Revista Colombiana de Estadística
Volumen 27 N^o 2. Págs. 123 a 137. Diciembre 2004

Un frailty gamma en supervivencia bivariante

Esteban Navarrete Álvarez*

Resumen

Consideramos un modelo de supervivencia bivariante basado en la presencia de una covariable aleatoria no observada, frailty, que induce dependencia entre los tiempos de fallo y actúa bajo hipótesis de riesgo proporcional. Consideramos una distribución gamma para el frailty y distribuciones marginales Weibull e introducimos covariables explicativas, que actúan bajo hipótesis de vida acelerada. Introducimos además covariables en el parámetro de asociación.

Palabras Clave: *Supervivencia bivariante, covariación, Weibull, estimación.*

Abstract

We consider a bivariate survival model with the presence of an unobserved random covariate, frailty, that induces dependency between the failure times and acts according to a proportional hazard model. We consider gamma distribution for this frailty and Weibull marginal distributions and we introduce explanatory covariates that act according to an accelerated life model. We consider covariates in asociation parameter.

Key words: *Bivariate survival, Frailty, Proportional hazard, Accelerated life, Covariates, Weibull, Estimation.*

*Departamento de Estadística e Investigación Operativa. Universidad de Granada. Facultad de Ciencias, Campus de Fuentenueva, s/n. 18071 Granada, España. Fax: 958243267, Tel: 958243156, E-Mail: estebang@ugr.es

1. Introducción

Uno de los procedimientos más importantes para la construcción de modelos de supervivencia multivariante se basa en la idea debida a Vaupel, Manton & Stallard (1979) para explicar la heterogeneidad de la población en modelos univariantes. Rápidamente esta idea fué extendida a supervivencia bivariante. La idea en la que se basa este método es la siguiente: la asociación posible entre los tiempos de supervivencia, T_1 y T_2, se explica a través de una característica común que poseen los individuos de una pareja. Esta característica común induce cierta dependencia entre los tiempos. La heterogeneidad se modeliza por un factor de riesgo, una covariable aleatoria no observable, denominada *frailty*. Para cada individuo de una pareja concreta, el frailty toma el mismo valor y distinto entre individuos de parejas distintas. Si se denota por Z la variable aleatoria frailty, se considera que T_1 y T_2 son condicionalmente independientes, es decir, son independientes para cada valor fijo de Z. En cierto sentido, como se verá a continuación, se podría considerar el frailty como una covariable aleatoria no observada y se suele considerar que actúa multiplicativamente en la función de riesgo, es decir, bajo la hipótesis de riesgo proporcional.

La idea intuitiva de este efecto aleatorio es la siguiente. La asociación entre los tiempos de supervivencia proviene de un factor desconocido denominado frailty y no porque la incidencia de T_1 influya directamente en el valor de T_2; es decir, por ejemplo, que un padre se muera de infarto no influye de una forma directa en que el hijo lo vaya a hacer pero sí existe un factor que poseen ambos, por ejemplo, un factor genético, que liga al padre con el hijo.

Vamos a considerar el caso bivariante (dos tiempos de fallo) e hipótesis de riesgo proporcional, es decir, cada función de riesgo marginal (univariante) se comporta, en cuanto al frailty, como un modelo de riesgo proporcional. En lo que sigue consideraremos que $E(Z) = 1$ lo cual no resta generalidad. Si no fuera así, en lugar de $h_i(t_i)$ habría de tomarse un riesgo base o inicial $h_{i0}(t_i)$. Es decir:

$$h_i(t_i/Z) = Zh_i(t_i) \ , \qquad i : 1, 2,$$

o, lo que es lo mismo, la función de supervivencia es:

$$S_i\big(t_i/Z\big) = \exp\Big[-z\int_0^{tsub} ih_i(t_i)dt_i\Big] = \exp\Big[-Z\Lambda_i(t_i)\Big] = \Big[S_i(t_i)\Big]^Z,$$

donde $\Lambda_i(t_i)$ es la función de riesgo acumulativa marginal correspondiente a la variable T_i.

En Hougaard (1986a) se demuestra que si el frailty actúa multiplicativamente en las componentes del vector función de riesgo, así lo hace también en las componentes del vector función de riesgo marginal aunque con distinta constante de proporcionalidad.

La función de supervivencia condicionada es:

$$\begin{aligned} S(t_1,\ t_2/Z) &= P(T_1 \geq t_1,\ T_2 \geq t_2/Z) \\ &= P(T_1 \geq t_1/Z)\ P(T_2 \geq t_2/Z) = \big[S_1(t_1)\big]^Z \big[S_2(t_2)\big]^Z \\ &= \big[S_1(t_1)S_2(t_2)\big] \sup Z = \exp\big[-Z\Lambda_1(t_1)\big] \exp\big[-Z\Lambda_2(t_2)\big] \\ &= \exp\big[-Z\big(\Lambda_1(t_1) + \Lambda_2(t_2)\big)\big], \end{aligned}$$

y realizando la mixtura respecto a la distribución del frailty Z, se verificará que la función de supervivencia conjunta es:

$$S(t_1,\ t_2) = \int S(t_1,\ t_2/Z) dF(z) = E\big[\exp\big(-Z(\Lambda_1(t_1) + \Lambda_2(t_2))\big)\big].$$

El tratamiento del modelo dependerá, por la expresión anterior, de la distribución del frailty Z y, sobre todo, de la resolubilidad de la integral anterior. Ésta no es fácil de resolver en general, pero veremos más adelante un método para solucionar este problema.

La función de supervivencia se expresa como:

$$S(t_1,\ t_2) = \int \big[S_1(t_1)S_2(t_2)\big]^Z dF(z).$$

Sea u una variable definida de la siguiente forma:

$$u = -\log\big[S_1(t_1)S_2(t_2)\big] = -\log S_1(t_1) - \log S_2(t_2) = \Lambda_1(t_1) + \Lambda_2(t_2),$$

y sea la transformada de Laplace de Z:

$$\phi(x) = E\big[\exp(-xZ)\big] = \int \exp(-xZ) dF(z),$$

con lo que resulta que:

$$\phi(u) = \int \exp(-uZ) dF(z) = S(t_1,\ t_2).$$

Así pues, el problema está en obtener la transformada de Laplace de la distribución del frailty.

En el caso de que el frailty siga una distribución gamma se tiene que:

$$S(t_1,\ t_2) = \left[\frac{1}{[S_1(t_1)]^{\theta-1}} + \frac{1}{[S_2(t_2)]^{\theta-1}} - 1\right]^{-\dfrac{1}{\theta-1}},$$

donde la constante θ es un parámetro de dicha distribución.

Existen ciertas variables explicativas o covariables que caracterizan a los distintos individuos, es decir, son características propias del individuo, factores exógenos, tratamientos experimentales aplicados, etc. Ejemplos de estas características son la edad, el sexo, clase social, etc. Si se tienen dos tiempos de supervivencia que miden, por ejemplo, la aparición de una enfermedad cardiovascular y la aparición de hipertensión, respectivamente, características asociadas a estos individuos que pueden afectar a los tiempos de fallo podrían ser la edad, hábitos de fumar, sexo, una medicación específica, índice de colesterol, etc. Un ejemplo ciertamente interesante puede verse en Wassel & Moeschberger (1993). Una clasificación detallada de los distintos tipos de covariables puede verse en Lara (1995).

Se denota por:

$$x_1 = (x_{11}, \ldots,\ x_{1p})^T$$
$$x_2 = (x_{21}, \ldots,\ x_{2q})^T,$$

a dos vectores de covariables correspondientes a dos individuos, por ejemplo, padre-hijo (o a un solo individuo al que se le miden dos tiempos de fallo de distinta índole). El efecto de las covariables suele introducirse en el modelo mediante una función que usualmente es:

$$\psi_i(x_i) = \exp(\beta_i x_i), \quad i : 1, 2,$$

donde $\beta_1 = (\beta_{11}, \ldots, \beta_{1p})$ y $\beta_2 = (\beta_{21}, \ldots, \beta_{2q})$ son los parámetros asociados a las covariables, también llamados coeficientes de regresión.

Se considera un modelo de riesgo proporcional (modelo de regresión de Cox (1972)), es decir, las covariables y el frailty actúan multiplicativamente en la función de riesgo marginal, es decir:

$$h_i\big(t_i,\ \psi_i(x_i),\ z\big) = z \cdot \psi_i(x_i) \cdot h_i(t_i) \quad i : 1, 2.$$

En este caso, las funciones de supervivencia marginales condicionadas a un valor fijo de Z son:

$$S(t_1,\ \psi_i(x_i)/z) = \exp\left(-z \int_0^{t_i} \psi_i(x_i) h_i(t_i) dt_i\right) = \exp\big(-z\psi_i(x_i)\Lambda_i(t_i)\big).$$

La función de supervivencia conjunta es:

$$S(t_1,\ t_2) = \int \exp\Big[-z\big(\exp(\beta_1 x_1)\Lambda_1(t_1) + \exp(\beta_2 x_2)\Lambda_2(t_2)\big)\Big] dF(z),$$

donde $\Lambda_i(t_i)$ corresponde a una función de riesgo base acumulativa.

Si el frailty Z puede modelizarse mediante una distribución gamma, la función de supervivencia conjunta condicionada a un valor fijo de Z es:

$$S(t_1,\ t_2/z) = \exp\Big[-z\big(\exp(\beta_1 x_1)\Lambda_1(t_1) + \exp(\beta_2 x_2)\Lambda_2(t_2)\big)\Big],$$

y la función de supervivencia conjunta queda:

$$S(t_1,\ t_2) = \left[1 + \frac{\exp(\beta_1 x_1)\Lambda_1(t_1) + \exp(\beta_2 x_2)\Lambda_2(t_2)}{\theta}\right]^{-\theta},$$

siendo la función de supervivencia marginal:

$$S_i(t_i) = \left[1 + \frac{\exp(\beta_i x_i)\Lambda_i(t_i)}{\theta}\right]^{-\theta}.$$

En este modelo, que coincide con el de Clayton, la función de supervivencia puede expresarse como:

$$S(t_1,\ t_2) = \left[\frac{1}{\big(S_1(t_1)\big)^{\frac{1}{\theta}}} + \frac{1}{\big(S_2(t_2)\big)^{\frac{1}{\theta}}} - 1\right]^{-\theta},$$

y si se consideran distribuciones Weibull para las marginales, es decir:

$$S_i(t_i) = \exp(-\varepsilon_i t_i^{\gamma_i}),$$

las funciones de supervivencia marginales quedan:

$$S_i(t_i) = \exp\big[-\exp(\beta_i x_i) t_i^{\gamma_i}\big],$$

donde el parámetro de escala queda integrado en las covariables.

Wassel & Moeschberger (1993) proponen la introducción de covariables también en el parámetro de asociación.

2. El modelo mixto

Nosotros vamos a construir un modelo, al que denominaremos modelo *mixto*, en el que las covariables actúan según un modelo de vida acelerada, en tanto el frailty lo hace multiplicativamente en la función de riesgo marginal, esto es, es de riesgo proporcional. Es decir, partiremos de la hipótesis siguiente:

$$h_i\big(t_i,\ z,\ \psi_i(x_i)\big) = z \cdot \exp(\beta_i x_i) h_i\big(t_i \cdot \exp(\beta_i x_i)\big).$$

La función de supervivencia marginal condicionada será:

$$S_i\big(t_i,\ \psi_i(x_i),\ z\big) = \exp\big[-z \cdot \Lambda_i\big(t_i \exp(\beta_i x_i)\big)\big].$$

La función de supervivencia conjunta condicionada será:

$$\begin{aligned} S\Big(t_1,\ \frac{t_2}{z},\ \psi_1(x_1),\ \psi_2(x_2)\Big) &= \exp\Big[-z \exp(\beta_1 x_1) \int_0^{t_1} h_1\big(t_1 \exp(\beta_1 x_1)\big) dt_1 \\ &\quad - z \exp(\beta_2 x_2) \int_0^{t_2} h_2\big(t_2 \exp(\beta_2 x_2)\big) dt_2\Big] \\ &= \exp\Big[-z\Big(\int_0^{t_1} h_1\big(t_1 \exp(\beta_1 x_1)\big) d\big(\exp(\beta_1 x_1) t_1\big) + \\ &\quad + \int_0^{t_2} h_2\big(t_2 \exp(\beta_2 x_2)\big) d\big(\exp(\beta_2 x_2) t_2\big)\Big)\Big] \\ &= \exp\Big[-z\Big(\Lambda_1\big(t_1 \exp(\beta_1 x_1)\big) + \Lambda_2\big(t_2 \exp(\beta_2 x_2)\big)\Big)\Big]. \end{aligned}$$

Realizando la mixtura de distribuciones, la función de supervivencia quedará:

$$S(t_1,\ t_2) = \int \exp\big[-z\big(\Lambda_1(t_1 \exp(\beta_1 x_1)) + \Lambda_2(t_2 \exp(\beta_2 x_2))\big)\big] dF(z).$$

Si el frailty Z se modeliza con una distribución gamma, la función de supervivencia quedará:

$$S(t_1,\ t_2) = \left[1 + \frac{\Delta_1\big(t_1 \exp(\beta x)\big) + \Delta_2\big(t_2 \exp(\gamma y)\big)}{\theta}\right]^{-\theta},$$

donde

$$S_1(t_1) = \left[1 + \frac{\varepsilon_1 t_1^{\delta_1} \exp(\delta_1 \beta x)}{\theta}\right]^{-\theta}.$$

3. Covariables en el parámetro de asociación

Se trata de introducir las covariables, además, a través del parámetro de asociación. De esta manera, involucramos a dichas covariables en la posible asociación entre los tiempos de fallo, es decir, estas covariables explicarán el grado de dependencia entre los tiempos de fallo.

En un modelo frailty gamma, mixto y marginales Weibull la expresión que introduce las covariables en el parámetro de asociación suponemos que es:

$$\theta - 1 = \exp(c + \gamma y),$$

$$\log(\theta - 1) = c + \gamma y = \alpha.$$

En este caso la expresión de la función de supervivencia es

$$S(t_1,\ t_2) = \left[\exp\Big(\varepsilon \exp\big(\eta_1 \beta_1 x_1 + \alpha\big) t_1^{\eta_1}\Big) + \exp\Big(\varepsilon \exp(\eta_2 \beta_2 x_2 + \alpha) t_2^{\eta_2}\Big) - 1\right]^{-\dfrac{1}{\exp\alpha}}.$$

4. Estimación

En supervivencia multivariante los tamaños muestrales han de ser relativamente más grandes que en supervivencia univariante. Es decir, se necesita un conjunto de datos multivariantes suficientemente alto. En la práctica se observa que la toma de datos presenta ciertos problemas. Uno de ellos, muy frecuente, es que no se conocen algunos tiempos de supervivencia. Los datos son “incompletos”. Se dice que se tienen datos censurados. En supervivencia bivariante esto ocurre cuando alguna componente o las dos del vector de datos no se conoce. La censura mas usual es la censura a la derecha en la cual el tiempo observado es menor que el tiempo de supervivencia real. Es decir, al medir aún no ha fallado tal componente. Sería censura a la izquierda si el tiempo observado es mayor que el tiempo de supervivencia real. Al ir a medir ya había fallado. Existen otros tipos de censuras, por intervalos, de tipo I, tipo II, aleatoria, etc. Consideremos censuras a la derecha en lo que sigue. Se dispone de un dato expresado por el vector $(t_1,\ t_2)$. Estableceremos cuatro situaciones distintas definidas por los siguientes grupos:

Grupo 1 si t_1 y t_2 no son censuras,
Grupo 2 si t_1 no es censura pero t_2 sí lo es,

Grupo 3 si t_1 es censura pero t_2 no lo es,
Grupo 4 si t_1 y t_2 son censuras.

Usaremos el indicador de censura siguiente:

$$I(j) = \begin{cases} 1 & \text{si } (t_1,\ t_2) \text{ está en el grupo } j, \\ 0 & \text{si } (t_1,\ t_2) \text{ no está en el grupo } j, \end{cases} \qquad j : 1, 2, 3, 4.$$

La estimación de los parámetros suele hacerse por el método de la máxima verosimilitud. Cada tipo de datos aporta a la verosimilitud una cantidad distinta. Así, un dato no censurado (grupo 1) aporta a la función de verosimilitud la expresión:

$$g1 = \frac{\partial^2 S(t_1,\ t_2)}{\partial t_1 \partial t_2},$$

es decir, el valor correspondiente de la función de densidad. Un dato del grupo 2 aporta la expresión:

$$g2 = \frac{-\partial S(t_1,\ t_2)}{\partial t_1}.$$

Un dato del grupo 3 aporta la expresión:

$$g3 = \frac{-\partial S(t_1,\ t_2)}{\partial t_2},$$

y un dato del grupo 4 aporta el valor correspondiente de la función de supervivencia, es decir, la expresión:

$$g4 = S(t_1, t_2).$$

En resumen, si se tiene la muestra:

$$(t_{1i},\ t_{2i}), \quad i : 1, \ldots, n$$

la función de verosimilitud es:

$$L = \sum_{i=1}^{n} \left[\frac{\partial^2 S(t_{1i},\ t_{2i})}{\partial t_{1i} \partial t_{2i}} \right]^{I(1)} \left[\frac{-\partial S(t_{1i},\ t_{2i})}{\partial t_{1i}} \right]^{I(2)} \left[\frac{-\partial S(t_{1i},\ t_{2i})}{\partial t_{2i}} \right]^{I(3)} \left[S(t_{1i},\ t_{2i}) \right]^{I(4)},$$

entendiendo tales derivadas respecto a t_1 ó t_2, y evaluadas en t_{1i} ó t_{2i}, respectivamente.

Considerando lo anterior, partamos de un frailty (riesgo proporcional) con distribución gamma, marginales con distribución Weibull, sujeto a censuras y en presencia de covariables que actúan según hipótesis de vida acelerada. Es decir, se trata de un modelo mixto y además, con covariables sobre el parámetro de asociación.

El modelo, sin concretar las marginales y sin covariables, tiene función de supervivencia:

$$S(t_1,\ t_2) = \left[S_1(t_1)^{1-\theta} + S_2(t_2)^{1-\theta} - 1\right]^{\frac{1}{1-\theta}}, \quad \theta \geq 1.$$

Casos especiales son los siguientes:

$$\begin{array}{ll} \theta = 1 & \text{independencia}, \\ \theta \to \infty & S(t_1,\ t_2) = \min\big(S_1(t_1),\ S_2(t_2)\big). \end{array}$$

Que haya covariables dentro del parámetro de asociación puede ser debido a que haya covariables que incidan o tengan cierta influencia en la correlación. Es de observar que la media del frailty gamma es 1 y la varianza $\exp(c+\gamma x_i)$, según la formulación que exponemos a continuación. Así, las covariables explican la asociación. Dado que en la expresión anterior del modelo aparece el factor $\theta-1$ expresaremos:

$$\begin{aligned} \theta_i - 1 &= \exp(c + \gamma y_i), \\ \log(\theta_i - 1) &= c + \gamma y_i = \alpha. \end{aligned}$$

Las funciones de supervivencia marginales son:

$$\begin{aligned} S_1(t_1) &= \exp\big[-\big(\varepsilon \exp(\eta_1\beta_1 x_1)\big)t_1^{\eta_1}\big] \\ S_2(t_2) &= \exp\big[-\big(\varepsilon \exp(\eta_2\beta_2 x_2)\big)t_2^{\eta_2}\big] \end{aligned}$$

y la función de supervivencia conjunta será:

$$\begin{aligned} S(t_1,\ t_2) &= \Bigg[\exp\Big(\big(\varepsilon \exp(\eta_1\beta_1 x_1)\big)t_1^{\eta_1} \exp(c+\gamma y)\Big) + \\ &\qquad + \exp\Big(\big(\varepsilon \exp(\eta_2\beta_2 x_2)\big)t_2^{\eta_2} \exp(c+\gamma y)\Big) - 1\Bigg]^{\frac{1}{-\exp(c+\gamma y)}} \\ &= \Bigg[\exp\Big(\varepsilon \exp\big(\eta_1\beta_1 x_1 + \alpha\big)t_1^{\eta_1}\Big) + \\ &\qquad + \exp\Big(\varepsilon \exp\big(\eta_2\beta_2 x_2 + \alpha\big)t_2^{\eta_2}\Big) - 1\Bigg]^{\frac{1}{-\exp\alpha}} \\ &= A^{\frac{1}{-\exp(c+\gamma y)}}. \end{aligned}$$

Consideramos una muestra:

$$(t_{1i},\ t_{2i}), \quad i:1,...,n$$

donde además dispondremos de los valores correspondientes de las covariables. Los parámetros a estimar son:

$$\eta_1,\ \eta_2,\ \varepsilon,\ \beta_1=(\beta_{11},\ldots,\beta_{1J}),\ \beta_2=(\beta_{21},\ldots,\beta_{2K}),\ c,\ \gamma=(\gamma_1,\ldots,\gamma_M)$$

es decir $1+1+1+J+K+1+M$ parámetros, y habremos de calcular:

$$\frac{\partial \log L}{\partial \eta_1},\ \frac{\partial \log L}{\partial \eta_2},\ \frac{\partial \log L}{\partial \varepsilon}$$

$$\frac{\partial \log L}{\partial \beta_{1j}},_{(j:1,\ldots,J)},\ \frac{\partial \log L}{\partial \beta_{2k}},_{(k:1,\ldots,K)}$$

$$\frac{\partial \log L}{\partial c},\ \frac{\partial \log L}{\partial \gamma_m},_{(m:1,\ldots,M)}\cdot$$

Comencemos calculando las siguientes derivadas:

$$\begin{aligned}
\frac{\partial S(t_1,\ t_2)}{\partial t_1} &= \frac{-1}{\exp(\alpha)} A^{\frac{-1}{\exp(\alpha)}-1} \exp\big(\varepsilon t_1^{\eta_1}\exp(\eta_1\beta_1 x_1+\alpha)\big)\exp(\eta_1\beta_1 x_1+\alpha)\varepsilon\eta_1 t_1^{\eta_1-1},\\
\frac{\partial S(t_1,\ t_2)}{\partial t_2} &= \frac{-1}{\exp(\alpha)} A^{\frac{-1}{\exp(\alpha)}-1} \exp\big(\varepsilon t_2^{\eta_2}\exp(\eta_2\beta_2 x_2+\alpha)\big)\exp(\eta_2\beta_2 x_2+\alpha)\varepsilon\eta_2 t_2^{\eta_2-1},
\end{aligned}$$

y

$$\begin{aligned}
\frac{\partial^2 S(t_1,\ t_2)}{\partial t_1 \partial t_2} &= -\frac{1}{\exp(\alpha)}\exp\big(\varepsilon t_1^{\eta_1}\exp(\eta_1\beta_1 x_1+\alpha)\big)\exp(\eta_1\beta_1 x_1+\alpha)\cdot\\
&\cdot \varepsilon\eta_1 t_1^{\eta_1-1}\exp\big(\varepsilon t_2^{\eta_2}\exp(\eta_2\beta_2 x_2+\alpha)\big)\exp(\eta_2\beta_2 x_2+\alpha)\cdot\\
&\cdot \varepsilon\eta_2 t_2^{\eta_2-1}\left(-\frac{1}{\exp(\alpha)}-1\right)A^{-\frac{1}{\exp(\alpha)}}-2.
\end{aligned}$$

Luego la log-verosimilitud es:

$$\begin{aligned}\log L = \sum_{i=1}^{n}\Big[& I(1)\log\frac{\partial^2 S(t_{1i},\ t_{2i})}{\partial t_{1i}\partial t_{2i}} + I(2)\log\frac{-\partial S(t_{1i},\ t_{2i})}{\partial t_{1i}} + \\ & + I(3)\log\frac{-\partial S(t_{1i},\ t_{2i})}{\partial t_{2i}} + I(4)\log S(t_{1i},\ t_{2i})\Big] \\ = I(1)\Big[& \sum\log\big(1+\exp(\alpha)\big) - 2\sum\alpha + \varepsilon\sum t_1^{\eta_1}\exp(\eta_1\beta_1 x_1+\alpha) + \\ & + \sum(\eta_1\beta_1 x_1+\alpha) + n\log\varepsilon + n\log\eta_1 + (\eta_1-1)\sum\log t_1 + \\ & + \varepsilon\sum t_2^{\eta_2}\exp(\eta_2\beta_2 x_2+\alpha) + \sum(\eta_2\beta_2 x_2+\alpha) + n\log\varepsilon + n\log\eta_2 + \\ & + (\eta_2-1)\sum\ln t_2 - \sum\left(\frac{1}{\exp(\alpha)}+2\right)\log A\Big] + \\ & + I(2)\Big[-\sum\alpha - \sum\left(\frac{1}{\exp(\alpha)}+1\right)\log A + \varepsilon\sum t_1^{\eta_1}\exp(\eta_1\beta_1 x_1+\alpha) + \\ & + \sum(\eta_1\beta_1 x_1+\alpha) + n\log\varepsilon + n\log\eta_1 + (\eta_1-1)\sum\log t_1\Big] + \\ & + I(3)\Big[-\sum\alpha - \sum\left(\frac{1}{\exp(\alpha)}+1\right)\log A + \varepsilon\sum t_2^{\eta_2}\exp(\eta_2\beta_2 x_2+\alpha) + \\ & + \sum(\eta_2\beta_2 x_2+\alpha) + n\log\varepsilon + n\log\eta_2 + (\eta_2-1)\sum\log t_2\Big] - \\ & - I(4)\sum\frac{1}{\exp(\alpha)}\log A.\end{aligned}$$

Calculemos las derivadas respecto los parámetros: llamemos

$$a = \exp\big[\varepsilon\exp(\eta_1\beta_1 x_1+\alpha)t_1^{\eta_1}\big]\cdot\big[\varepsilon\exp(\eta_1\beta_1 x_1+\alpha)t_1^{\eta_1}\log t_1 + \varepsilon t_1^{\eta_1}\exp(\eta_1\beta_1 x_1+\alpha)\beta_1 x_1\big],$$

entonces:

$$\begin{aligned}\frac{\partial\log L}{\partial\eta_1} = I(1)\Big[& \varepsilon\sum t_1^{\eta_1}\log t_1\exp(\eta_1\beta_1 x_1+\alpha) + \varepsilon\sum t_1^{\eta_1}\exp(\eta_1\beta_1 x_1+\alpha)\beta_1 x_1 + \\ & + \sum\beta_1 x_1 + \frac{n}{\eta_1} + \sum\ln t_1 - \sum\left(\frac{1}{\exp(\alpha)}+2\right)\frac{a}{A}\Big] + \\ & + I(2)\Big[-\sum\left(\frac{1}{\exp(\alpha)}+1\right)\frac{a}{A} + \varepsilon\sum t_1^{\eta_1}\log t_1\exp(\eta_1\beta_1 x_1+\alpha) + \\ & + \varepsilon\sum t_1^{\eta_1}\exp(\eta_1\beta_1 x_1+\alpha)\beta_1 x_1 + \sum\beta_1 x_1 + \frac{n}{\eta_1} + \sum\log t_1\Big] + \\ & + I(3)\Big[-\sum\left(\frac{1}{\exp(\alpha)}+1\right)\frac{a}{A}\Big] - I(4)\Big[\sum\frac{1}{\exp(\alpha)}\frac{a}{A}\Big],\end{aligned}$$

e igualmente si llamamos

$$b=\exp\big[\varepsilon(\eta_2\beta_2x_2+\alpha)t_2^{\eta_2}\big]\cdot\big[\varepsilon\exp(\eta_2\beta_2x_2+\alpha)t_2^{\eta_2}\log t_2+\varepsilon t_2^{\eta_2}\exp(\eta_2\beta_2x_2+\alpha)\beta_2x_2\big],$$

entonces:

$$\begin{aligned}\frac{\partial\log L}{\partial\eta_2}=I(1)\Big[&\varepsilon\sum t_2^{\eta_2}\log t_2\exp(\eta_2\beta_2x_2+\alpha)+\\&+\varepsilon\sum t_2\sum\eta_2\exp(\eta_2\beta_2x_2+\alpha)\beta_2x_2+\sum\beta_2x_2+\frac{n}{\eta_2}+\sum\log t_2-\\&-\sum\left(\frac{1}{\exp\alpha}+2\right)\frac{b}{A}\Big]+I(2)\Big[-\sum\left(\frac{1}{\exp\alpha}+1\right)\frac{b}{A}\Big]+\\&+I(3)\Big[-\sum\left(\frac{1}{\exp\alpha}+1\right)\frac{b}{A}+\varepsilon\sum t_2^{\eta_2}\log t_2\exp(\eta_2\beta_2x_2+\alpha)+\\&+\varepsilon\sum t_2^{\eta_2}\exp(\eta_2\beta_2x_2+\alpha)\beta_2x_2\Big]-I(4)\Big[\sum\frac{1}{\exp\alpha}\frac{b}{A}\Big].\end{aligned}$$

Sea ahora

$$\begin{aligned}d=&\exp\big[\varepsilon\exp(\eta_1\beta_1x_1+\alpha)t_1^{\eta_1}\big]\exp\big[(\eta_1\beta_1x_1+\alpha)t_1^{\eta_1}\big]+\\&+\exp\big[\varepsilon\exp(\eta_2\beta_2x_2+\alpha)t_2^{\eta_2}\big]\exp\big[(\eta_2\beta_2x_2+\alpha)t_2^{\eta_2}\big],\end{aligned}$$

entonces:

$$\begin{aligned}\frac{\partial\log L}{\partial\varepsilon}=I(1)\Big[&\sum t_1^{\eta_1}\exp(\eta_1\beta_1x_1+\alpha)+\frac{n}{\varepsilon}+\sum t_2^{\eta_2}\exp(\eta_2\beta_2x_2+\alpha)+\\&+\frac{n}{\varepsilon}-\sum(\frac{1}{\exp\alpha}+2)\frac{d}{A}\Big]+\\&+I(2)\Big[-\sum\left(\frac{1}{\exp\alpha}+1\right)\frac{d}{A}+\sum t_1^{\eta_1}\exp(\eta_1\beta_1x_1+\alpha)+\frac{n}{\varepsilon}\Big]+\\&+I(3)\Big[-\sum\left(\frac{1}{\exp\alpha}+1\right)\frac{d}{A}+\sum t_2^{\eta_2}\exp(\eta_2\beta_2x_2+\alpha)+\frac{n}{\varepsilon}\Big]-\\&-I(4)\Big[\sum\frac{1}{\exp\alpha}\frac{d}{A}\Big].\end{aligned}$$

Sea ahora

$$\begin{aligned}f=&\exp\big[\varepsilon\exp(\eta_1\beta_1x_1+\alpha)t_1^{\eta_1}\big]\varepsilon t_1^{\eta_1}\exp(\eta_1\beta_1x_1+\alpha)+\\&+\exp\big[\varepsilon\exp(\eta_2\beta_2x_2+\alpha)t_2^{\eta_2}\big]\varepsilon t_2^{\eta_2}\exp(\eta_2\beta_2x_2+\alpha),\end{aligned}$$

entonces:

$$\begin{aligned}\frac{\partial \log L}{\partial c} = I(1)\Big[&\sum \frac{\exp\alpha}{1+\exp\alpha} - 2n + \varepsilon \sum t_1^{\eta_1} \exp(\eta_1\beta_1 x_1 + \alpha) + n + \\ &+ \varepsilon \sum t_2^{\eta_2} \exp(\eta_2\beta_2 x_2 + \alpha) + n - \sum\Big(- \exp(-\alpha)\log A + \\ &+ \Big(\frac{1}{\exp\alpha} + 2\Big)\frac{f}{A}\Big)\Big] + I(2)\Big[-n - \sum\Big(-\exp(-\alpha)\log A + \\ &+ \Big(\frac{1}{\exp\alpha}+1\Big)\frac{f}{A}\Big) + \varepsilon\sum t_1^{\eta_1}\exp(\eta_1\beta_1 x_1+\alpha) + n\Big] + \\ &+ I(3)\Big[-\sum\Big(-\exp(-\alpha)\log A + \Big(\frac{1}{\exp\alpha}+1\Big)\frac{f}{A}\Big) + \\ &+ \varepsilon\sum t_2^{\eta_2}\exp(\eta_2\beta_2 x_2+\alpha)\Big] - I(4)\Big[-\exp(-\alpha)\log A + \frac{1}{\exp\alpha}\frac{f}{A}\Big].\end{aligned}$$

De manera análoga, sea

$$h = \exp\big[\varepsilon \exp(\eta_1\beta_1 x_1 + \alpha) t_1^{\eta_1}\big] \varepsilon t_1^{\eta_1} \exp(\eta_1\beta_1 x_1 + \alpha)\eta_1 x_{11},$$

entonces:

$$\begin{aligned}\frac{\partial \log L}{\partial \beta_{11}} = I(1)\Big[&\varepsilon\sum t_1^{\eta_1}\exp(\eta_1\beta_1 x_1+\alpha)\eta_1 x_{11} + \sum \eta_1 x_{11} - \\ &- \sum\Big(\frac{1}{\exp}\alpha + 2\Big)\frac{h}{A}\Big] + I(2)\Big[-\sum\Big(\frac{1}{\exp\alpha}+1\Big)\frac{h}{A} + \\ &+ \varepsilon\sum t_1^{\eta_1}\exp(\eta_1\beta_1 x_1+\alpha)\eta_1 x_{11} + \sum\eta_1 x_{11}\Big] + \\ &+ I(3)\Big[-\sum\Big(\frac{1}{\exp\alpha}+1\Big)\frac{h}{A}\Big] - I(4)\Big[\sum\frac{1}{\exp\alpha}\frac{h}{A}\Big],\end{aligned}$$

y así los demás coeficientes de β_1 y β_2.

Por último sea

$$\begin{aligned}k = &\exp\big[\varepsilon\exp(\eta_1\beta_1 x_1+\alpha)t_1^{\eta_1}\big]\varepsilon t_1^{\eta_1}\exp(\eta_1\beta_1 x_1+\alpha) y_1 + \\ &+ \exp\big[\varepsilon\exp(\eta_2\beta_2 x_2+\alpha)t_2^{\eta_2}\big]\varepsilon t_2^{\eta_2}\exp(\eta_2\beta_2 x_2+\alpha) y_1\ ,\end{aligned}$$

entonces:

$$
\begin{aligned}
\frac{\partial \log L}{\partial \gamma_1} = I(1)\Big[&\sum \frac{y_1 \exp\alpha}{1+\exp\alpha} - 2\sum y_1 + \varepsilon \sum t_1^{\eta_1} \exp(\eta_1\beta_1 x_1 + \alpha) y_1 + \sum y_1 + \\
&+ \varepsilon \sum t_2^{\eta_2} \exp(\eta_2\beta_2 x_2 + \alpha) y_1 + \sum y_1 - \sum \Big(-y_1 \exp(-\alpha) \log A + \\
&+ \Big(\frac{1}{\exp\alpha} + 2\Big)\frac{k}{A}\Big)\Big] + I(2)\Big[- \sum \Big((-y_1 e^{-\alpha}) \log A + \Big(\frac{1}{\exp\alpha} + 1\Big)\frac{k}{A}\Big) + \\
&+ \varepsilon \sum t_1^{\eta_1} \exp(\eta_1\beta_1 x_1 + \alpha) y_1\Big] + I(3)\Big[- \sum \Big(-y_1 e^{-\alpha} \log A + \\
&+ \Big(\frac{1}{\exp\alpha} + 1\Big)\frac{k}{A}\Big) + \varepsilon \sum t_2^{\eta_2} \exp(\eta_2\beta_2 x_2 + \alpha) y_1\Big] - \\
&- I(4)\Big[\sum \Big(-y_1 e^{-\alpha} \log A + \frac{k}{A}\frac{1}{e^{\alpha}}\Big)\Big]
\end{aligned}
$$

e igualmente los demás coeficientes del vector γ.

Si

$$
\begin{aligned}
\phi &= (\eta_1,\ \eta_2,\ \varepsilon,\ \beta_{11}, \ldots,\ \beta_{1J},\ \beta_{21}, \ldots,\ \beta_{2K},\ c,\ \gamma_1, \ldots,\ \gamma_M) \\
&= (\phi_1, \ldots,\ \phi_{4+J+K+M})\ ,
\end{aligned}
$$

las soluciones del sistema:

$$
\frac{\partial \log L}{\partial \phi_i} = 0, \quad i : 1, \ldots, 4 + J + K + M\ ,
$$

proporcionan los estimadores de máxima verosimilitud. Para que correspondan a un máximo habría de ser definida positiva la matriz:

$$
I(\phi) = \left(E\left(\frac{\partial L}{\partial \phi_i}\right)^2 \right) = \left(-E \frac{\partial^2 \log L}{\partial \phi_i \partial \phi_j} \right).
$$

Un problema abierto es que

$$
\hat{\phi} \to N\Big(\phi,\ \big(nI(\phi)\big)^{-1}\Big).
$$

Las soluciones de las ecuaciones de verosimilitud son bastante complejas. Para su tratamiento se necesitan paquetes matemáticos (Mathematica...).

Bibliografía

Clayton, D. G. & Cuzick, J. (1985), 'Multivariate generalizations of the proportional hazard model', *Journal of the Royal Statistical Society* **A, 148**, 82–117.

Cox, D. R. (1972), 'Regression models and life-tables (with discussion)', *Journal of the Royal Statistical Society* **B, 34**, 187–220.

Hougaard, P. (1986a), 'A class of multivariate failure time distributions', *Biometrika* **73**, 671–8.

Hougaard, P. (1986b), 'Survival models for heterogeneous populations derived from stable distributions', *Biometrika* **73**, 387–96.

Hougaard, P., Harvald, B. & Holm., N. V. (1992), 'Measuring the similarities between the lifetimes of adult danish twins born between 1881-1930', *Journal of the American Statistical Association* **87, 417**, 17–24.

Lara, A. M. (1995), *Aportaciones a modelos de supervivencia: distribuciones base con puntos de cambio y covariables dependientes del tiempo*, Tesis doctoral. Universidad de Granada, Granada.

Oakes, D. (1989), 'Bivariate survival models induced by frailties', *Journal of the American Statistical Association* **84, 406**, 487–493.

Vaupel, J. W., Manton, K. G. & Stallard, E. (1979), 'The impact of heterogeneity in individual frailty and the dynamics of mortality', *Demography* **16**, 439–454.

Wassel, J. T. & Moeschberger, M. L. (1993), 'A bivariate survival model with modified gamma frailty for assessing the impact of interventions', *Statistics in Medicine* **12**, 241–248.