

---

## Enfoque bayesiano para obtener las tasas de transición en un modelo de estados múltiples. Aplicación a datos sobre artritis reumatoide

Bayesian approach for obtaining the transition rates in a multi-state model. Application to Rheumatoid Arthritis data

Juan Carlos Salazar U.<sup>a</sup>  
jcsalaza@unal.edu.co

René Iral P.<sup>b</sup>  
riral@unal.edu.co

Juan Carlos Correa M.<sup>c</sup>  
jccorrea@unal.edu.co

Adriana Rojas V.<sup>d</sup>  
adrirojas@gmail.com

Juan M. Anaya<sup>e</sup>  
juan.anaya@urosario.edu.co

---

### Resumen

Los modelos de estados múltiples han demostrado ser de utilidad para el análisis de datos longitudinales, particularmente aquellos que involucran información acerca de la progresión de una enfermedad a través del tiempo. Por otra parte, los métodos bayesianos son útiles en situaciones de alta complejidad cuando se usan técnicas como Monte Carlo Markov Chain. En este trabajo se implementa un método bayesiano basado en el muestreador de Gibbs con el fin de obtener las tasas de transición que gobiernan un modelo de tres estados con estructura markoviana de primer orden. Estas tasas de transición se vinculan con las covariables por medio de un modelo del tipo Andersen-Gill. De esta manera, la estimación óptima de los efectos de las covariables permitirá obtener mejores estimaciones de las tasas de transición. Esta técnica bayesiana se compara vía simulación con la técnica de estimación estudiada por Iral & Salazar (2007) y con un método basado en la discretización del soporte de la distribución posterior. Finalmente, estas técnicas de estimación se ilustran usando datos reales sobre pacientes colombianos con artritis reumatoide.

---

<sup>a</sup>Profesor Asociado. Universidad Nacional - Sede Medellín, Colombia.

<sup>b</sup>Profesor Asociado. Universidad Nacional - Sede Medellín, Colombia.

<sup>c</sup>Profesor Asociado. Universidad Nacional - Sede Medellín, Colombia.

<sup>d</sup>Profesora Asociada y Coordinadora Centro de Estudio de Enfermedades Autoinmunes (CREA), Universidad del Rosario, Colombia.

<sup>e</sup>Profesor Asociado y Director Centro de Estudio de Enfermedades Autoinmunes (CREA), Universidad del Rosario, Colombia

**Palabras clave:** artritis reumatoide, métodos bayesianos, modelos de estados, muestreador de Gibbs, múltiples, tasas de transición.

### Abstract

Multi-state models have shown to be useful to analyze longitudinal data, especially those involving information about the progression of a disease through time. On the other hand, bayesian methods are useful in highly complex situations where Monte Carlo Markov Chain based techniques are used. In this work, a bayesian method for obtaining the transition rates that govern a three state model with first order markovian structure is implemented. These transition rates are linked to the covariates by means of a Andersen-Gill type model. In this way, the optimal estimation of the covariates effects will improve the estimated values of the transition rates. This bayesian technique based on the Gibbs sampler is compared, using a simulation study, both with a technique proposed by Iral and Salazar and a method based on the discretization of the domain of the posterior distribution. Finally, these techniques are illustrated using real data about Rheumatoid Arthritis collected in Colombian patients.

**Keywords:** bayesian methods, Gibbs sampler, multi-state models, transition rates, rheumatoid arthritis.

## 1. Introducción

De acuerdo con la literatura estadística, se evidencia que los modelos de estados múltiples conforman una importante familia de herramientas estadísticas que son apropiadas para el análisis de datos longitudinales con respuesta categórica (ver por ejemplo, Gao (2004), Ten Have et al. (2000) y Salazar, Schmitt, Yu, Mendiondo & Kryscio (2007)). Este tipo de datos surgen cuando se estudia, por ejemplo, la progresión de una enfermedad incurable, tal como la enfermedad de Alzheimer o artritis reumatoide. Los modelos de estados múltiples aplicados a datos longitudinales han tenido una exitosa acogida en campos de la ciencia tan diversos como la biología, la física, la farmacia, la epidemiología, las ciencias sociales y la medicina (ver Singer & Spilerman (1976), Wasserman (1980), Bartholomew (1983), Selke (1984), Lindsey & Ryan (1993), Commenges (1999), Ten Have et al. (2000), Harezlak et al. (2003), Gao (2004), Vauquelin & Van-Liefde (2005), Tyas et al. (2007), Stephen et al. (2010)). Este trabajo está motivado, básicamente, por los aportes realizados por Kay (1986), Marshall et al. (1995), Frydman (1992), Frydman (1995), Joly & Commenges (1999) y Salazar, Schmitt, Yu, Mendiondo & Kryscio (2007), Salazar, Iral, Calvo, Rojas, Hincapié, Anaya & Díaz (2007).

La información acerca de la progresión de un fenómeno, tal como una enfermedad incurable, usualmente se recolecta por medio de datos longitudinales. Con esto se busca registrar el cambio en el tiempo de una respuesta de interés. Es por esto que obtener las tasas de transición que se asocian con cada uno de estos cambios resulta importante para entender e identificar cuáles factores se relacionan con

el riesgo que un paciente tiene de transitar a través de diferentes estados de un proceso o enfermedad. Por lo tanto, es necesario contar con métodos eficientes que estimen dichas tasas con un grado de precisión aceptable. La metodología bayesiana es muy usada para obtener aproximaciones a los valores de parámetros de interés (ver Hans & Dunson (2005), Gordon (2001)). Una de las ventajas de la metodología bayesiana es que las soluciones que se obtienen no dependen de supuestos asintóticos y todo el trabajo inferencial se realiza usando la distribución posterior. La vigencia y utilidad de dichos métodos justifican su uso para obtener tasas de transición en un modelo de tres estados.

En muchos estudios interesa modelar la manera como los individuos en cierta población transitan de un estado a otro a través del tiempo (ver por ejemplo, Kay (1986), Joly & Commenges (1999) y Salazar, Schmitt, Yu, Mendiondo & Kryscio (2007)). Es muy usual que la información que se recolecta en estos estudios sea de tipo categórico (Aitkin & Alfó (1998)), debido principalmente a que dicha información solo está disponible a través de estados de una enfermedad, los cuales se registran aleatoriamente en el tiempo. La valoración de un estado avanzado en algunas enfermedades puede ser muy compleja y requiere de mucho tiempo. Además el estado en el cual se encuentra un paciente en una visita específica se observa a tiempos irregulares y diferentes para cada paciente. Por estas razones, es necesario usar y proponer metodologías que permitan estudiar, entre otras cosas, la manera como los individuos pasan de un estado a otro de cierta enfermedad, a través de cantidades conocidas como tasas de transición (Bhat (1994), Kao (1997)), pues estas proporcionan información directa del riesgo de transición asociado con cada estado de la enfermedad. En el caso particular de una respuesta categórica (por ejemplo, estados de una enfermedad), las tasas de transición se entiende como la velocidad instantánea de transición de un estado  $i$  a un estado  $j$  en el intervalo de tiempo  $[t, t + \Delta t]$ :

$$\lambda_{ij} = \lim_{\Delta t \rightarrow 0} \frac{P(\text{transición del estado } i \text{ al } j \text{ en } [t, t + \Delta t] \mid \text{estado } i \text{ en el tiempo } t)}{\Delta t}.$$

Algunas veces se opta por asumir que tales tasas de transición son funciones constantes del tiempo debido a que esto permite simplificar la evaluación de la función de verosimilitud, pero es posible generalizar este supuesto y trabajar con tasas de transición dependientes del tiempo, aunque en este último escenario pueden surgir problemas numéricos adicionales a los del proceso de estimación.

El interés de esta propuesta de trabajo está en modelar las tasas relativas de transición entre estados como funciones constantes del tiempo pero que dependen de algunas covariables relacionadas con la evolución de la enfermedad o del fenómeno en cuestión. El contexto básico sobre el cual se fundamenta este trabajo es el siguiente: diferentes participantes son monitoreados en el tiempo a intervalos irregulares y el estado de una enfermedad particular (que puede pensarse como una respuesta categórica) es registrado (estudio longitudinal). Además, se conoce el estado de la enfermedad en un tiempo base para todos los pacientes y el estado en

cada visita particular así como ciertas características propias de cada participante. El problema consiste en estimar las tasas de transición (o tasas de intensidad de transición) de un estado a otro cuando los datos presentan estas características. La importancia de estas tasas radica en el hecho de que ellas cuantifican el riesgo de transición entre estados y los parámetros que se usan para su cálculo permiten cuantificar el efecto de ciertos factores (por ejemplo, genéticos y/o ambientales) en la evolución de una enfermedad particular. Un método que no dependa de supuestos asintóticos, puede mejorar notoriamente la estimación de estas tasas y es precisamente lo que se espera lograr con el uso de métodos bayesianos.

El objetivo general de este trabajo es investigar la precisión en la estimación de las tasas de intensidad asociadas a un modelo de tres estados, usando tres aproximaciones. Es de nuestro interés:

1. Formular y maximizar la función de verosimilitud.
2. Comparar, vía simulación, los resultados obtenidos mediante estrategias bayesianas con el método de estimación basado en el algoritmo de Newton-Raphson estudiado en Iral y Salazar (Iral & Salazar (2006), Iral & Salazar (2007)).
3. Implementar la técnica de estimación bayesiana propuesta usando datos reales.

El presente artículo está organizado en 7 secciones y un apéndice. En la primera sección se hace una breve introducción sobre el tipo de datos a ser trabajados y al modelo de tres estados. En la segunda sección se presenta en más detalle el modelo de tres estados, en el cual se incorpora información de covariables asociadas al fenómeno de estudio, a través de las tasas de intensidad de transición. En la sección tres se muestra cómo se llega a la distribución posterior que será utilizada para obtener las distribuciones condicionales completas, necesarias para la estimación basada en el muestreador Gibbs. En la sección cuatro se describen tres métodos utilizados en la estimación de los parámetros del modelo. En la sección cinco se realiza un estudio de simulación usando los tres métodos descritos en la sección cuatro. En la sección seis se ilustran las metodologías usando datos de artritis reumatoide de un estudio con pacientes colombianos. Finalmente, se presentan algunas conclusiones y recomendaciones.

## 2. Modelo de estados múltiples

Suponga un proceso estocástico  $\{Y(t), t > 0\}$ , con estructura markoviana de primer orden: si el estado de un individuo es conocido para un valor particular del parámetro tiempo  $t$ , esta información es suficiente para predecir el comportamiento del proceso más allá de ese punto. En este proceso se consideran  $k$  estados transitorios donde uno de ellos es absorbente. Se asume que la transición de un

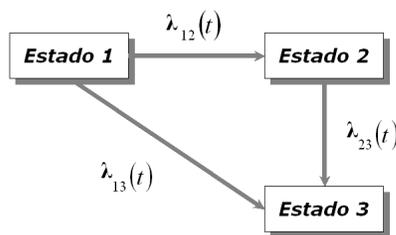


Figura 1: modelo de tres estados de movimiento estrictamente a derecha con un estado absorbente. Fuente: elaboración propia.

estado a otro se hace en una sola dirección (hacia la derecha). Sean,

- $\lambda_{ij}$  = Tasa de transición del estado  $i$  al  $j$
- $P_{ij}$  = Probabilidad de transición del estado  $i$  al  $j$ ,  
dado que se parte del estado  $i$
- $\mathbf{Q}$  =  $[\lambda_{ij}]$  (Matriz de transiciones)
- $\mathbf{P}$  =  $[P_{ij}]$  (Matriz de probabilidades)

La transición entre estados se puede cuantificar por medio de tasas de transición usando un modelo del tipo Andersen-Gill, (Andersen et al. 1993):

$$\lambda_{ij} = \lambda_{ij}^* e^{\sum_{\ell=1}^p \beta_{ij,\ell} x_{\ell}}$$

donde  $\lambda_{ij}^*$  puede verse como una intensidad de referencia;  $\beta_{ij}$  representa el efecto de las variables explicativas  $x$  de acuerdo a la transición. Es posible relacionar las probabilidades de transición con las tasas de transición por medio de un sistema de ecuaciones hacia adelante de Kolmogorov (Bhat (1994), Kao (1997), Chung. (1983)). Para el modelo de tres estados:

Sea  $S = (S_0, S_1, \dots, S_m)$  un vector de estados observados para un sujeto  $w$  en particular durante  $m$  visitas a los tiempos  $t_0, t_1, \dots, t_m$ .

Si se observa  $s = (s_0, s_1, \dots, s_m)$ , por notación,

$$P(s_i | s_0, s_1, \dots, s_{i-1}) = P(S_i = s_i | S_0 = s_0, S_1 = s_1, \dots, S_{i-1} = s_{i-1}) .$$

Usando la propiedad de Markov se tiene que:

$$P(S = s | S_0 = s_0) = P(s_1 | s_0) P(s_2 | s_1) \dots P(s_m | s_{m-1}) .$$

$P_{s_i, s_{i+1}}^{(w)}$  denotará la probabilidad de transición del estado  $s_i$  al estado  $s_{i+1}$  en el intervalo de tiempo  $(t_i, t_{i+1})$  para el sujeto  $w$ .

Para el modelo de tres estados, donde  $\lambda_{ij} = \lambda_{ij}^* e^{\sum_{\ell=1}^p \beta_{ij,\ell} x_{\ell}}$ , las probabilidades de transición se relacionan con las tasas de transición a través de un sistema de ecuaciones hacia adelante de Kolmogorov:

$$\frac{d}{dt} \mathbf{P}(t) = \mathbf{P}(t) \mathbf{Q} ; \mathbf{P}(0) = \mathbf{I}, \text{ con } \mathbf{Q} = [\lambda_{ij}] ,$$

En el modelo de tres estados discutido en este trabajo, las ecuaciones hacia adelante de Kolmogorov se cumplen y están dadas por:

$$\begin{bmatrix} p'_{11}(t) & p'_{12}(t) & p'_{13}(t) \\ 0 & p'_{22}(t) & p'_{23}(t) \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} p_{11}(t) & p_{12}(t) & p_{13}(t) \\ 0 & p_{22}(t) & p_{23}(t) \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} -(\lambda_{12} + \lambda_{13}) & \lambda_{12} & \lambda_{13} \\ 0 & -\lambda_{23} & \lambda_{23} \\ 0 & 0 & 0 \end{bmatrix}$$

Las ecuaciones resultantes y sus soluciones exactas son:

$$\begin{aligned} p'_{11}(t) &= -(\lambda_{12} + \lambda_{13})p_{11}(t) & p_{11}(t) &= \exp(-(\lambda_{12} + \lambda_{13})t) \\ p'_{12}(t) &= \lambda_{12}p_{11}(t) - \lambda_{23}p_{12}(t) & p_{12}(t) &= \frac{\lambda_{12}}{\lambda_{**}} [1 - \exp(-\lambda_{**}t)] \exp(-\lambda_{23}t) \\ p'_{13}(t) &= \lambda_{13}p_{11}(t) + \lambda_{23}p_{12}(t) & p_{13}(t) &= 1 - p_{11}(t) - p_{12}(t) \\ p'_{22}(t) &= -\lambda_{23}p_{22}(t) & p_{22}(t) &= \exp(-\lambda_{23}t) \\ p'_{23}(t) &= \lambda_{23}p_{22}(t) = -\lambda_{23}p_{23}(t) & p_{23}(t) &= 1 - \exp(-\lambda_{23}t) \\ \lambda_{**} &= \lambda_{12} + \lambda_{13} - \lambda_{23} \end{aligned} \quad (1)$$

Por ejemplo, la probabilidad de transición  $p_{11}(t)$ , en términos de las covariables, está dada por:

$$\begin{aligned} p_{11}(t) &= \exp(-(\lambda_{12} + \lambda_{13})t) \\ &= \exp\left(-\left(\lambda_{12}^* e^{\sum_{l=1}^p \beta_{12,l} x_l} + \lambda_{13}^* e^{\sum_{l=1}^p \beta_{13,l} x_l}\right) t\right) \end{aligned}$$

### 3. La verosimilitud

Para la implementación de técnicas bayesianas se asumen las distribuciones *a priori* no informativas. El uso de distribuciones *a priori* no informativas busca que ellas tengan un impacto mínimo sobre la distribución posterior del parámetro de interés y que sean relativamente planas con relación a la verosimilitud. Esto busca que sean los datos los que tengan un claro dominio en la distribución posterior, y, por lo tanto, en todas las inferencias que de ellas se obtengan. Si el espacio parametral es finito se puede utilizar una distribución *a priori* uniforme para reflejar ignorancia total. Si se selecciona una distribución *a priori* impropia o aplanada de la forma  $\xi(\theta) \propto 1$ , con  $\theta = (\lambda_{12}^*, \lambda_{13}^*, \lambda_{23}^*, \beta_{12}, \beta_{13}, \beta_{23})$ , la distribución posterior es proporcional a la función de verosimilitud:

$$\xi(\theta|\text{datos}) \propto L(\theta|\text{datos})$$

Para un sujeto  $w$ , que ha sido monitoreado  $M_w$  veces, la contribución de este sujeto a la función de verosimilitud está dada por (Iral & Salazar (2007)):

$$\prod_{i=1}^{M_w} P_{s_i, s_{i+1}}^{(w)}(t_i, t_{i+1})$$

Por lo tanto, la verosimilitud para los  $n$  sujetos está dada por:

$$\prod_{w=1}^n \prod_{i=1}^{M_w} P_{s_i, s_{i+1}}^{(w)}(t_i, t_{i+1})$$

Con el fin de hacer más explícita esta expresión en términos de las componentes del vector  $\theta$ , considere de nuevo al sujeto  $w$  y denote por  $\mathbf{\Lambda}_{iw}$  a la matriz cuyas componentes son  $\delta_{jk}^{(i,w)}$ , donde:

$$\delta_{jk}^{(i,w)} = \begin{cases} 1, & \text{si en la visita } i \text{ el sujeto } w \text{ pasa del estado } j \text{ al } k \\ 0, & \text{otro caso} \end{cases}$$

Haciendo

$$a_{jk}^{(w)} = \sum_{i=1}^{M_w} \delta_{jk}^{(i,w)} \quad , \quad \text{para } j = 1, 2; k = 1, 2, 3; j \leq k \quad ,$$

En esta expresión,  $k = 3$ , corresponde a un estado absorbente.

Si en una visita particular el sujeto  $w$  pasó del estado  $j$  al  $k$ , la matriz  $\Lambda_{iw}$  asociada a esa visita es cuadrada y tiene ceros excepto en la posición donde  $\delta_{jk} = 1$ .

Sean  $\mathbf{A}$  y  $\mathbf{B}$  matrices cuadradas de orden  $p \times p$  con  $\mathbf{A} = (a_{ij})$  y  $\mathbf{B} = (b_{ij})$ . Defina:

$$\mathbf{A}^{\mathbf{B}} = \left( a_{ij}^{b_{ij}} \right) = \begin{bmatrix} a_{11}^{b_{11}} & \cdots & a_{1p}^{b_{1p}} \\ \vdots & \vdots & \vdots \\ a_{p1}^{b_{p1}} & \cdots & a_{pp}^{b_{pp}} \end{bmatrix}$$

y

$$|\mathbf{A}|_p = \prod_{i=1}^p \prod_{j=1}^p a_{ij} \quad (\text{Producto de todas las entradas de } \mathbf{A})$$

De esta manera se obtiene una expresión general para la función de verosimilitud que involucra todas las probabilidades de transición, lo cual nos permite incorporar distribuciones *a priori* para estimar los parámetros del modelo usando métodos bayesianos.

La verosimilitud para el modelo de tres estados, con movimiento estrictamente a la derecha y un estado absorbente, se puede escribir como:

$$\begin{aligned}
& \prod_{w=1}^n \prod_{i=1}^{M_w} P_{s_i, s_{i+1}}^{(w)}(t_i, t_{i+1}) \\
= & \prod_{w=1}^n \prod_{i=1}^{M_w} |\mathbf{P}^{\Lambda_{iw}}|_P \\
= & \prod_{w=1}^n \prod_{i=1}^{M_w} p_{11}^{\delta_{11}^{(i,w)}} p_{12}^{\delta_{12}^{(i,w)}} p_{13}^{\delta_{13}^{(i,w)}} p_{22}^{\delta_{22}^{(i,w)}} p_{23}^{\delta_{23}^{(i,w)}} \\
= & \prod_{w=1}^n \left\{ p_{11}^{\sum_{i=1}^{M_w} \delta_{11}^{(i,w)}} p_{12}^{\sum_{i=1}^{M_w} \delta_{12}^{(i,w)}} p_{13}^{\sum_{i=1}^{M_w} \delta_{13}^{(i,w)}} p_{22}^{\sum_{i=1}^{M_w} \delta_{22}^{(i,w)}} p_{23}^{\sum_{i=1}^{M_w} \delta_{23}^{(i,w)}} \right\} \\
= & \prod_{w=1}^n \left\{ p_{11}^{a_{11}^{(w)}} p_{12}^{a_{12}^{(w)}} p_{13}^{a_{13}^{(w)}} p_{22}^{a_{22}^{(w)}} p_{23}^{a_{23}^{(w)}} \right\},
\end{aligned}$$

donde cada una de las probabilidades de transición está dada por la solución del sistema de ecuaciones diferenciales hacia adelante de Kolmogorov presentada en la sección anterior. Por lo tanto, la distribución posterior está dada por:

$$\begin{aligned}
\xi(\theta|\text{datos}) & \propto L(\theta|\text{datos}) \\
& \propto \prod_{w=1}^n \left\{ p_{11}^{a_{11}^{(w)}} p_{12}^{a_{12}^{(w)}} p_{13}^{a_{13}^{(w)}} p_{22}^{a_{22}^{(w)}} p_{23}^{a_{23}^{(w)}} \right\}.
\end{aligned}$$

Reemplazando las probabilidades de transición que resultaron de resolver el sistema de ecuaciones diferenciales hacia adelante de Kolmogorov se pueden obtener las densidades condicionales completas (ver apéndice **A**).

## 4. Métodos comparados

### 4.1. Método de Iral y Salazar

Iral & Salazar (2007) proponen optimizar la función de verosimilitud usando el algoritmo de Newton-Raphson en conjunto con el sistema de ecuaciones hacia adelante de Kolmogorov, a través de las soluciones exactas asociadas al modelo de tres estados descritas en la sección anterior.

### 4.2. Discretización del soporte de la posterior

A fin de obtener estimaciones de las tasas de transición se propone un método de estimación bayesiana basado en el método MCMC. El método usado se basa en

la discretización del soporte de la distribución. En otras palabras, si se tiene una variable aleatoria  $X$  con f.d.p.  $f(x)$ , el soporte de  $X$  se define como el siguiente conjunto:  $S = \{x : f(x) > 0\}$ . Para obtener valores al azar de esta distribución, se genera una rejilla sobre  $S$  con valores  $x_0, x_1, x_2, \dots, x_K$ .

Se calculan los valores  $f(x_0), f(x_1), f(x_2), \dots, f(x_K)$ . Luego se extraen muestras con reemplazo de  $\{x_0, x_1, x_2, \dots, x_K\}$  con probabilidades proporcionales a  $\{f(x_0), f(x_1), f(x_2), \dots, f(x_K)\}$ . Para que este procedimiento sea bueno se requiere que la rejilla sea lo suficientemente fina.

### 4.3. Muestreador de Gibbs

El muestreador de Gibbs (ver Casella & George (1992), Ritter & Tanner (1992), Tanner (1996)) es una poderosa herramienta bayesiana para obtener distribuciones posteriores en situaciones complejas, por ejemplo las que surgen cuando se trabaja con modelos de regresión de estados múltiples para datos longitudinales con respuestas categóricas. Con el propósito de obtener una muestra de la distribución conjunta  $p(\theta_1, \dots, \theta_d)$  el muestreador Gibbs se basa en el siguiente algoritmo:

- Muestree  $\theta_1^{(i+1)}$  de  $p\left(\theta_1 \mid \theta_2^{(i)}, \dots, \theta_d^{(i)}\right)$
- Muestree  $\theta_2^{(i+1)}$  de  $p\left(\theta_2 \mid \theta_1^{(i+1)}, \theta_3^{(i)}, \dots, \theta_d^{(i)}\right)$
- $\vdots$
- Muestree  $\theta_d^{(i+1)}$  de  $p\left(\theta_d \mid \theta_1^{(i+1)}, \dots, \theta_{d-1}^{(i+1)}\right)$

Es decir, se muestrea cada variable a partir de la distribución de dicha variable condicionada en las otras variables haciendo uso de los valores más recientes y actualizando la variable con su nuevo valor una vez ha sido muestreado. Este proceso se detiene tan pronto la cadena de Markov alcanza la distribución límite. Sin embargo, en la práctica esta convergencia puede ser muy lenta y el mayor problema es saber si se ha logrado una convergencia razonable (esto se conoce como un *burn-in*). Por lo tanto las muestras obtenidas hasta el punto de *burn-in* son descartadas. En el apéndice A se muestra la manera como se deduce la verosimilitud y posteriormente la obtención de la distribución posterior, de la cual se derivan las densidades condicionales completas.

## 5. Estudio de simulación

Para cada método estudiado se consideraron los siguientes tamaños de muestra 30, 50, 100, 200 y 400, número máximo de visitas  $m = 3$  y 4 y con cada combinación se simuló 1000 veces. Algunos tamaños muestrales no fueron tenidos en cuenta para

el método de Iral y Salazar ya que en los trabajos referenciados no se consideraron esos tamaños por ser esta una metodología que requiere de tamaños muestrales grandes.

Estas muestras se generaron usando unos valores de referencia descritos en un artículo de Harezlak et.al (2003). Para el método de Iral y Salazar, el proceso de estimación se basó en el algoritmo de Newton-Raphson, mientras que para el método basado en la discretización del soporte de la posterior ya no fue necesario optimizar sino muestrear sucesivamente de la distribución posterior.

A fin de implementar el muestreador de Gibbs se usaron cadenas de tamaño 12500 y posteriormente se descartaron los primeros 2500. Esto se realizó para cada uno de los 6 parámetros del vector  $\theta$  definido en la sección 2. Usando el modelo de Andersen-Gill, se calculan los  $\lambda$ 's de interés, por cada grupo de edad, respectivamente. Los grupos de edad se construyeron generando aleatoriamente edades de una distribución normal y definiendo tres categorías ordinales, que por conveniencia se tomaron como 1, 2, 3.

El valor aleatorio para cada uno de los parámetros fue escogido usando la distribución empírica que se generaba a partir de una rejilla previamente establecida. Este proceso se repitió para cada combinación de tamaño de sujetos y de número máximo de monitoreos (o visitas).

Para generar los  $\theta$ 's se utilizó un proceso de Gibbs en combinación con discretización de la distribuciones condicionales completas. Con las matrices resultantes se calcularon las estimaciones de los lambdas para cada grupo de edad y para  $m = 3$  y  $m = 4$ , siendo  $m$  el número máximo de visitas o monitoreos.

### 5.1. Evaluación de una cadena típica ( $n = 100$ , $m = 4$ )

A continuación se presenta el monitoreo de una cadena de Markov típica de las 1000 generadas, discriminadas por grupos de edad. Los gráficos resultantes se muestran en las figuras 3, 4 y 5 respectivamente.

De acuerdo con los gráficos de las cadenas, los ACF y las densidades, no se observan problemas con la estabilidad de las respectivas cadenas de Markov.

Las estimaciones obtenidas para los tres métodos, para  $m = 3$  y  $m = 4$  visitas, se muestran en las tablas 1 y 2.

Para el caso de  $m = 3$  monitoreos o visitas, las estimaciones obtenidas para las tasas de transición  $\lambda_{12}$  y  $\lambda_{13}$  muestran que los métodos bayesianos subestiman los valores de referencia para los tres grupos de edad. El método Iral-Salazar también subestima estos valores a partir de un tamaño de muestra de 200. De las tres metodologías la basada en el muestreador de Gibbs presenta las peores estimaciones, sobre todo para  $\lambda_{13}$ . En el caso de  $\lambda_{23}$  se evidencia que en general los métodos discretización e Iral-Salazar, subestiman los valores de referencia, mientras que el método basado en el muestreador de Gibbs, sobre estima dichos valores de referencia. Estos resultados son muy similares a los que se obtienen con  $m = 4$  visitas. En

Cadena Típica n=100 y m=4, Grupo de Edad 1

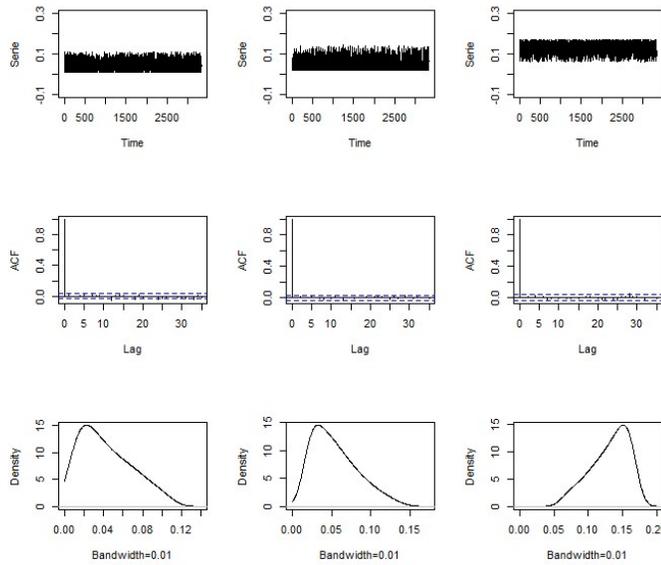


Figura 2: cadena para  $\lambda_{12}$ ,  $\lambda_{13}$  y  $\lambda_{23}$  para grupo de edad 1. Fuente: elaboración propia.

Tabla 1: estimaciones de tasas de intensidad usando los métodos comparados ( $m=3$  visitas). Fuente: elaboración propia.

Descripción	Valores Covariable								
	1			2			3		
n	$\lambda_{12}$	$\lambda_{13}$	$\lambda_{23}$	$\lambda_{12}$	$\lambda_{13}$	$\lambda_{23}$	$\lambda_{12}$	$\lambda_{13}$	$\lambda_{23}$
<b>30</b>									
Harezlak	0.0538	0.0760	0.1238	0.0574	0.0803	0.1335	0.0612	0.0848	0.1439
Iral	---	---	---	---	---	---	---	---	---
Discretizacion	0.0506	0.0726	0.1189	0.0539	0.0766	0.1279	0.0574	0.0808	0.1375
Gibbs	0.0421	0.0551	0.1318	0.0444	0.0590	0.1428	0.0469	0.0631	0.1548
<b>50</b>									
Harezlak	0.0538	0.0760	0.1238	0.0574	0.0803	0.1335	0.0612	0.0848	0.1439
Iral	---	---	---	---	---	---	---	---	---
Discretizacion	0.0486	0.0706	0.1174	0.0517	0.0744	0.1262	0.0551	0.0785	0.1357
Gibbs	0.0420	0.0551	0.1317	0.0443	0.0590	0.1428	0.0468	0.0631	0.1548
<b>100</b>									
Harezlak	0.0538	0.0760	0.1238	0.0574	0.0803	0.1335	0.0612	0.0848	0.1439
Iral	0.0674	0.0931	0.1411	0.0486	0.0729	0.1134	0.0475	0.0730	0.1050
Discretizacion	0.0491	0.0699	0.1144	0.0522	0.0737	0.1228	0.0556	0.0777	0.1319
Gibbs	0.0420	0.0551	0.1317	0.0443	0.0589	0.1427	0.04678	0.0631	0.1547
<b>200</b>									
Harezlak	0.0538	0.0760	0.1238	0.0574	0.0803	0.1335	0.0612	0.0848	0.1439
Iral	0.0585	0.0792	0.1309	0.0469	0.0713	0.1144	0.0449	0.0721	0.1056
Discretizacion	0.0455	0.0693	0.1109	0.0484	0.0730	0.1189	0.0514	0.0769	0.1276
Gibbs	0.0420	0.0551	0.1318	0.0443	0.0589	0.1428	0.0468	0.0631	0.1548
<b>400</b>									
Harezlak	0.0538	0.0760	0.1238	0.0574	0.0803	0.1335	0.0612	0.0848	0.1439
Iral	0.0507	0.0764	0.1249	0.0448	0.0720	0.1136	0.0432	0.0716	0.1059
Discretizacion	0.0438	0.0680	0.1074	0.0465	0.0716	0.1150	0.0495	0.0754	0.1232
Gibbs	0.0420	0.0551	0.1318	0.0443	0.0589	0.1428	0.0468	0.0631	0.1548

Cadena Típica n=100 y m=4, Grupo de Edad 2

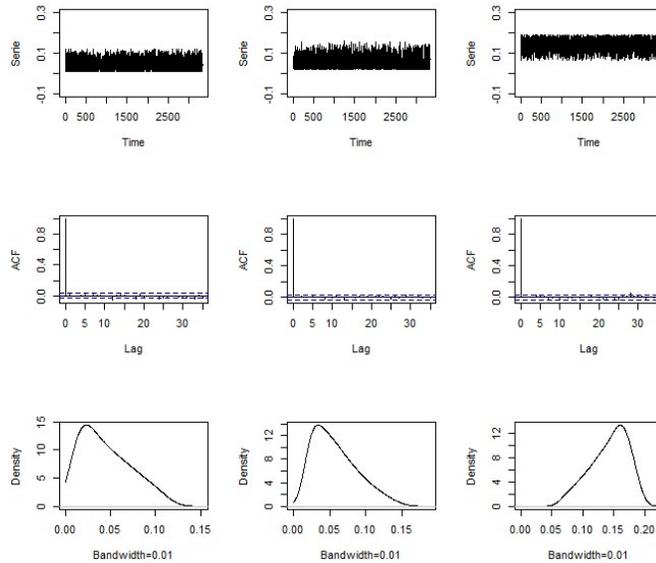


Figura 3: cadena para  $\lambda_{12}$ ,  $\lambda_{13}$  y  $\lambda_{23}$  para grupo de edad 2. Fuente: elaboración propia.

Tabla 2: estimaciones de tasas de intensidad usando los métodos comparados ( $m=4$  visitas). Fuente: elaboración propia.

Descripción	Valores Covariable								
	1			2			3		
n	$\lambda_{12}$	$\lambda_{13}$	$\lambda_{23}$	$\lambda_{12}$	$\lambda_{13}$	$\lambda_{23}$	$\lambda_{12}$	$\lambda_{13}$	$\lambda_{23}$
<b>30</b>									
Harezlak	0.0538	0.0760	0.1238	0.0574	0.0803	0.1335	0.0612	0.0848	0.1439
Iral	---	---	---	---	---	---	---	---	---
Discretizacion	0.0497	0.0730	0.1200	0.0529	0.0769	0.1290	0.0563	0.0812	0.1388
Gibbs	0.0420	0.0552	0.1318	0.0443	0.0590	0.1428	0.0468	0.0632	0.1549
<b>50</b>									
Harezlak	0.0538	0.0760	0.1238	0.0574	0.0803	0.1335	0.0612	0.0848	0.1439
Iral	---	---	---	---	---	---	---	---	---
Discretizacion	0.0510	0.0737	0.1181	0.0543	0.0777	0.1269	0.0578	0.0820	0.1365
Gibbs	0.0420	0.0551	0.1317	0.0443	0.0589	0.1428	0.0468	0.0631	0.1548
<b>100</b>									
Harezlak	0.0538	0.0760	0.1238	0.0574	0.0803	0.1335	0.0612	0.0848	0.1439
Iral	0.0708	0.0957	0.1377	0.0530	0.0782	0.1195	0.0538	0.0791	0.1172
Discretizacion	0.0496	0.0718	0.1164	0.0528	0.0757	0.1251	0.0562	0.0798	0.1344
Gibbs	0.0420	0.0551	0.1318	0.0443	0.0590	0.1428	0.0468	0.0631	0.1549
<b>200</b>									
Harezlak	0.0538	0.0760	0.1238	0.0574	0.0803	0.1335	0.0612	0.0848	0.1439
Iral	0.0583	0.0805	0.1302	0.0499	0.0750	0.1213	0.0501	0.0781	0.1191
Discretizacion	0.0490	0.0710	0.1151	0.0521	0.0748	0.1237	0.0555	0.0789	0.1328
Gibbs	0.0420	0.0551	0.1318	0.0443	0.0590	0.1428	0.0468	0.0631	0.1549
<b>400</b>									
Harezlak	0.0538	0.0760	0.1238	0.0574	0.0803	0.1335	0.0612	0.0848	0.1439
Iral	0.0546	0.0752	0.1290	0.0502	0.0739	0.1219	0.0499	0.0768	0.1182
Discretizacion	0.0460	0.0708	0.1119	0.0490	0.0746	0.1200	0.0521	0.0787	0.1288
Gibbs	0.0420	0.0551	0.1318	0.0443	0.0590	0.1428	0.0468	0.0631	0.1549

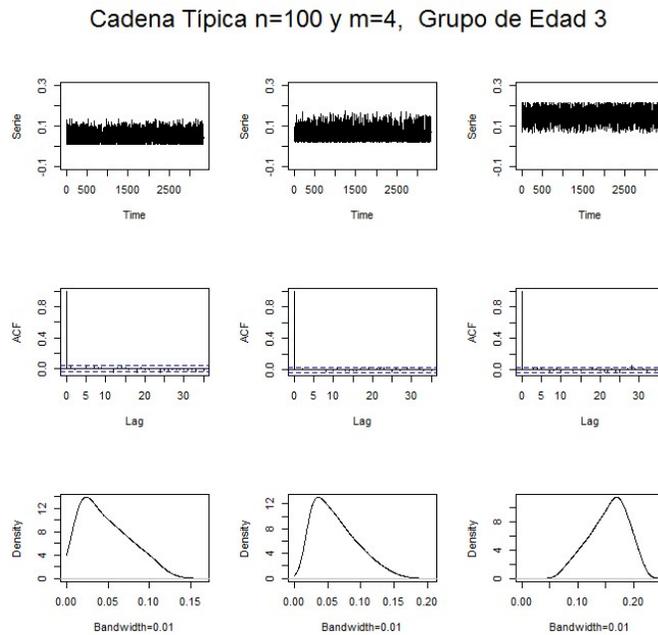


Figura 4: cadena para  $\lambda_{12}$ ,  $\lambda_{13}$  y  $\lambda_{23}$  para grupo de edad 3. Fuente: elaboración propia.

general, el método de Discretización proporcionó mejores aproximaciones a los valores de referencia que el método basado en el muestreador de Gibbs, para las tasas  $\lambda_{12}$  y  $\lambda_{13}$ , pero la discrepancia aumentaba a medida que se aumentaba el tamaño de la muestra. Para  $\lambda_{23}$  el muestreador de Gibbs proporcionó mejores aproximaciones que los demás. Cabe anotar que las aproximaciones proporcionadas por el método Iral-Salazar, fueron en promedio mejores que las demás aproximaciones para tamaños de muestra superiores a 100.

## 6. Aplicación: progresión radiográfica en artritis reumatoide

Los datos utilizados para ilustrar los métodos incluyen 464 radiografías de 146 pacientes diagnosticados con artritis reumatoide, 84.9% mujeres. La edad promedio de los pacientes era de 47.1 años con una desviación estándar de 13.4 años. Se tenía un promedio de 3 radiografías por paciente de manos y pies (Rojas-Villarraga et al. 2009). Los datos fueron proporcionados por el grupo CREA de la Universidad del Rosario.

Estos pacientes fueron monitoreados durante un promedio de  $4.2 \pm 1.6$  años por el mismo equipo médico y siguiendo un régimen terapéutico estándar. Las radio-

grafías de cada paciente fueron leídas y evaluadas por dos profesionales calificados siguiendo el método de Sharp van der Heijde (Van-der Heijde et al. 1999). El objetivo consistía en registrar el promedio de erosiones en manos y pies. El puntaje de erosión, correspondiente al promedio de manos y pies de acuerdo con el método de Sharp van der Heijde, fue categorizado en tres niveles en relación con su severidad: 1=Leve ( $\leq 4$ ), 2=Moderado ( $> 4$  y  $\leq 16$ ) y 3=Severo ( $> 16$ ) (modelo de tres estados).

Debido a la naturaleza progresiva de la enfermedad la transición del estado leve al severo no es posible y por lo tanto esta ilustración es un caso particular del modelo expuesto en el estudio de simulación (modelo de tres estados de movimiento estrictamente a derecha).

Para la estimación de las funciones de intensidad ( $\lambda_{LM}, \lambda_{MS}$ ) 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 reumatoide (ver Delgado-Vega et al. (2006) y Anaya et al. (2006)).

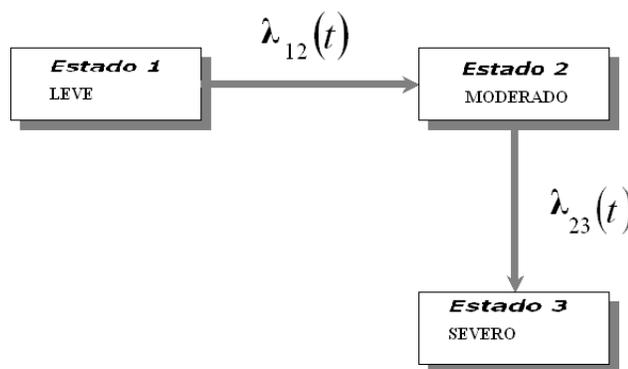


Figura 5: *modelo de tres estados con un estado absorbente para artritis reumatoide.*  
Fuente: elaboración propia.

En la Tabla 3 se observan los valores estimados de las funciones de intensidad para los grupos con presencia y ausencia de la secuencia genética del SE. para los métodos estudiados (N=146).

Al comparar los resultados obtenidos para las tres metodologías, se observa que en ausencia de SE, los resultados son muy similares. La mayor discrepancia se observa entre las tasas de transición de Moderado a Severo donde el método bayesiano de discretización de la posterior tiende a subestimar las tasas de transición tal y como se observó en el estudio de simulación.

Tabla 3: estimación de las tasas de transición (error estándar)  $\lambda_{LM}$  y  $\lambda_{MS}$  .  
Fuente: elaboración propia.

Función de intensidad estimada usando el método de Iral y Salazar	Ausencia de SE	Presencia de SE
<i>Leve</i> $\rightarrow$ <i>Moderado</i> , ( $\hat{\lambda}_{LM}$ )	0.117 (0.028)	0.110 (0.029)
<i>Moderado</i> $\rightarrow$ <i>Severo</i> , ( $\hat{\lambda}_{MS}$ )	0.097 (0.042)	0.264 (0.078)
Función de intensidad estimada usando discretización de la posterior	Ausencia de SE	Presencia de SE
<i>Leve</i> $\rightarrow$ <i>Moderado</i> , ( $\hat{\lambda}_{LM}$ )	0.110 (0.022)	0.093 (0.016)
<i>Moderado</i> $\rightarrow$ <i>Severo</i> , ( $\hat{\lambda}_{MS}$ )	0.09 (0.020)	0.204 (0.034)
Función de intensidad estimada usando muestreador de Gibbs	Ausencia de SE	Presencia de SE
<i>Leve</i> $\rightarrow$ <i>Moderado</i> , ( $\hat{\lambda}_{LM}$ )	0.124 (0.028)	0.117 (0.029)
<i>Moderado</i> $\rightarrow$ <i>Severo</i> , ( $\hat{\lambda}_{MS}$ )	0.102 (0.043)	0.307 (0.135)

## 7. Conclusiones

El muestreador de Gibbs presentó mucha consistencia para todos los tamaños muestrales y número de medidas repetidas.

Las tasas de intensidad obtenidas usando el muestreador de Gibbs tienden a subestimar los valores de referencia para  $\lambda_{12}$  y para  $\lambda_{13}$  y a sobreestimar los correspondientes a  $\lambda_{23}$ , para todas las combinaciones de tamaños de muestra y número de visitas

Los errores estándar asociados al muestreador de Gibbs fueron inferiores a los obtenidos con el método de discretización y al propuesto por Iral & Salazar.

Los métodos bayesiano aquí estudiados son efectivos aún con tamaños de muestra pequeños. Este comportamiento no pudo ser evaluado en el método de Iral y Salazar para tamaños de muestra pequeño, debido a que se requiere de tamaños de muestra grandes para su implementación.

Si se van a evaluar varias covariables y se dispone de un buen tamaño muestral (al menos 100 unidades) es aconsejable usar el algoritmo de Iral y Salazar, ya que no requiere de demasiadas ajustes en valores iniciales a medida que se consideran covariables adicionales.

## 8. Agradecimientos

Este trabajo fue realizado con recursos del proyecto de investigación 9201 de la Vicerrectoría de Investigación de la Universidad Nacional de Colombia, código DIME 20101007954. A la Facultad de Ciencias de la Universidad Nacional de Colombia,

Sede Medellín. Los autores también agradecen al CREA de la Universidad del Rosario por facilitar los datos usados en el ejemplo de aplicación.

Recibido: 29 de abril de 2014  
Aceptado: 25 de agosto de 2014

## Referencias

- Aitkin, M. & Alfó, M. (1998), 'Regression models for binary longitudinal responses', *Statistics and Computing* **8**, 289–307.
- Anaya, J. M., Pineda-Tamayo, R., Gómez, L. M., Galarza-Maldonado, C., Rojas-Villarraga, A. & Martín, J. (2006), *Artritis Reumatoide: Bases Moleculares, Clínicas y Terapéuticas*, 1 edn, FUNPAR, Medellín.
- Andersen, P. K., Borgan, o., Gill, R. D. & Keiding, N. (1993), *Statistical Models Based on Counting Processes*, 1 edn, Springer-Verlag, New York.
- Bartholomew, D. J. (1983), 'Some recent developments in social statistics', *International Statistical Review* **51**, 1–9.
- Bhat, U. N. (1994), *Elements of applied stochastic processes*, 2 edn, Wiley, New York.
- Casella, G. & George, E. I. (1992), 'Explaining the Gibbs sampler.', *The American Statistician* **46**, 167–174.
- Chung., K. L. (1983), *Teoría elemental de la probabilidad y de los procesos estocásticos*, Reverté, Barcelona.
- Commenges, D. (1999), 'Multi-State Models in Epidemiology', *Lifetime Data Analysis* **5**, 315–327.
- Delgado-Vega, 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.
- Frydman, H. (1992), 'A Nonparametric estimation procedure for a periodically observed three state Markov process, with application to AIDS', *J.R Statist. Soc B* **54**, 853–866.
- 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', *Biometrics* **23**, 211–219.

- Gordon, P. (2001), *Bayesian statistical modelling*, 1 edn, John Wiley & Sons, Chichester.
- Hans, C. & Dunson, D. B. (2005), 'Bayesian inference on umbrella orderings', *Biometrics* **61**, 1018–1026.
- Harezlak, J., Gao, S. & Hui, S. L. (2003), 'An illness-death scholastic model in the analysis of longitudinal dementia data', *Statistics in Medicine* **22**, 1465–1475.
- 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', Bucaramanga, Colombia, Universidad Nacional de Colombia.
- Iral, R. & Salazar, J. C. (2007), Estimación de funciones de intensidad en un modelo de Markov de tres estados bajo el efecto de covariables con datos longitudinales, Master's thesis, Escuela de Estadística, Universidad Nacional de Colombia, Sede Medellín.
- 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.
- Kao, E. P. (1997), *An introduction to stochastic processes*, Duxbury Press, Belmont.
- Kay, R. (1986), 'A Markov Model for Analyzing Cancer Markers and Disease States in Survival Studies', *Biometrics* **42**(4), 855–865.
- Lindsey, J. C. & Ryan, L. M. (1993), 'A three-state multiplicative model for rodent tumorigenicity experiments', *Applied Statistics* **42**, 283–300.
- 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.
- Ritter, C. & Tanner, M. A. (1992), 'The Gibbs stopper and the gridy Gibbs sampler', *Journal of the American Statistical Association* **87**, 861–868.
- Rojas-Villarraga, A., Diaz, F., Calvo, E., Salazar, J. C., Iglesias-Gamarra, A., Mantilla, R. D. & Anaya, J. M. (2009), 'Familial disease, the HLA-DRB1 shared epitope and anti-CCP antibodies influence time at appearance of substantial joint damage in rheumatoid arthritis', *Journal of Autoimmunity* **32**, 64–69.
- Salazar, J. C., Iral, R., Calvo, E., Rojas, A., Hincapié, M. E., Anaya, J. M. & Díaz, F. J. (2007), 'Modelo de Markov de tres estados: comparación de parametrizaciones de la tasa de intensidad de transición. Aplicación a datos de artritis reumatoidea', *Revista Colombiana de Estadística* **30**(2), 213–229.

- Salazar, J. C., Schmitt, F. A., Yu, L., Mendiondo, M. & Kryscio, R. J. (2007), ‘Shared random effects analysis of multistate Markov models: application to a longitudinal study of transitions to dementia’, *Statist. Med.* **3**(26), 568–580.
- Selke, W. (1984), ‘Monte Carlo studies of interfacial adsorption in multi-state models’, *Surface Science* **144**(1), 176–181.
- Singer, B. & Spilerman, S. (1976), ‘The representation of social processes by Markov models’, *American Journal of Sociology* **82**, 1–54.
- Stephen, D. U., Kenneth, I. J., Robert, F. L., Peter, P. L., Scott, E. S. & Mychailo, B. T. (2010), Physics-Based Stress Corrosion Cracking Component Reliability Model cast in an R7-Compatible Cumulative Damage Framework, Technical Report PNNL-20596, US Department of Energy.
- Tanner, M. A. (1996), *Tools for statistical inference: Methods for the exploration of posterior distributions and likelihood functions*, 3 edn, Springer, New York.
- Ten Have, 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.
- Tyas, S. L., Salazar, J. C., Snowdon, D. A., Desrosiers, M. F., Riley, K. P., Mendiondo, M. S. & Kryscio, R. J. (2007), ‘Transitions to Mild Cognitive Impairments, Dementia, and Death: Findings from the Nun Study’, *American Journal of Epidemiology* **165**(11), 1231–1238.
- Van-der Heidje, D., Dankert, T. & Nieman, F. (1999), ‘Reliability and Sensitivity to Change of a Simplification of the Shap van der Heidje radiological assessment in Rheumatoid Arthritis’, *Rheumatology* **38**, 938–941.
- Vauquelin, G. & Van-Liefde, I. (2005), ‘G protein-coupled receptors: a count of 1001 conformations’, *Fundamental & clinical pharmacology* **19**(1), 45–56.
- Wasserman, S. (1980), ‘Analyzing social networks as stochastic processes’, *Journal of the American Statistical Association* **75**, 280–294.

## A. Apéndice

Considere un modelo de tres estados, donde el tercer estado es absorbente (la probabilidad de transición a otro estado, dado que se está en este estado es cero). Se supone una transición estrictamente a derecha. Las soluciones al sistema de ecuaciones hacia adelante de Kolmogorov se muestran en la sección 2. Usando un modelo multiplicativo tipo Andersen-Gill con una covariable, se obtienen las siguientes expresiones en términos del vector  $\theta = (\lambda_{12}^*, \lambda_{13}^*, \lambda_{23}^*, \beta_{12}, \beta_{13}, \beta_{23})$ :

$$p_{11}(t) = \exp\{-\{\lambda_{12}^* e^{\beta_{12} X} + \lambda_{13}^* e^{\beta_{13} X}\} t\}$$

$$\begin{aligned}
p_{12}(t) &= \left[ \frac{\lambda_{12}^* e^{\beta_{12} X}}{\lambda_{12}^* e^{\beta_{12} X} + \lambda_{13}^* e^{\beta_{13} X} - \lambda_{23}^* e^{\beta_{23} X}} \right] \\
&\times \left[ 1 - \exp\{-\{\lambda_{12}^* e^{\beta_{12} X} + \lambda_{13}^* e^{\beta_{13} X} - \lambda_{23}^* e^{\beta_{23} X}\} t\} \right] \exp\{-\lambda_{23}^* t e^{\beta_{23} X}\} \\
p_{13}(t) &= 1 - \exp\{-\{\lambda_{12}^* e^{\beta_{12} X} + \lambda_{13}^* e^{\beta_{13} X}\} t\} \\
&- \left[ \frac{\lambda_{12}^* e^{\beta_{12} X}}{\lambda_{12}^* e^{\beta_{12} X} + \lambda_{13}^* e^{\beta_{13} X} - \lambda_{23}^* e^{\beta_{23} X}} \right] \\
&\times \left[ 1 - \exp\{-\{\lambda_{12}^* e^{\beta_{12} X} + \lambda_{13}^* e^{\beta_{13} X} - \lambda_{23}^* e^{\beta_{23} X}\} t\} \right] \exp\{-\lambda_{23}^* t e^{\beta_{23} X}\} \\
p_{22}(t) &= \exp\{-\lambda_{23}^* t e^{\beta_{23} X}\} \quad ; \quad p_{23}(t) = 1 - \exp\{-\lambda_{23}^* t e^{\beta_{23} X}\} .
\end{aligned}$$

Haciendo

$$\delta_{jk}^{(i,w)} = \begin{cases} 1, & \text{si en la visita } i \text{ el sujeto } w \text{ pasa del estado } j \text{ al } k \\ 0, & \text{otro caso} \end{cases}$$

y

$$a_{jk}^{(w)} = \sum_{i=1}^{M_w} \delta_{jk}^{(i,w)} \quad , \quad \text{para } j = 1, 2; k = 1, 2, 3;$$

la función de verosimilitud posterior se expresa como:

$$L(\boldsymbol{\theta} | X) = \prod_{w=1}^n \left\{ p_{11}^{a_{11}^{(w)}} p_{12}^{a_{12}^{(w)}} p_{13}^{a_{13}^{(w)}} p_{22}^{a_{22}^{(w)}} p_{23}^{a_{23}^{(w)}} \right\} ,$$

Reemplazando las soluciones descritas anteriormente en esta última expresión, se obtiene la función de verosimilitud posterior:

$$\begin{aligned}
&\xi(\boldsymbol{\theta} | X_w) \\
&\propto \prod_{w=1}^n \exp\left\{-\left\{\lambda_{12}^* e^{\beta_{12} X_w} + \lambda_{13}^* e^{\beta_{13} X_w}\right\} t_w a_{11}^{(w)}\right\} \times A_w^{a_{12}^{(w)}} \times B_w^{a_{12}^{(w)}} \\
&\times \exp\left\{-\lambda_{23}^* t_w a_{12}^{(w)} e^{\beta_{23} X_w}\right\} \times C_w^{a_{13}^{(w)}} \times \exp\left\{-\lambda_{23}^* t_w a_{22}^{(w)} e^{\beta_{23} X_w}\right\} \\
&\times \left[1 - \exp\left\{-\lambda_{23}^* t_w e^{\beta_{23} X_w}\right\}\right]^{a_{23}^{(w)}} \\
\xi(\boldsymbol{\theta} | X_w) &= \prod_{w=1}^n \exp\left\{-\lambda_{12}^* t_w a_{11}^{(w)} e^{\beta_{12} X_w}\right\} \times \exp\left\{-\lambda_{13}^* t_w a_{11}^{(w)} e^{\beta_{13} X_w}\right\} \\
&\times \exp\left\{-\lambda_{23}^* t_w \left(a_{12}^{(w)} + a_{22}^{(w)}\right) e^{\beta_{23} X_w}\right\} \\
&\times A_w^{a_{12}^{(w)}} \times B_w^{a_{12}^{(w)}} \times C_w^{a_{13}^{(w)}} \times \left[1 - \exp\left\{-\lambda_{23}^* t_w e^{\beta_{23} X_w}\right\}\right]^{a_{23}^{(w)}}
\end{aligned}$$

donde:

$$A_w = \left[ \frac{\lambda_{12}^* e^{\beta_{12} X_w}}{\lambda_{12}^* e^{\beta_{12} X_w} + \lambda_{13}^* e^{\beta_{13} X_w} - \lambda_{23}^* e^{\beta_{23} X_w}} \right] ,$$

$$B_w = 1 - \exp \left\{ - \{ \lambda_{12}^* e^{\beta_{12} X_w} + \lambda_{13}^* e^{\beta_{13} X_w} - \lambda_{23}^* e^{\beta_{23} X_w} \} t_w \right\} ,$$

$$C_w = 1 - \exp \left\{ - \{ \lambda_{12}^* e^{\beta_{12} X_w} + \lambda_{13}^* e^{\beta_{13} X_w} \} t_w \right\} - A_w B_w \exp \{ - \lambda_{23}^* t_w e^{\beta_{23} X_w} \} .$$

Las expresiones para las densidades condicionales completas están dadas por:

$$\xi(\lambda_{12}^* | \lambda_{13}^*, \lambda_{23}^*, \beta_{12}, \beta_{13}, \beta_{23}, X) = \exp \left\{ - \lambda_{12}^* \sum_{w=1}^n t_w a_{11}^{(w)} e^{\beta_{12} X_w} \right\} \times \prod_{w=1}^n A_w^{a_{12}^{(w)}} B_w^{a_{12}^{(w)}} C_w^{a_{13}^{(w)}} .$$

$$\xi(\lambda_{13}^* | \lambda_{12}^*, \lambda_{23}^*, \beta_{12}, \beta_{13}, \beta_{23}, X) = \exp \left\{ - \lambda_{13}^* \sum_{w=1}^n t_w a_{11}^{(w)} e^{\beta_{13} X_w} \right\} \times \prod_{w=1}^n A_w^{a_{12}^{(w)}} B_w^{a_{12}^{(w)}} C_w^{a_{13}^{(w)}} .$$

$$\begin{aligned} \xi(\lambda_{23}^* | \lambda_{12}^*, \lambda_{13}^*, \lambda_{23}^*, \beta_{12}, \beta_{13}, \beta_{23}, X) &= \exp \left\{ - \lambda_{23}^* \sum_{w=1}^n t_w (a_{12}^{(w)} + a_{22}^{(w)}) e^{\beta_{23} X_w} \right\} \\ &\times \prod_{w=1}^n [1 - \exp \{ - \lambda_{23}^* t_w e^{\beta_{23} X_w} \}] A_w^{a_{12}^{(w)}} B_w^{a_{12}^{(w)}} C_w^{a_{13}^{(w)}} \end{aligned}$$

$$\xi(\beta_{12} | \lambda_{12}^*, \lambda_{13}^*, \lambda_{23}^*, \beta_{13}, \beta_{23}, X) = \exp \left\{ - \lambda_{12}^* \sum_{w=1}^n t_w a_{11}^{(w)} e^{\beta_{12} X_w} \right\} \times \prod_{w=1}^n A_w^{a_{12}^{(w)}} B_w^{a_{12}^{(w)}} C_w^{a_{13}^{(w)}} .$$

$$\xi(\beta_{13} | \lambda_{12}^*, \lambda_{13}^*, \lambda_{23}^*, \beta_{12}, \beta_{23}, X) = \exp \left\{ - \lambda_{13}^* \sum_{w=1}^n t_w a_{11}^{(w)} e^{\beta_{13} X_w} \right\} \times \prod_{w=1}^n A_w^{a_{12}^{(w)}} B_w^{a_{12}^{(w)}} C_w^{a_{13}^{(w)}} .$$

$$\begin{aligned} \xi(\beta_{23} | \lambda_{12}^*, \lambda_{13}^*, \lambda_{23}^*, \beta_{12}, \beta_{13}, X) &= \exp \left\{ - \lambda_{23}^* \sum_{w=1}^n t_w (a_{12}^{(w)} + a_{22}^{(w)}) e^{\beta_{23} X_w} \right\} \\ &\times \prod_{w=1}^n [1 - \exp \{ - \lambda_{23}^* t_w e^{\beta_{23} X_w} \}] A_w^{a_{12}^{(w)}} B_w^{a_{12}^{(w)}} C_w^{a_{13}^{(w)}} \end{aligned}$$