Revista Colombiana de Estadística
Volumen 30 No. 2. pp. 213 a 229. Diciembre 2007

Modelo de Markov de tres estados: comparación de parametrizaciones de la tasa de intensidad de transición. Aplicación a datos de artritis reumatoidea

Three state Markov model: comparing three parameterizations of the transition intensity rate. Application to rheumatoid arthritis data

Juan Carlos Salazar[1,a], René Iral Palomino[1,b], Enrique Calvo[5,c], Adriana Rojas[2,d], María Eugenia Hincapié[2,e], Juan Manuel Anaya[2,3,4,f], Francisco Javier Díaz[1,g]

[1]Escuela de Estadística, Universidad Nacional de Colombia, Medellín, Colombia
[2]Corporación para Investigaciones Biológicas, CIB, Medellín, Colombia
[3]Universidad del Rosario, Bogotá, Colombia
[4]Clínica Universitaria Bolivariana, Medellín, Colombia
[5]Departamento Imágenes Diagnósticas, Facultad de Medicina, Universidad Nacional de Colombia, Bogotá, Colombia

Resumen

Se considera un modelo múltiple de tres estados donde uno de ellos es absorbente. Se asume que la dependencia entre las observaciones registradas para un mismo sujeto sigue un proceso de Markov. Se comparan, vía simulación, tres diferentes parametrizaciones de la tasa de intensidad de transición: la primera está basada en el modelo de *hazard* multiplicativo de Andersen-Gill (Andersen et al. 1993), la segunda, en el modelo logístico, y la tercera depende del modelo log-log complementario. El método de estimación de parámetros se basa en la función de verosimilitud la cual se optimiza usando las soluciones exactas de un sistema de ecuaciones de Kolmogorov hacia adelante junto con el algoritmo de Newton-Raphson (Abramowitz & Stegun 1972). Usando el sesgo relativo, se selecciona el mejor método de parametrización y se ilustra usando datos recopilados en la Corporación para Investigaciones

[a]Profesor asistente. E-mail: jcsalaza@unalmed.edu.co
[b]Profesor asociado. E-mail: riral@unalmed.edu.co
[c]Profesor asociado. E-mail: ecalvopa@yahoo.com
[d]Médica reumatóloga. E-mail: arojas@cib.org.co
[e]Investigadora asistente. E-mail: mehincapie@cib.org.co
[f]Profesor adjunto. E-mail: janaya@cib.org.co
[g]Profesor asociado. E-mail: fjdiaz@unalmed.edu.co

Biológicas, CIB[1], acerca de pacientes con artritis reumatoidea.

***Palabras clave*:** procesos estocásticos, tasas de intensidad, datos longitudinales, artritis reumatoidea.

Abstract

We consider a three state model with an absorbing state assuming an underlying Markov process to explain the dependence among observations within subjects. We compare, using a simulation study, three different parameterizations of the transition intensity rate: the first one is based on the Andersen-Gill's multiplicative hazard model (Andersen et al. 1993), the second one is based on the logistic model, and the third one depends on the complementary log-log model. The method to estimate the effect of the parameters is based on the likelihood function which can be optimized using the exact solutions of a Kolmogorov forward differential equations system in conjunction with the Newton-Raphson algorithm (Abramowitz & Stegun 1972). We use the relative bias to select the best estimation estrategy. The methodology is ilustrated using longitudinal data about rheumatoid arthritis (RA) from the Corporación para Investigaciones Biológicas, CIB.

***Key words*:** Stochastic processes, Intensity rates, Longitudinal data, Rheumatoid Arthritis.

1. Introducción

Los modelos de estados múltiples[2] conforman una importante familia de herramientas estadísticas que son apropiadas para el análisis de datos longitudinales con respuesta categórica; por ejemplo, la progresión de una enfermedad incurable, tal como la enfermedad de Alzheimer, artritis reumatoidea o la preferencia de un usuario de telefonía celular por un plan específico. Los modelos de estados múltiples han recibido gran atención en la literatura en años recientes. Con los notables avances en estadística, algunos autores han podido aplicar estos modelos en áreas tales como el estudio de datos longitudinales con valores perdidos y en el campo de los modelos lineales mixtos; ver por ejemplo, Gao (2004), Ten et al. (2000) y (Aitkin & Alfó 1998). Este tipo de modelos de estados múltiples han tenido una exitosa acogida en campos de la ciencia tan diversos como biología, física, farmacia, epidemiología, ciencias sociales y medicina. Este trabajo está motivado, básicamente, por los trabajos realizados por Kay (1986), Marshall et al. (1995), Frydman (1995), Joly & Commenges (1999) y Salazar et al. (2007).

La información acerca de la progresión de un fenómeno, tal como una enfermedad incurable, usualmente se recolecta por medio de un proceso de mediciones repetidas tomadas en diferentes ocasiones en el tiempo (datos longitudinales, Diggle et al. (2002)). Con esto se busca registrar el cambio en el tiempo de una respuesta de interés. Es por esto que determinar la función de intensidad de transición[3]

[1] www.cib.org.co/

[2] También conocidos como modelos de compartimientos (Jacquez 1972).

[3] Conocida también como tasa de intensidad de transición. Se define formalmente en la sección 2.

(Bhat 1994) que se asocia con cada uno de estos cambios, resulta importante para entender e identificar que factores se relacionan con el riesgo que un paciente o una unidad experimental particular tiene de transitar a través de diferentes estados de un proceso. Por tanto, es necesario contar con herramientas que permitan estimar estas tasas con un grado de precisión aceptable. Puesto que la manera de formular las tasas de intensidad en términos de las covariables no es única, es pertinente preguntarse cuál de esas formas resulta más adecuada para el proceso de estimación.

Un proceso cuidadoso de identificación de factores de riesgo debe incluir modelos estadísticos que permitan detectar características que se relacionan con los cambios de estado que una persona puede experimentar en el tiempo y medir su grado de asociación con la respuesta de interés (ver Woodward (1999)). Adicionalmente, debe tener en cuenta el proceso de recolección de datos, la estructura de asociación de las medidas repetidas y la naturaleza de la respuesta. Los modelos de Markov basados en parametrizaciones de las tasas de intensidad son de gran utilidad ya que tienen en cuenta estos elementos con la ventaja adicional que permiten cuantificar el papel de las covariables en las distintas transiciones. Debido a que existen diferentes métodos para incorporar el efecto de las covariables sobre las tasas de intensidad de transición en este tipo de modelos es importante identificar cuál de ellos proporciona respuestas más razonables. El interés específico de este trabajo es estudiar metodologías para evaluar las tasas de intensidad de transición cuando el fenómeno bajo estudio se puede idealizar por medio de un modelo de tres estados con un estado absorbente como el que se ilustra en la figura 1.

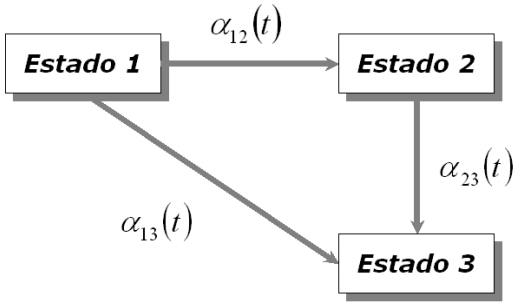

Figura 1: Modelo de tres estados con un estado absorbente.

Como objetivo general se pretende evaluar diferentes metodologías para estimar el conjunto de tasas de intensidad de transición (por brevedad se denominarán de ahora en adelante tasas de intensidad) que gobiernan las transiciones de los sujetos dentro de un modelo múltiple de tres estados con un estado absorbente, y decidir cuál de ellas es más adecuada. Específicamente, se quiere:

1. Cuantificar el cambio, en términos del sesgo relativo, que experimentan los estimadores de máxima verosimilitud de las tasas de intensidad en presencia de distintas parametrizaciones.

2. Identificar condiciones con las cuales se pueden obtener estimadores de máxima verosimilitud de las tasas de intensidad más precisos dentro del modelo múltiple de tres estados con un estado absorbente.

Este trabajo está organizado de la siguiente manera: en la sección 2 se presenta el modelo junto con el sistema de ecuaciones de Kolmogorov hacia adelante; luego, en la sección 3, se detallan las distintas parametrizaciones que se van a comparar; posteriormente, en la sección 4, se exponen los tipos de datos junto con la función de verosimilitud. La sección 5 está dedicada a un estudio de simulación, mientras que en la sección 6 se ilustra el modelo usando datos longitudinales de artritis reumatoidea. Finalmente, en la sección 7, se analizan los méritos y las limitaciones de los modelos.

2. El modelo

Considere un proceso estocástico $\{Y(t) : t \geq 0\}$ con espacio de estados finito $S = \{1, 2, \ldots, k\}$, ($Y(t) \in S$ con probabilidad 1). Se asume que este proceso satisface la propiedad de Markov, esto es:

$$P\big(Y(t_n) = i_n \mid Y(t_1) = i_1, Y(t_2) = i_2, \ldots, Y(t_m) = i_m\big) = P\big(Y(t_n) = i_n \mid Y(t_m) = i_m\big)$$

donde $t_1 < t_2 < \cdots < t_m < t_n$.

Esto implica que el proceso es homogéneo. Adicionalmente, se asume que el estado k es absorbente y que t representa el tiempo transcurrido desde la primera visita. Por ejemplo, en el modelo de tres estados, el espacio de estados es $S = \{1, 2, 3\}$ y solo se admiten las transiciones $1 \to 2$, $1 \to 3$ y $2 \to 3$.

Los estados de este proceso se pueden describir por medio de una matriz de probabilidades de transición dependientes del tiempo (Bhat 1994). Suponga que se toman n historias independientes de este proceso y que cada una de ellas se organiza de acuerdo con su patrón de transición. Este patrón de transición se observa a través de k estados previamente definidos y varía de historia a historia.

El interés es estimar las tasas de intensidad asociadas con los distintos estados del proceso. Formalmente hablando, una tasa de intensidad para una transición de un estado i a un estado j se define como:

$$\alpha_{ij}(t) = \lim_{\Delta_t \to 0} \frac{p\big(Y(t + \Delta_t) = j \mid Y(t) = i\big)}{\Delta_t}$$

donde, $i, j = 1, \ldots, k$ y $\alpha_{ij}(t) \geq 0$, $i, j = 1, \ldots, k$.

Asumir la propiedad de Markov implica que las funciones de intensidad son funciones constantes del tiempo, y por esto resulta apropiado llamarlas tasas de intensidad. Además, esta propiedad sirve para vincular en el modelo la correlación entre las mediciones repetidas de una misma unidad. La relación entre estas tasas

de intensidad y las probabilidades de transición se establece a partir de un sistema de ecuaciones de Kolmogorov hacia adelante (Bhat 1994),

$$\frac{d}{dt}\mathbf{P}(t) = \mathbf{P}(t)\mathbf{Q}, \quad \mathbf{P}(0) = \mathbf{I}_{k+1}, \quad \text{con } \mathbf{Q} = [\alpha_{ij}]$$

donde: $\mathbf{P}(t)$ es la matriz de probabilidades de transición, α_{ij} la tasa de intensidad del estado i al estado j y $\mathbf{Q}$ la matriz de tasas de intensidad. Cuando se considera el modelo de tres estados (ver figura 1) este sistema tiene una solución exacta dada por

$$\begin{aligned}
p_{11}(t) &= \exp\big(-(\alpha_{12}+\alpha_{13})t\big)\\
p_{12}(t) &= \frac{\alpha_{12}}{\alpha_{**}}\big[1-\exp(-\alpha_{**}t)\big]\exp(-\alpha_{23}t)\\
p_{13}(t) &= 1-p_{11}(t)-p_{12}(t)\\
p_{22}(t) &= \exp(-\alpha_{23}t)\\
p_{23}(t) &= 1-\exp(-\alpha_{23}t)\\
\alpha_{**} &= \alpha_{12}+\alpha_{13}-\alpha_{23}
\end{aligned}$$

En la práctica, las tasas de intensidad podrían estar influidas por covariables, y por esta razón se deben incorporar dentro del siguiente conjunto:

$$\mathbf{A} = \{\alpha_{i,j}(\theta;\mathbf{z}) \mid i,j = 1,\ldots,k\}$$

donde θ es un vector de parámetros y $\mathbf{z}^T = [z_1,\ldots,z_p]^T$ es un vector de covariables.

La forma en que estas covariables se vinculan con las tasas de intensidad no es única. El plan de trabajo consiste en evaluar, vía simulación, tres diferentes metodologías para estimar el conjunto de tasas de intensidad que gobiernan las transiciones de los sujetos de un estado a otro. El objetivo de este trabajo es evaluar cuál de esas tres parametrizaciones consideradas resulta ser la más recomendable.

A continuación se detallan las parametrizaciones estudiadas.

3. Parametrizaciones de la tasa de intensidad

Con el fin de estudiar el comportamiento de los estimadores de máxima verosimilitud de las tasas de intensidad se hace una comparación, por medio de un estudio de simulación, de tres diferentes parametrizaciones de la tasa de intensidad.

***Modelo de Andersen-Gill*:** la primera parametrización está basada en el modelo de *hazard* multiplicativo de Andersen-Gill (Andersen et al. 1993):

$$\alpha_{i,j}(\theta;\mathbf{Z}) = \alpha^*_{ij}\, e^{\boldsymbol{\beta}_{ij}^T\mathbf{Z}}$$

aquí, α^*_{ij} es un número positivo no especificado por estimar y $\boldsymbol{\beta}_{ij}$ es un vector que representa los efectos desconocidos de las covariables en una transición del estado i al estado j, $i,j = 1,\ldots,k$.

***Modelo logístico*:** la segunda parametrización está basada en la distribución logística y se define como:

$$\alpha_{i,j}(\theta;\mathbf{Z}) = \frac{\alpha^*_{ij}\ e^{\boldsymbol{\beta}^T_{ij}\mathbf{Z}}}{1+\alpha^*_{ij}\ e^{\boldsymbol{\beta}^T_{ij}\mathbf{Z}}}$$

donde α^*_{ij} es un número positivo no especificado por estimar y $\boldsymbol{\beta}_{ij}$ es un vector que representa los efectos desconocidos de las covariables en una transición del estado i al estado j, $i, j = 1, \ldots, k$.

***Modelo log-log complementario*:** la tercera parametrización está basada en la transformación log-log complementario:

$$\alpha_{i,j}(\theta;\mathbf{Z}) = 1 - e^{-\alpha^*_{ij}e^{\boldsymbol{\beta}^T_{ij}\mathbf{Z}}}$$

donde α^*_{ij} es un número positivo no especificado por estimar y $\boldsymbol{\beta}_{ij}$ es un vector que representa los efectos desconocidos de las covariables en una transición del estado i al estado j, $i, j = 1, \ldots, k$.

Estos tres métodos se comparan con el método no paramétrico (*naive*) para obtener estimaciones de las funciones de intensidad descrito en Kay (1986); básicamente, este método propone estimar las tasas de intensidad usando la siguiente relación:

$$\widehat{\alpha}_{ij} = \frac{m_{ij}}{T_i}$$

donde m_{ij} es el total de transiciones del estado i al j y T_i es el total del tiempo que las unidades permanecieron en el estado i.

4. Función de verosimilitud

El método de estimación de las tasas de intensidad dentro de cada parametrización se basa en la función de verosimilitud. Por tratarse de un proceso homogéneo en el tiempo, las probabilidades de transición de un estado a otro para los individuos solo dependen de la diferencia de tiempos entre visitas sucesivas. La tabla 1 muestra el esquema general de la información recopilada longitudinalmente para un individuo particular.

Tabla 1: Esquema de recolección de datos para un participante con historia de transición: $1 \to 2$, $2 \to 2$, y $2 \to 3$.

Visita	V_1	V_2	V_3	V_4
Tiempos de transición	t_1	t_2	t_3	t_4
Estado	1	2	2	3
Probabilidad de transición	$p_{12}(t_2 - t_1)$	$p_{22}(t_3 - t_2)$	$p_{23}(t_4 - t_3)$	

Para este ejemplo, la contribución de este individuo a la verosimilitud será:

$$p_{12}(t_2 - t_1) * p_{22}(t_3 - t_2) * p_{23}(t_4 - t_3)$$

En general, si $S_i^{(k)}$ representa el estado del k-ésimo participante en la i-ésima visita, n el número de participantes bajo estudio, m_k número de visitas para el k-ésimo participante y $p^{(k)}_{S_i,S_{i+1}}$ la probabilidad de que el k-ésimo participante pase del estado S_i al estado S_{i+1} en el intervalo de tiempo (t_i, t_{i+1}), entonces la contribución del k-ésimo participante a la verosimilitud está dada por

$$\prod_{i=1}^{m_k} p^{(k)}_{S_i,S_{i+1}}(t_{i+1} - t_i)$$

Note que para obtener estimaciones de las tasas de intensidad en el modelo markoviano no es necesario conocer el momento exacto en el que el participante emigra de un estado a otro.

La función de verosimilitud para los n participantes está dada por

$$\prod_{k=1}^{n} \prod_{i=1}^{m_k} p^{(k)}_{S_i,S_{i+1}}(t_{i+1} - t_i)$$

Para estimar las tasas de intensidad, es necesario derivar esta última función de verosimilitud con respecto a los parámetros del modelo. La dificultad radica en que es muy complicado obtener una solución explícita para la ecuación de verosimilitud. Esto es especialmente cierto para modelos con más de tres estados, en donde la solución de las ecuaciones de Kolmogorov hacia adelante se obtienen en términos de la descomposición espectral de la matriz de tasas de intensidad, la cual puede generar soluciones que no son números reales (ver Kalbfleish et al. 1983). Es por esto que se debe usar un método numérico en asocio con las ecuaciones diferenciales para obtener estas estimaciones.

5. Estudio de simulación

Para llevar a cabo el estudio de simulación se asume un proceso estocástico de Markov de tres estados que se denotarán 1, 2 y 3, y donde el estado 3 es absorbente; las únicas transiciones que se admiten son $1 \to 2$, $1 \to 3$ y $2 \to 3$.

Las condiciones en las cuales se ejecutaron las simulaciones se describen a continuación. Primero, se simularon 1000 muestras de tamaños $n = 200$, 300, 400, 500 y 600 unidades, respectivamente, que contenían historias aleatorias de transiciones en el modelo de tres estados para los n participantes; luego para cada tamaño muestral se generaron un máximo de cinco o seis medidas repetidas por unidad y una covariable dicotómica, que para fines prácticos puede indicar la presencia o ausencia de una característica genética; la variable edad se incorpora en el modelo con tres categorías: ≤ 75 años, $75 - 85$ años y ≥ 85 años. Para calcular los valores de referencia se usaron estimaciones de las tasas de intensidad reportadas en un artículo de Harezlak et al. (2003) obtenidas a partir de datos del *Indianapolis-Ibadan Dementia Project* (Hendrie et al. 2001), el cual es un estudio longitudinal acerca del envejecimiento y demencia realizado con participantes de

dos poblaciones ubicadas una en Indianapolis, Estados Unidos, y la otra en Nigeria, África. Estos valores se modificaron ligeramente para incorporar la dependencia de las transiciones en la variable dicotómica generada.

Cada ronda de simulación estimaba los efectos de las covariables y con estos se obtenían estimadores de α_{12}, α_{13} y α_{23} para cada grupo de edad. Todas las simulaciones se llevaron a cabo usando el software SAS IML© (SAS Institute Inc 1990).

El proceso de selección del mejor modelo de parametrización se basa en el sesgo relativo (SR), que se define como:

$$SR = \frac{\widehat{\alpha}_{ij} - \alpha_{ij}}{\alpha_{ij}}, \qquad i, j = 1, 2, 3$$

El modelo asociado con los menores sesgos relativos se considera el más recomendable para obtener los estimadores de máxima verosimilitud de las tasas de intensidad.

En el proceso de simulación, la función de verosimilitud se optimizó iterativamente usando las soluciones exactas de un sistema de ecuaciones de Kolmogorov hacia adelante (asociado con el modelo de tres estados) junto con el método de Newton-Raphson (Abramowitz & Stegun 1972).

La tabla 2 muestra los resultados correspondientes a los tres métodos cuando se simuló con tamaños muestrales de 200, 300, 400, 500 y 600 y un número máximo de visitas de 5, mientras que la tabla 3 muestra los resultados con los mismos tamaños muestrales pero un número máximo de visitas por participante de 6.

En ambas tablas se nota que el método no paramétrico (*naive*) está asociado con las peores estimaciones y que, consistentemente, tiende a subestimar el valor de referencia. También, el modelo de Andersen-Gill subestima consistentemente los valores de referencia, pero en general los SR están dentro de un rango razonablemente cercano a los valores de referencia.

En la tabla 2 se observa una sobreestimación de la tasa α_{12}, excepto para el método no paramétrico, en el grupo de edad de ≤ 75 años. A medida que se aumenta el tamaño muestral, las estimaciones de α_{12} se acercan más a los valores de referencia. Con $n = 200$ y usando el modelo logístico se observa el SR más alto (0.758), mientras que el modelo de Andersen-Gill generó el SR más bajo (0.053). Para el grupo de edad de $75 - 85$ años también se registran valores estimados cercanos a los de referencia siendo el modelo logístico el que produce mejores estimaciones. En este mismo grupo y el de los ≥ 85 el desempeño del modelo logístico y el log-log complementario es mejor que el del modelo de Andersen-Gill. Para el caso de los ≤ 75, los tres métodos producen estimaciones similares.

En el grupo de edad de los ≥ 85 se observan estimaciones cercanas a los valores de referencia; de hecho, el SR oscila entre -0.292 y 0.136 con valores tan pequeños como 0.003 y 0.008.

Cuando el número de visitas máximo se incrementa a seis (ver tabla 3) y se aumentan los tamaños muestrales, se observa una mejora en las estimaciones, pero las tendencias son las mismas que las observadas en el caso de cinco visitas.

Tabla 2: Estimación (sesgo relativo). Número máximo de visitas = 5.

		Grupo de edad								
n	Método	≤ 75			$75-85$			≥ 85		
		α_{12}	α_{13}	α_{23}	α_{12}	α_{13}	α_{23}	α_{12}	α_{13}	α_{23}
200	**Referencia**	**0.019**	**0.050**	**0.079**	**0.051**	**0.078**	**0.125**	**0.119**	**0.130**	**0.208**
	Naive	0.015	0.043	0.054	0.031	0.062	0.068	0.040	0.087	0.083
		(-0.201)	(-0.146)	(-0.315)	(-0.401)	(-0.201)	(-0.459)	(-0.660)	(-0.333)	(-0.599)
	Andersen-Gill	0.030	0.051	0.069	0.053	0.077	0.111	0.107	0.124	0.148
		(0.589)	(0.020)	(-0.126)	(0.033)	(-0.014)	(-0.114)	(-0.101)	(-0.050)	(-0.287)
	Logístico	0.033	0.056	0.076	0.058	0.088	0.128	0.120	0.148	0.175
		(0.758)	(0.123)	(-0.034)	(0.136)	(0.128)	(0.026)	(0.008)	(0.136)	(-0.158)
	Log-Log	0.032	0.052	0.072	0.057	0.082	0.117	0.113	0.134	0.196
		(0.695)	(0.043)	(-0.084)	(0.110)	(0.045)	(-0.062)	(-0.047)	(0.033)	(-0.059)
300	**Referencia**	**0.019**	**0.050**	**0.079**	**0.051**	**0.078**	**0.125**	**0.119**	**0.130**	**0.208**
	Naive	0.013	0.041	0.056	0.024	0.059	0.074	0.033	0.085	0.094
		(-0.342)	(-0.180)	(-0.295)	(-0.522)	(-0.249)	(-0.408)	(-0.719)	(-0.348)	(-0.550)
	Andersen-Gill	0.025	0.047	0.072	0.045	0.072	0.110	0.093	0.116	0.161
		(0.292)	(-0.051)	(-0.093)	(-0.111)	(-0.073)	(-0.121)	(-0.222)	(-0.109)	(-0.228)
	Logístico	0.026	0.051	0.079	0.051	0.079	0.127	0.109	0.130	0.197
		(0.384)	(0.012)	(-0.003)	(-0.009)	(0.008)	(0.015)	(-0.084)	(0.003)	(-0.055)
	Log-Log	0.026	0.049	0.074	0.048	0.076	0.117	0.100	0.124	0.189
		(0.353)	(-0.028)	(-0.059)	(-0.055)	(-0.032)	(-0.064)	(-0.156)	(-0.046)	(-0.093)
400	**Referencia**	**0.019**	**0.050**	**0.079**	**0.051**	**0.078**	**0.125**	**0.119**	**0.130**	**0.208**
	Naive	0.012	0.041	0.057	0.022	0.058	0.075	0.030	0.085	0.098
		(-0.392)	(-0.189)	(-0.283)	(-0.571)	(-0.251)	(-0.398)	(-0.749)	(-0.350)	(-0.531)
	Andersen-Gill	0.022	0.047	0.073	0.043	0.072	0.110	0.089	0.115	0.166
		(0.178)	(-0.067)	(-0.072)	(-0.160)	(-0.082)	(-0.117)	(-0.248)	(-0.118)	(-0.203)
	Logístico	0.024	0.049	0.080	0.046	0.078	0.126	0.098	0.129	0.201
		(0.255)	(-0.016)	(0.016)	(-0.099)	(-0.004)	(0.012)	(-0.178)	(-0.007)	(-0.035)
	Log-Log	0.023	0.047	0.076	0.045	0.074	0.119	0.095	0.122	0.189
		(0.232)	(-0.063)	(-0.033)	(-0.116)	(-0.052)	(-0.049)	(-0.201)	(-0.063)	(-0.092)
500	**Referencia**	**0.019**	**0.050**	**0.079**	**0.051**	**0.078**	**0.125**	**0.119**	**0.130**	**0.208**
	Naive	0.011	0.040	0.056	0.021	0.057	0.076	0.029	0.083	0.099
		(-0.412)	(-0.195)	(-0.288)	(-0.587)	(-0.275)	(-0.393)	(-0.758)	(-0.358)	(-0.524)
	Andersen-Gill	0.021	0.046	0.072	0.041	0.070	0.109	0.084	0.111	0.166
		(0.105)	(-0.085)	(-0.091)	(-0.205)	(-0.101)	(-0.129)	(-0.292)	(-0.145)	(-0.202)
	Logístico	0.023	0.048	0.079	0.045	0.076	0.125	0.095	0.126	0.203
		(0.186)	(-0.048)	(-0.005)	(-0.125)	(-0.028)	(0.004)	(-0.201)	(-0.033)	(-0.024)
	Log-Log	0.022	0.047	0.074	0.043	0.074	0.116	0.090	0.121	0.185
		(0.144)	(-0.063)	(-0.065)	(-0.162)	(-0.054)	(-0.071)	(-0.244)	(-0.072)	(-0.110)
600	**Referencia**	**0.019**	**0.050**	**0.079**	**0.051**	**0.078**	**0.125**	**0.119**	**0.130**	**0.208**
	Naive	0.011	0.040	0.057	0.021	0.057	0.076	0.029	0.084	0.099
		(-0.435)	(-0.204)	(-0.281)	(-0.587)	(-0.269)	(-0.393)	(-0.758)	(-0.355)	(-0.525)
	Andersen-Gill	0.020	0.045	0.073	0.040	0.070	0.109	0.085	0.114	0.165
		(0.053)	(-0.104)	(-0.079)	(-0.208)	(-0.097)	(-0.127)	(-0.287)	(-0.126)	(-0.205)
	Logístico	0.023	0.048	0.080	0.046	0.077	0.126	0.099	0.128	0.202
		(0.186)	(-0.038)	(0.010)	(-0.105)	(-0.013)	(0.007)	(-0.164)	(-0.018)	(-0.031)
	Log-Log	0.021	0.046	0.075	0.042	0.074	0.116	0.090	0.121	0.184
		(0.120)	(-0.077)	(-0.054)	(-0.167)	(-0.057)	(-0.069)	(-0.245)	(-0.069)	(-0.117)

Tabla 3: Estimación (sesgo relativo). Número máximo de visitas = 6.

		Grupo de edad								
n	Método	≤ 75			$75-85$			≥ 85		
		α_{12}	α_{13}	α_{23}	α_{12}	α_{13}	α_{23}	α_{12}	α_{13}	α_{23}
200	**Referencia**	**0.019**	**0.050**	**0.079**	**0.051**	**0.078**	**0.125**	**0.119**	**0.130**	**0.208**
	Naive	0.015	0.044	0.056	0.031	0.064	0.072	0.038	0.088	0.090
		(-0.235)	(-0.128)	(-0.288)	(-0.389)	(-0.181)	(-0.421)	(-0.682)	(-0.320)	(-0.566)
	Andersen-Gill	0.031	0.053	0.073	0.055	0.080	0.118	0.109	0.126	0.167
		(0.616)	(0.066)	(-0.072)	(0.070)	(0.022)	(-0.053)	(-0.080)	(-0.031)	(-0.195)
	Logístico	0.033	0.063	0.082	0.060	0.093	0.138	0.125	0.150	0.211
		(0.741)	(0.268)	(0.038)	(0.174)	(0.196)	(0.108)	(0.049)	(0.157)	(0.013)
	Log-Log	0.033	0.056	0.076	0.058	0.086	0.126	0.117	0.141	0.217
		(0.726)	(0.116)	(-0.037)	(0.146)	(0.106)	(0.011)	(-0.019)	(0.088)	(0.042)
300	**Referencia**	**0.019**	**0.050**	**0.079**	**0.051**	**0.078**	**0.125**	**0.119**	**0.130**	**0.208**
	Naive	0.013	0.041	0.058	0.024	0.060	0.076	0.032	0.087	0.099
		(-0.340)	(-0.178)	(-0.267)	(-0.521)	(-0.226)	(-0.391)	(-0.734)	(-0.332)	(-0.525)
	Andersen-Gill	0.024	0.048	0.076	0.046	0.074	0.118	0.097	0.120	0.176
		(0.269)	(-0.047)	(-0.042)	(-0.096)	(-0.054)	(-0.059)	(-0.182)	(-0.080)	(-0.153)
	Logístico	0.026	0.051	0.083	0.050	0.081	0.135	0.108	0.136	0.215
		(0.347)	(0.023)	(0.045)	(-0.024)	(0.042)	(0.078)	(-0.091)	(0.049)	(0.034)
	Log-Log	0.026	0.048	0.079	0.049	0.076	0.127	0.103	0.127	0.207
		(0.355)	(-0.040)	(-0.003)	(-0.037)	(-0.025)	(0.016)	(-0.136)	(-0.023)	(-0.006)
400	**Referencia**	**0.019**	**0.050**	**0.079**	**0.051**	**0.078**	**0.125**	**0.119**	**0.130**	**0.208**
	Naive	0.012	0.041	0.058	0.022	0.060	0.078	0.030	0.085	0.101
		(-0.371)	(-0.183)	(-0.264)	(-0.559)	(-0.235)	(-0.374)	(-0.748)	(-0.343)	(-0.512)
	Andersen-Gill	0.023	0.046	0.076	0.047	0.070	0.117	0.100	0.112	0.180
		(0.225)	(-0.080)	(-0.040)	(-0.085)	(-0.097)	(-0.068)	(-0.156)	(-0.138)	(-0.134)
	Logístico	0.024	0.049	0.082	0.048	0.078	0.133	0.106	0.129	0.216
		(0.271)	(-0.016)	(0.042)	(-0.052)	(-0.005)	(0.060)	(-0.113)	(-0.007)	(0.036)
	Log-Log	0.024	0.047	0.078	0.047	0.074	0.124	0.102	0.123	0.201
		(0.247)	(-0.056)	(-0.011)	(-0.072)	(-0.046)	(-0.004)	(-0.144)	(-0.056)	(-0.033)
500	**Referencia**	**0.019**	**0.050**	**0.079**	**0.051**	**0.078**	**0.125**	**0.119**	**0.130**	**0.208**
	Naive	0.011	0.041	0.059	0.022	0.059	0.079	0.029	0.086	0.102
		(-0.396)	(-0.180)	(-0.257)	(-0.578)	(-0.238)	(-0.371)	(-0.755)	(-0.341)	(-0.510)
	Andersen-Gill	0.022	0.046	0.077	0.044	0.071	0.117	0.092	0.113	0.180
		(0.159)	(-0.082)	(-0.026)	(-0.142)	(-0.090)	(-0.061)	(-0.223)	(-0.129)	(-0.132)
	Logístico	0.024	0.050	0.085	0.047	0.079	0.136	0.102	0.130	0.220
		(0.237)	(-0.007)	(0.070)	(-0.076)	(0.008)	(0.085)	(-0.142)	(0.004)	(0.059)
	Log-Log	0.023	0.046	0.080	0.047	0.073	0.126	0.101	0.120	0.203
		(0.217)	(-0.071)	(0.010)	(-0.081)	(-0.059)	(-0.010)	(-0.154)	(-0.074)	(-0.023)
600	**Referencia**	**0.019**	**0.050**	**0.079**	**0.051**	**0.078**	**0.125**	**0.119**	**0.130**	**0.208**
	Naive	0.011	0.041	0.058	0.021	0.059	0.079	0.029	0.085	0.102
		(-0.404)	(-0.181)	(-0.264)	(-0.583)	(-0.245)	(-0.369)	(-0.754)	(-0.345)	(-0.508)
	Andersen-Gill	0.022	0.045	0.076	0.044	0.071	0.116	0.094	0.112	0.179
		(0.155)	(-0.090)	(-0.033)	(-0.135)	(-0.095)	(-0.069)	(-0.209)	(-0.135)	(-0.138)
	Logístico	0.024	0.048	0.083	0.049	0.077	0.135	0.108	0.127	0.222
		(0.243)	(-0.033)	(0.049)	(-0.044)	(-0.015)	(0.078)	(-0.094)	(-0.024)	(0.065)
	Log-Log	0.023	0.047	0.079	0.046	0.074	0.126	0.100	0.122	0.202
		(0.185)	(-0.062)	(-0.002)	(-0.097)	(-0.045)	(0.005)	(-0.160)	(-0.065)	(-0.026)

A medida que el tamaño muestral aumenta, se evidencia una reducción en los sesgos relativos de la función α_{12} para el grupo de edad de menores de 75 años. Los sesgos relativos asociados con los otros grupos de edad y las otras funciones se mantienen en un rango razonablemente bajo ($-29.2\,\%$ y $14.8\,\%$).

De acuerdo con este estudio de simulación, tamaños muestrales cercanos a 200 y un seguimiento de aproximadamente cinco visitas producen estimadores dentro de un rango razonablemente cercano a los valores de referencia.

El desempeño de los tres métodos es similar en todas las combinaciones de tamaños muestrales y número de visitas. Esto implica que cualquiera de ellos se puede recomendar en la práctica notándose una ligera ventaja en el modelo de Andersen-Gill por ser éste un estimador no acotado por encima; además, este método es más estable computacionalmente como se puede observar en las tablas 4 y 5 donde se reportan las tasas de convergencia de cada simulación asociadas a cada modelo.

Según este estudio de simulación, el cambio que experimentan los estimadores de máxima verosimilitud de las funciones de intensidad bajo distintas parametrizaciones no es muy significativo.

Tabla 4: Tasas de convergencia asociadas a los tres métodos de estimación. Número máximo de visitas = 5.

	Tamaño muestral				
Modelo	200	300	400	500	600
Andersen-Gill	97.1 %	99.4 %	99.5 %	99.6 %	99.6 %
Logístico	94.8 %	98.1 %	98.1 %	98.8 %	98.7 %
Log-Log	75.1 %	90.5 %	95.0 %	97.5 %	97.7 %

Tabla 5: Tasas de convergencia asociadas a los tres métodos de estimación. Número máximo de visitas = 6.

	Tamaño muestral				
Modelo	200	300	400	500	600
Andersen-Gill	98.1 %	99.0 %	98.9 %	99.6 %	98.9 %
Logístico	96.1 %	97.3 %	98.1 %	99.1 %	97.5 %
Log-Log	81.3 %	91.6 %	96.2 %	97.7 %	96.7 %

6. Ejemplo: progresión radiográfica en artritis reumatoidea

El registro radiográfico de daño en las articulaciones de manos y pies de un paciente con artritis reumatoidea (AR) es de considerable interés para entender esta enfermedad. Puesto que los patrones de daño radiográfico varían dentro del mismo paciente y también de paciente a paciente, es razonable modelar esta evolución asumiendo la existencia de estados predefinidos a través de los cuales la enfermedad progresa. Una pregunta de interés se relaciona con la identificación de

factores que pueden incidir en el tránsito por los distintos estados de severidad de la enfermedad.

Los datos utilizados para ilustrar incluyen 464 radiografías de 146 pacientes diagnosticados con artritis reumatoidea, 84.9 % mujeres, edad promedio 47.1 ± 13.4 años, un promedio de tres radiografías de manos y pies.

La artritis reumatoidea (AR) es una enfermedad crónica autoinmune e inflamatoria que compromete las articulaciones que tienen movimiento (Anaya et al. 2006). Afecta principalmente a las mujeres entre la cuarta y quinta décadas de la vida. Con frecuencia compromete otros órganos distintos a las articulaciones. Dadas las características mencionadas, la AR tiene un impacto adverso en la esfera biopsicosocial y su costo es alto (Anaya et al. 2006).

Estos pacientes fueron monitoreados durante un promedio de 4.2 ± 1.6 años por el mismo equipo médico y siguiendo un régimen terapéutico estándar. Las radiografías de cada paciente fueron leídas y evaluadas por dos profesionales calificados siguiendo el método de Sharp van der Heijde (Van der Heijde 1999). El objetivo consistía en registrar el promedio de erosiones en manos y pies. Las lecturas se hicieron de manera independiente y ninguno de los dos lectores conocía el nombre del paciente al cual pertenecía una radiografía específica (ver figura 2). El acuerdo en sus lecturas fue evaluado posteriormente haciendo uso del coeficiente de correlación intraclase (ICC) y se observó un acuerdo notable entre ambas lecturas; de hecho, el ICC para las lecturas de manos fue de 0.95 mientras que para pies fue de 0.80. El puntaje de erosión, correspondiente al promedio de manos y pies de acuerdo al método de Sharp van der Heijde, fue categorizado en tres niveles de acuerdo con su severidad: 1 = Leve (≤ 4), 2 = Moderado (> 4 y ≤ 16) y 3 = Severo (> 16).

FIGURA 2: Radiografía de manos comparativas en proyección oblicua (1) y anteroposterior (2) con compromiso moderado, según puntaje Sharp van der Heijde. Se destacan erosiones a nivel de las articulaciones metacarpofalángicas (flecha blanca).

Debido a la naturaleza progresiva de la enfermedad, la transición del estado leve al severo no es posible, y por tanto esta ilustración es un caso particular del modelo expuesto en el estudio de simulación (ver figura 1). Treinta pacientes que fueron observados en el estado severo al momento de la primera radiografía fueron

excluidos del análisis ya que no aportaban información relevante para el proceso de estimaciones de las funciones de intensidad.

Para la estimación de las funciones de intensidad $(\alpha_{LM}, \alpha_{MS})$ se incluyó la covariable ausencia o presencia de la secuencia del Share Epitope (SE): 0 = No, 1 = Sí. Esta característica genética ha sido reportada en la literatura como un factor de riesgo para la artritis reumatoidea (Delgado et al. 2006). La historia familiar no se incluyó en el análisis puesto que se contaba con muy poca información.

La función de verosimilitud se optimizó usando el algoritmo de Newton-Raphson implementado en el SAS IML© (SAS Institute Inc 1990) usando las tres parametrizaciones. En la tabla 6 se observan los valores estimados de las funciones de intensidad para los grupos con presencia y ausencia de la secuencia genética del SE.

Tabla 6: Estimación de las funciones de intensidad (error estándar) α_{LM} y α_{MS} usando las tres parametrizaciones ($n = 146$).

Función de intensidad estimada usando el modelo de Andersen-Gill	Ausencia de SE	Presencia de SE	Estadístico z	Valor-p
Leve → *Moderado*, $(\widehat{\alpha}_{LM})$	0.117 (0.028)	0.110 (0.029)	−0.174	0.862
Moderado → *Severo*, $(\widehat{\alpha}_{MS})$	0.097 (0.042)	0.264 (0.078)	1.885	0.059
Función de intensidad estimada usando transformación logit	Ausencia de SE	Presencia de SE	Estadístico z	Valor-p
Leve → *Moderado*, $(\widehat{\alpha}_{LM})$	0.132 (0.028)	0.124 (0.029)	−0.198	0.843
Moderado → *Severo*, $(\widehat{\alpha}_{MS})$	0.107 (0.043)	0.358 (0.122)	1.940	0.052
Función de intensidad estimada usando transformación log-log	Ausencia de SE	Presencia de SE	Estadístico z	Valor-p
Leve → *Moderado*, $(\widehat{\alpha}_{LM})$	0.124 (0.028)	0.117 (0.029)	−0.174	0.862
Moderado → *Severo*, $(\widehat{\alpha}_{MS})$	0.102 (0.043)	0.307 (0.135)	1.447	0.148

La tabla 6 muestra los valores estimados de las tasas de intensidad junto con el valor del estadístico z y su correspondiente Valor-p. Los errores estándar se calcularon usando un método reportado en Iral & Salazar (2006) que está basado en el método delta.

En el modelo de Andersen-Gill no se observa una diferencia significativa en las tasas de intensidad de Leve a Moderado en ausencia y presencia del SE (Valor-p = 0.862); sin embargo, se observa una diferencia importante en las tasas de intensidad de Moderado a Severo en ausencia y presencia del SE (Valor-p = 0.059).

En el modelo logístico no se observa una diferencia significativa en las tasas de intensidad de Leve a Moderado en ausencia y presencia del SE (Valor-p = 0.843) pero el modelo sí detecta una diferencia importante en las tasas de intensidad de Moderado a Severo en ausencia y presencia del SE (Valor-p = 0.052).

Finalmente, en el modelo log-log complementario no se detecta ninguna diferencia importante entre las tasas de intensidad de Leve a moderado y de Moderado a Severo (Valor-p = 0.862 y Valor-p = 0.148, respectivamente).

Según estos resultados, la influencia del SE es más importante en transiciones del estado Moderado a Severo.

7. Conclusiones y recomendaciones

En este trabajo se ha discutido el problema de estimar las tasas de intensidad en un proceso de Markov de múltiples estados al comparar tres métodos distintos para parametrizar la tasa de intensidad. Por medio de un estudio de simulación intensivo se evaluó el comportamiento de cada una de estas parametrizaciones según diferentes especificaciones de tamaño muestral y número máximo de visitas. Se recomienda la parametrización basada en el modelo de Andersen-Gill ya que no está acotada por encima y es más estable computacionalmente; esto la hace más apropiada en la práctica. Sin embargo, las otras parametrizaciones se pueden usar en situaciones donde se sospecha que las tasas de intensidad son pequeñas. Una de las desventajas más notables del modelamiento discutido en este trabajo es que requiere una gran cantidad de datos. Otros autores han propuesto modelos que tienen en cuenta censura arbitraria, pero la complejidad en su implementación los hacen poco prácticos (Joly & Commenges 1999). No obstante, para respuestas categóricas recolectadas longitudinalmente este tipo de modelamiento ha demostrado efectividad siempre y cuando se cuente con un número apropiado de datos y visitas.

Agradecimientos

Agradecimiento especial a todos los pacientes que hicieron parte del estudio radiográfico sobre progresión en artritis reumatoidea. Su aporte es esencial para avanzar en el entendimiento de este mal que afecta a una cantidad considerable de personas en la actualidad. Proyecto patrocinado por UNC-DIME Número 30802921.

Recibido: mayo de 2007
Aceptado: octubre de 2007

Referencias

Abramowitz, M. & Stegun, I. A. (1972), *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, Dover Publications, Inc., New York.

Aitkin, M. & Alfó, M. (1998), 'Regression Models for Binary Longitudinal Responses', *Statistics and Computing* **8**, 289–307.

Anaya, J. M., Pineda, R., Gómez, L. M., Galarza, C., Rojas, A. & Martín, J. (2006), *Artritis reumatoide: bases moleculares, clínicas y terapéuticas*, CIB, Universidad del Rosario, FUNPAR, Medellín, Colombia.

Andersen, P. K., Borgan, O., Gill, R. D. & Keiding, N. (1993), *Statistical Models Based on Counting Processes*, Springer-Verlag, New York.

Bhat, U. N. (1994), *Elements of Applied Stochastic Processes*, second edn, Wiley, New York.

Delgado, A. M., Martín, J., Granados, J. & Anaya, J. M. (2006), 'Epidemiología genética de la artritis reumatoide: ¿Qué esperar en América Latina?', *Biomédica* **26**, 562–584.

Diggle, P., Heagerty, P., Liang, K. Y. & Zeger, S. (2002), *Analysis of Longitudinal Data*, Oxford Statistical Science Series 25, second edn, Oxford University Press Inc., New York.

Frydman, H. (1995), 'Semiparametric Estimation in a Three-State Duration Dependent Markov Model from Interval-Censored Observations with Application to Aids Data', *Biometrics* **51**, 502–511.

Gao, S. (2004), 'A Shared Random Effect Parameter Approach for Longitudinal Dementia Data with Non-Ignorable Missing Data', *Statist. Med.* **23**, 211–219.

Harezlak, J., Gao, S. & L., H. S. (2003), 'An Illness-death Scholastic Model in the Analysis of Longitudinal Dementia Data', *Statistics in Medicine* **22**, 1465–1475.

Hendrie, H. C., Ogunniyi, A., Hall, K. S., Baiyewu, O., Unverzagt, F. W. & Gureje, O. (2001), 'Incidence of Dementia and Alzheimer Disease in 2 Communities: Yoruba Residing in Ibadan, Nigeria, and African Americans Residing in Indianapolis', *JAMA* **285**(6), 739–747.

Iral, R. & Salazar, J. C. (2006), Efecto de las covariables en la estimación de intervalos de confianza para las tasas de transición en un modelo de Markov de tres estados, *in* 'Memorias XVI Simposio de Estadística 2006: Estadística en la Industria. III encuentro Colombia-Venezuela de Estadística', Universidad Nacional de Colombia, Bucaramanga, Colombia.

Jacquez, J. A. (1972), *Compartmental Analysis in Biology and Medicine: Kinetics of Distribution of Tracer-Labeled Materials*, Elsevier Pub. Co., Amsterdam, New York.

Joly, P. & Commenges, D. (1999), 'A Penalized Likelihood Approach for a Progressive Three-State Model with Censored and Truncated Data: Application to AIDS', *Biometrics* **55**, 887–890.

Kalbfleish, J. D., Lawless, J. F. & Vollmer, W. M. (1983), 'Estimation in Markov Models from Aggregate Data', *Biometrics* **39**, 907–919.

Kay, R. (1986), 'A Markov Model for Analyzing Cancer Markers and Disease States in Survival Studies', *Biometrics* **42**(4), 855–865.

Marshall, G., Guo, W. & Jones, R. H. (1995), 'MARKOV: A Computer Program for Multi-State Markov Models with Covariables', *Computer Methods and Programs in Biomedicine* **47**, 147–156.

Salazar, J. C., Schmitt, F. A., Yu, L., Mendiondo, M. & Kryscio, R. J. (2007), 'Shared Random Effects Analysis of Multi-State Markov Models: Application to a Longitudinal Study of Transitions to Dementia', *Statist. Med.* **26**, 568–580.

SAS Institute Inc (1990), *SAS/IML Software: Usage and Reference*, Version 6, 1st edn, SAS Institute Inc., Cary, NC.

Ten, T. R., Miller, M. E., Reboussin, B. A. & James, M. K. (2000), 'Mixed Effects Logistic Regression Models for Longitudinal Ordinal Functional Response Data with Multiple-Cause Drop-Out from the Longitudinal Study of Aging', *Biometrics* **56**, 279–287.

Van der Heijde, D. (1999), 'How to Read Radiographs According to the Sharp/Van Der Heijde Method', *J Rheumatol* **26**, 743–745.

Woodward, M. (1999), *Epidemiology: Study Design and Data Analysis*, Chapman and Hall/CRC, New York.

Apéndice A.
Solución del sistema de ecuaciones de Kolmogorov hacia adelante en el modelo de tres estados

En el modelo de tres estados, el sistema de ecuaciones de Kolmogorov hacia adelante resultante es:

$$\begin{aligned}
p'_{11}(t) &= -(\lambda_{12}+\lambda_{13})p_{11}(t)\\
p'_{12}(t) &= \lambda_{12}p_{11}(t)-\lambda_{23}p_{12}(t)\\
p'_{13}(t) &= \lambda_{13}p_{11}(t)+\lambda_{23}p_{12}(t)\\
p'_{22}(t) &= -\lambda_{23}p_{22}(t)\\
p'_{23}(t) &= \lambda_{23}p_{22}(t) = -\lambda_{23}p_{23}(t)\\
\lambda_{**} &= \lambda_{12}+\lambda_{13}-\lambda_{23}
\end{aligned}$$

Para la primera ecuación, observe que:

$$p'_{11}(t) = -(\lambda_{12}+\lambda_{13})p_{11}(t) \Longleftrightarrow \frac{p'_{11}(t)}{p_{11}(t)} = \exp\big(-(\lambda_{12}+\lambda_{13})t\big)$$

Integrando a ambos lados respecto a t se tiene:

$$\log p_{11}(t) = C-(\lambda_{12}+\lambda_{13})t \Longleftrightarrow p_{11}(t) = \exp(C)\exp\big(-(\lambda_{12}+\lambda_{13})t\big)$$

Por la condición inicial, se tiene que $p_{11}(t)=1$ y así $C=0$. Con esto la solución a esta primera ecuación es $p_{11}(t)=\exp\big(-(\lambda_{12}+\lambda_{13})t\big)$.

Para la segunda ecuación observe que:

$$p'_{12}(t) = \lambda_{12}p_{11}(t)-\lambda_{23}p_{12}(t) \Longleftrightarrow p'_{12}(t)+\lambda_{23}p_{12}(t) = \lambda_{12}p_{11}(t)$$

Multiplicando por $\exp(\lambda_{23}t)$ se tiene que

$$\exp(\lambda_{23}t)p'_{12}(t) + \lambda_{23}\exp(\lambda_{23}t)p_{12}(t) = \lambda_{12}\exp(\lambda_{23}t)p_{11}(t)$$
$$\iff \frac{d}{dt}\big[\exp(\lambda_{23}t)p_{12}(t)\big] = \lambda_{12}\exp(\lambda_{23}t)p_{11}(t)$$

Remplazando $p_{11}(t)$ e integrando respecto a t en ambos lados se tiene que:

$$\exp(\lambda_{23}t)p_{12}(t) = -\frac{\lambda_{12}}{\lambda_{**}}\exp(-\lambda_{**}t) + C, \quad \text{con } \lambda_{**} = \lambda_{12} + \lambda_{13} - \lambda_{23}$$

Como $p_{12}(t) = 0$, entonces

$$\exp(\lambda_{23}t)p_{12}(t) = \frac{\lambda_{12}}{\lambda_{**}}\big[1-\exp(-\lambda_{**}t)\big] \iff p_{12}(t) = \frac{\lambda_{12}}{\lambda_{**}}\big[1-\exp(-\lambda_{**}t)\big]\exp(-\lambda_{23}t)$$

Dado que $p_{11}(t) + p_{12}(t) + p_{13}(t) = 1$, entonces $p_{13}(t) = 1 - p_{11}(t) - p_{12}(t)$.

Para la cuarta ecuación, se tiene que:

$$p'_{22}(t) = -\lambda_{23}p_{22}(t) \iff p_{22}(t) = \exp(C)\exp(-\lambda_{23}t)$$

Por la condición inicial $p_{22}(0) = 1$, con esto se obtiene que

$$p_{22}(t) = \exp(-\lambda_{23}t)$$

Como $p_{22}(t) + p_{23}(t) = 1$, entonces $p_{23}(t) = 1 - p_{22}(t) = 1 - \exp(-\lambda_{23}t)$.