BJ, Peterson AL. On the regression analysis of multivariate failure time data. *Biometrika* 1981;68:373-9.
13. Petersen JH. An additive frailty model for correlated life times. *Biometrics* 1998;54:646-61.
14. Nielsen G, Gill R, Andersen P, Sorensen T. A counting process approach to maximum likelihood estimation in frailty models. *Scan J Stat* 1992;19:25-43.
15. Nielsen G, Tabanera-Palacios R. GNFCN: a computer program for gamma frailty analysis. 1993; p. 1-21.
16. Barceló MA, Sáez M. A modification of the EM algorithm to estimate an Andersen-Gill gamma frailty model for multivariate failure time data. Mimeo, 2001.
17. Nelder JA, Lee Y. Generalized linear models for the analysis of Taguchi-type experiments. *Applied Stochastic Models and Data Analysis* 1991;7:107-20.
18. Therneau TM, Grambsch PM. Modeling survival data. Extending the Cox model. Statistics for biology and health. New York: Springer, 2000.
19. S-PLUS 2000 User's Guide. Seattle: Data Analysis Products Division, MathSoft, 1999.
20. Cook RJ, Lawless JF. Discussion of paper by Wei and Glidden. *Stat Med* 1997;16:841-3.
21. Lipschutz KH, Snapinn SM. Discussion of paper by Wei and Glidden. *Stat Med* 1997;16:846-8.
22. Clayton D. Some approaches to the analysis of recurrent event data. *Stat Med* 1994;3:244-63.
23. Oakes D. Frailty models for multiple event times. En: Klein YP, Goel PK, editors. *Survival analysis: state of the art*. Boston: Kluwer Academic Publishers, 1992; p. 415.
24. Oakes D. Discussion of paper by Wei and Glidden. *Stat Med* 1997;16:843.