



"El saber de mis hijos
hará mi grandeza"

UNIVERSIDAD DE SONORA

DIVISIÓN DE CIENCIAS EXACTAS Y NATURALES

Programa de Licenciado en Matemáticas

Estimación en procesos markovianos de salto

T E S I S

Que para obtener el título de:

Licenciado en Matemáticas

Presenta:

Francisco Figueroa Ramírez

Director de Tesis: Dr. José Arturo Montoya Laos

Asesor de Tesis: Dr. Oscar Vega Amaya

Hermosillo, Sonora, México, Enero, 2013.

Universidad de Sonora

Repositorio Institucional UNISON



**"El saber de mis hijos
hará mi grandeza"**



Excepto si se señala otra cosa, la licencia del ítem se describe como openAccess

SINODALES

Dra. Gudelia Figueroa Preciado
Universidad de Sonora

Dr. Jesús Adolfo Minjárez Sosa
Universidad de Sonora

Dr. José Arturo Montoya Laos
Universidad de Sonora

Dr. Oscar Vega Amaya
Universidad de Sonora

Agradecimientos

En primer lugar a Dios, que me ha brindado una vida llena de alegrías y aprendizaje, permitiéndome vivir una muy grata experiencia en mi etapa universitaria.

A toda mi familia que siempre me ha apoyado, guiado y cuidado con mucho amor, en especial a mis padres, Francisco y Cecilia, gracias de corazón por todas las oportunidades que me han brindado, espero estén orgullosos de mi. A mis hermanos, Oscar y Cecilia, por la compañía y el apoyo que me brindaron de una forma u otra, se que cuento con ellos siempre. Soy afortunado por contar siempre con su amor, comprensión y ejemplo. Esta tesis es suya.

Quiero agradecerles también a mi director de tesis, Dr. José Arturo Montoya Laos, y asesor de tesis, Dr. Oscar Vega Amaya, por sus enseñanzas, su paciencia y por haberme dedicado su tiempo, siempre con entusiasmo, por su apoyo y amistad que me permitieron aprender más que lo estudiado en el trabajo.

Muchas gracias a mis sinodales Dra. Gudelia Figueroa Preciado y Dr. Jesús Adolfo Minjárez Sosa, por haber destinado parte de su tiempo a la revisión de este trabajo, gracias por sus observaciones y correcciones.

A todos mis profesores de la licenciatura que me enseñaron tanto de la profesión como de la vida, impulsándome siempre a seguir adelante. Estoy muy agradecido por ayudarme a enriquecer mis conocimientos, gracias también por su amistad.

A todos mis amigos con quienes he compartido muchos momentos que siempre llevaré conmigo. Ustedes han enriquecido mi vida con su cariño y su alegría, y todos los personajes que han dejado huella en mi vida. Ustedes han sido fuente de alegría y muchos de ustedes fueron de gran apoyo.

*Francisco Figueroa Ramírez
Enero de 2013*

Índice general

Introducción	1
1. El proceso de Poisson	3
1.1. Procesos de conteo	3
1.2. Definición del proceso de Poisson	4
1.3. Distribuciones de tiempos de espera y entre arribos	9
1.4. Otras propiedades del proceso de Poisson	11
2. Procesos markovianos de salto	13
2.1. Construcción de los procesos de salto	13
2.2. Procesos de nacimiento y muerte	19
2.2.1. Procesos de ramificación	20
2.2.2. Proceso de nacimiento y muerte con dos estados	22
2.2.3. Proceso de Poisson	23
2.2.4. Proceso de nacimiento	25
3. Inferencia estadística	27
3.1. La función de verosimilitud	27
3.1.1. Plausibilidad	29
3.1.2. Estimador de máxima verosimilitud (emv)	29
3.2. Invarianza	31
3.3. Función de verosimilitud relativa	31
3.4. Regiones de verosimilitud	33
3.4.1. Intervalos de verosimilitud	33
3.5. Regiones de verosimilitud-confianza	34
4. Enfoque de verosimilitud para hacer inferencias en procesos markovianos de salto	37
4.1. Inferencias en un proceso de Poisson homogéneo: Caso sin censura	37
4.2. Inferencias en un proceso de Poisson homogéneo: Caso con censura	39
4.3. Inferencias en un proceso de nacimiento lineal: Caso sin censura	41
4.4. Inferencias en un proceso de nacimiento lineal: Caso con censura	44
4.5. Estudio de simulación	46

4.5.1. Parte 1: proceso de Poisson homogéneo	48
4.5.2. Parte 2: proceso de nacimiento lineal	48
5. Conclusiones	51
Bibliografía	53

Introducción

Los procesos estocásticos son una herramienta fundamental para el análisis de fenómenos reales que fluctúan aleatoriamente en el tiempo. Por ejemplo, la cantidad de clientes que entran a una tienda, el número de coches que pasan por una autopista, la llegada de personas a una fila de espera, el número de llamadas que llegan a una central telefónica, el número de partículas emitidas por un material radiactivo, entre otros. Un proceso estocástico es una familia de variables aleatorias $X(t)$ indexadas por un índice t , con $t \in T \subseteq \mathbb{R}$. Cada una de las variables aleatorias del proceso tiene su propia función de distribución de probabilidad y entre ellas pueden estar correlacionadas o no. Una clase muy importante de este tipo de procesos la constituyen los procesos de Markov, también llamados procesos markovianos. Estos tienen la característica general que dado el valor actual del proceso $X(t)$, los valores futuros $X(s)$ para $s > t$ son independientes de los valores pasados $X(u)$ para $u < t$. Dos procesos markovianos ampliamente estudiados en la literatura de la Probabilidad son: el proceso de Poisson y el proceso de nacimiento lineal.

En la mayoría de las aplicaciones reales los parámetros involucrados en los modelos markovianos no pueden ser establecidos a priori y tienen que ser inferidos a partir de un comportamiento histórico del proceso bajo estudio. Más aún, en muchas situaciones prácticas no es posible observar continuamente el proceso y lo único que se tiene para hacer inferencia estadística sobre los parámetros es información parcial o información censurada del proceso completo. Así, un problema estadístico relevante es hacer inferencia sobre los parámetros involucrados en los procesos markovianos de salto, con base en escenarios observacionales tanto sin censura como con censura.

El objetivo de este trabajo de tesis es aplicar el enfoque de verosimilitud para hacer inferencia estadística sobre los parámetros involucrados en las tasas de intensidad de dos procesos markovianos de salto, el proceso de Poisson homogéneo y el proceso de nacimiento lineal, con base en escenarios observacionales tanto sin censura como con censura. Además, explorar algunas propiedades estadísticas de los intervalos de verosimilitud-confianza, como la frecuencia de cobertura y la longitud media, a través de estudios de simulación.

A continuación se describen los aspectos más importantes abordados en cada capítulo.

En el Capítulo 1 se define el proceso de Poisson, se discuten diferentes formas de definirlo y las demostraciones de algunas de sus propiedades, usando conceptos y resultados básicos de probabilidad.

En el Capítulo 2 se expone la construcción de los procesos markovianos de salto; es decir, procesos de salto que poseen la propiedad de Markov. Algunos de estos procesos de salto son: el proceso de nacimiento y muerte, el proceso de ramificación, el proceso de Poisson y el proceso de nacimiento.

En el Capítulo 3 se desarrollan conceptos relevantes del enfoque de verosimilitud utilizado aquí para estimar los parámetros involucrados en los procesos markovianos de salto bajo estudio. Algunos de estos conceptos son: la función de verosimilitud, la propiedad de invarianza de la verosimilitud, la función de verosimilitud relativa, las regiones de verosimilitud y los intervalos de verosimilitud-confianza.

En el Capítulo 4 se aplican los conceptos definidos en el Capítulo 3 para hacer inferencia estadística sobre los parámetros involucrados en las tasas de intensidad de dos procesos markovianos estudiados en esta tesis, el proceso de Poisson homogéneo y el proceso de nacimiento lineal, con base en escenarios observacionales tanto sin censura como con censura. Además, se presenta un estudio de simulación donde se muestra el comportamiento de las frecuencias de cobertura y la longitud media de los intervalos de verosimilitud-confianza de los parámetros que caracterizan las tasas de intensidad de los procesos bajo estudio. Por último, en el Capítulo 5 se presentan las conclusiones generales de esta tesis.

Las principales aportaciones de este trabajo de tesis son:

- Exposición adecuada de la construcción de los procesos markovianos de salto, utilizando conceptos y resultados básicos de probabilidad, para hacer inferencia estadística via el enfoque de verosimilitud en el caso del proceso de Poisson y del proceso de nacimiento lineal.
- Ilustración del uso del enfoque de verosimilitud para hacer inferencia estadística sobre los parámetros involucrados en las tasas de intensidad de dos procesos markovianos, el proceso de Poisson homogéneo y el proceso de nacimiento lineal, con base en escenarios observacionales tanto sin censura como con censura.
- Exploración de algunas propiedades estadísticas de los intervalos de verosimilitud-confianza, como la frecuencia de cobertura y la longitud media, a través de estudios de simulación que consideran escenarios observacionales tanto sin censura como con censura.

Posibles extensiones inmediatas de este trabajo de tesis son utilizar el enfoque de verosimilitud para hacer inferencia estadística sobre parámetros asociados a procesos markovianos más complejos. Por ejemplo, los procesos de nacimiento y muerte, donde se tienen dos parámetros (la tasa de nacimiento y la tasa de muerte); los procesos de Poisson no homogéneos, donde la intensidad del proceso es una función paramétrica; los procesos markovianos escondidos; entre otros.

Las simulaciones, cálculos de funciones de verosimilitud, optimizaciones y las figuras mostradas en esta tesis se realizaron con la ayuda del lenguaje de programación R versión 2.13.1 instalado en una computadora personal Intel Core Duo.

Capítulo 1

El proceso de Poisson

El *proceso de Poisson* es un proceso de conteo y existen varias formas equivalentes de presentarlo. El objetivo de este capítulo es discutir algunas de estas presentaciones y algunas de sus propiedades.

1.1. Procesos de conteo

Un *proceso de conteo* $N(t)$, $t \geq 0$, es un proceso estocástico tal que $N(t)$ cuenta o registra el número total de veces que cierto evento de interés ocurre en el intervalo de tiempo $(0, t]$. Para ilustrar esta idea, Considere los siguientes ejemplos.

Ejemplo 1. Sea $N(t)$, $t \geq 0$, un proceso de conteo para las siguientes situaciones:

- i) El número de personas que entran a una tienda en particular en el periodo $(0, t]$.
- ii) El número de nacimientos hasta el tiempo t inclusive.
- iii) El número de cuadrángulos que Willie Mays ha pegado al tiempo t inclusive. Si $t = 0$ corresponde al tiempo en el cual Mays inicia su carrera en grandes ligas, entonces $N(19) \approx 600$.

Observe que en los ejemplos anteriores los tiempos en que ocurren los eventos son aleatorios. Entonces, para precisar ideas, denote por T_n el tiempo aleatorio en el que ocurre el n -ésimo evento para cada $n \in \mathbb{N} := \{1, 2, \dots\}$ y defina $T_0 \equiv 0$. Además, como lo sugieren los ejemplos, resulta natural suponer que estos tiempos son no-decrecientes, es decir,

$$T_n \leq T_{n+1}, \quad \forall n \in \mathbb{N}_0 := \mathbb{N} \cup \{0\}.$$

Nótese que esto implica que la sucesión es no-negativa. Entonces, usando esta notación, se tiene que

$$N(t) = \max\{n \geq 0 : T_n \leq t\}, \quad \forall t \geq 0. \quad (1.1)$$

Puede verificarse directamente, usando la propiedad (1.1) por ejemplo, que un proceso de conteo $N(t)$, $t \geq 0$, tiene las siguientes propiedades.

Proposición 1. Sea $N(t)$, $t \geq 0$, un proceso de conteo y defina $N(s, t] := N(t) - N(s)$ para $0 \leq s \leq t$. Entonces:

- i) $N(t) \in \mathbb{N}_0$ para todo $t \geq 0$.
- ii) $N(s) \leq N(t)$ para todo $0 \leq s \leq t$.
- iii) $N(s, t]$ es el número de eventos que ocurren en el intervalo $(s, t]$, para todo $0 \leq s \leq t$.

1.2. Definición del proceso de Poisson

Uno de los tipos más importantes de proceso de conteo es el proceso de Poisson, que se define de la siguiente manera.

Definición 1. El proceso de conteo $N(t)$, $t \geq 0$, es un proceso de Poisson con intensidad $\lambda > 0$ si satisface las siguientes propiedades:

- i) $N(0) = 0$.
- ii) El proceso tiene incrementos independientes, es decir, las variables aleatorias

$$N(t_0, t_1], N(t_1, t_2], \dots, N(t_{n-1}, t_n]$$

son independientes para cada colección finita de tiempos $0 \leq t_0 < t_1 < \dots < t_{n-1} < t_n$.

- iii) Para cada par de tiempos $s, t \geq 0$, la variable aleatoria $N(s, s+t]$ tiene distribución de Poisson con parámetro λt , es decir,

$$P \{N(s, s+t] = n\} = \exp(-\lambda t) \frac{(\lambda t)^n}{n!}, \quad \forall n = 0, 1, \dots \quad (1.2)$$

Un proceso que satisface estas tres condiciones se le llama proceso de Poisson *homogéneo*, tal adjetivo, se refiere a que el parámetro de intensidad λ no cambia con el tiempo, es decir, es homogéneo en el tiempo.

Para determinar si un proceso de conteo arbitrario es un proceso de Poisson se debe mostrar que las condiciones en la Definición 1 se satisfacen. La condición *i*) es una convención y solamente establece que el conteo de eventos comienza en el tiempo $t = 0$, mientras que las condiciones *ii*) y *iii*) se refieren a la estructura probabilística del proceso. Generalmente, la condición *ii*) puede verificarse directamente desde nuestro conocimiento del proceso. Sin embargo, en muchos casos no es del todo claro cómo podría determinarse que la condición *iii*) se satisface, requiriéndose entonces formulaciones equivalentes del proceso de Poisson.

La hipótesis de incrementos independientes puede ser razonable para algunos procesos de conteo, como el del Ejemplo 1 *i*), pero podría no serlo para otros, como el del

Ejemplo 1 *ii*). Recuerde que en este ejemplo, $N(t)$ es el número de nacimientos en el intervalo $(0, t]$. Si $N(t)$ es muy grande, entonces es posible que haya mucha gente viva en el tiempo t y esto nos lleva a creer que el número de nuevos nacimientos entre el tiempo t y $t+s$ también tenderá a ser grande, por lo que no parece razonable que $N(t)$ sea independiente de $N(t+s) - N(t)$. La hipótesis de incrementos independientes en el Ejemplo 1 *iii*) será justificada si se considera que las oportunidades de que Mays pegue un cuadrángular hoy no depende de cómo le esté llando. No será justificado si se considera en las malas y buenas rachas de Mays.

Obsérvese que la Propiedad *iii*) en la Definición 1 implica que los incrementos de un proceso de Poisson son *estacionarios*, es decir,

$$P\{N(0, t] = n\} = P\{N(s, t+s] = n\}, \quad \forall s, t \geq 0, n \in \mathbb{N}_0.$$

En otras palabras, la distribución de un incremento solamente depende de la longitud del intervalo de tiempo y no de su posición. Nótese también que

$$EN(s, t+s] = \lambda t,$$

lo cual justifica que al parámetro $\lambda > 0$ se le llame la *intensidad* del proceso de Poisson.

La hipótesis de incrementos estacionarios sólo será razonable en el Ejemplo 1 *i*) si no hubiese horas pico (es decir, de las 12 a.m. a la 1 p.m.) en el que un número mayor de personas acudan a la tienda. En el caso contrario, la hipótesis de estacionariedad del proceso no será justificada. En el Ejemplo 1 *ii*), la hipótesis de estacionariedad estará justificada, por ejemplo, si se considera que la población de la tierra es esencialmente constante (una hipótesis poco razonable en tiempos actuales). De igual forma, la estacionariedad del proceso de conteo del Ejemplo 1 *iii*) no parece razonable; de hecho, la mayoría de las personas estarían de acuerdo en que Mays podría hacer más cuadrángulares entre los 25 – 30 años que entre los 35 – 40 años.

Para dar presentaciones alternativas del proceso de Poisson se requiere del concepto de funciones $o(h)$ que se introduce a continuación.

Definición 2. Se dirá que una función $f(\cdot)$ es $o(h)$ si

$$\lim_{h \rightarrow 0} \frac{f(h)}{h} = 0.$$

En otras palabras, para que la función $f(\cdot)$ sea $o(h)$ es necesario que $f(h)/h$ tienda a cero cuando h tiende a cero. Entonces, para tales funciones, si h tiende a cero, la única manera para que $f(h)/h$ tienda a cero, es que $f(h)$ se vaya más rápido a cero que h . Es decir, para h pequeña, $f(h)$ debe de ser pequeña en comparación a h . Los siguientes ejemplos sencillos ilustran este concepto.

Ejemplo 2. *i*) La función $f(x) = x^2$, $x \in \mathbb{R}$, es $o(h)$ ya que

$$\lim_{h \rightarrow 0} \frac{f(h)}{h} = \lim_{h \rightarrow 0} \frac{h^2}{h} = \lim_{h \rightarrow 0} h = 0.$$

ii) La función $f(x) = x$, $x \in \mathbb{R}$, no es $o(h)$ ya que

$$\lim_{h \rightarrow 0} \frac{f(h)}{h} = \lim_{h \rightarrow 0} \frac{h}{h} = \lim_{h \rightarrow 0} 1 = 1 \neq 0.$$

iii) Si $f(\cdot)$ y $g(\cdot)$ son funciones $o(h)$, entonces $f(\cdot) + g(\cdot)$ es $o(h)$, puesto que

$$\lim_{h \rightarrow 0} \frac{f(h) + g(h)}{h} = \lim_{h \rightarrow 0} \frac{f(h)}{h} + \lim_{h \rightarrow 0} \frac{g(h)}{h} = 0.$$

Entonces, cualquier combinación lineal finita de funciones $o(h)$ también es una función $o(h)$.

iv) Si $f(\cdot)$ es $o(h)$, entonces también lo es $cf(\cdot)$. Esto es inmediato ya que

$$\lim_{h \rightarrow 0} \frac{cf(h)}{h} = c \lim_{h \rightarrow 0} \frac{f(h)}{h} = 0.$$

El siguiente teorema proporciona una presentación alternativa del proceso de Poisson a la dada en la Definición 1.

Teorema 1. Un proceso de conteo $N(t)$, $t \geq 0$, es un proceso de Poisson con intensidad $\lambda > 0$ si y sólo si satisface las siguientes propiedades:

- i) $N(0) = 0$.
- ii) El proceso tiene incrementos estacionarios e independientes.
- iii) $P\{N(h) = 1\} = \lambda h + o(h)$.
- iv) $P\{N(h) \geq 2\} = o(h)$.

Esta caracterización del proceso de Poisson hace uso de las probabilidades “infinitesimales” del proceso, lo cual tiene algunas ventajas desde el punto de vista de la interpretación de lo que sucede en un intervalo “infinitesimal” de tiempo $(t, t + h]$. Dado que el proceso empieza en cero, por las propiedades del Teorema 1 iii) y iv), la probabilidad de que pase al estado uno al final de un intervalo de tiempo pequeño $[0, h]$ es $\lambda h + o(h)$, y la probabilidad de que el proceso no sufra ningún cambio en dicho intervalo es $1 - \lambda h + o(h)$, mientras que la probabilidad de que el proceso tenga dos o más incrementos en tal intervalo es $o(h)$. Es decir, en un intervalo cualquiera de longitud “infinitesimal” h esencialmente sólo pueden ocurrir dos situaciones: que haya un incremento o que haya ninguno.

Demostración. Primero se mostrará que las condiciones i)-iv) son necesarias para que el proceso de conteo sea de Poisson. Si $N(t)$, $t \geq 0$, es un proceso de Poisson, entonces claramente se cumplen las condiciones i) y ii) del teorema. Ahora se mostrará que también se cumplen iii) y iv). Para obtener la propiedad iii) obsérvese que de (1.2) se tiene que

$$\begin{aligned} P\{N(h) = 1\} &= (\lambda h) \exp(-\lambda h) \\ &= \lambda h + \lambda h [\exp(-\lambda h) - 1]. \end{aligned}$$

Luego, como

$$\lim_{h \rightarrow 0} \frac{\lambda h [\exp(-\lambda h) - 1]}{h} = \lambda \lim_{h \rightarrow 0} [\exp(-\lambda h) - 1] = 0,$$

se concluye que

$$P\{N(h) = 1\} = \lambda h + o(h).$$

Para obtener la propiedad *iii*) recuerde que

$$\lim_{h \rightarrow 0} \frac{1 - \exp(-\lambda h)}{h} = \lambda.$$

Luego, de (1.2) se tiene que

$$\begin{aligned} P\{N(h) \geq 2\} &= 1 - P\{N(h) = 0\} - P\{N(h) = 1\} \\ &= 1 - \exp(-\lambda h) - \lambda h \exp(-\lambda h). \end{aligned}$$

Entonces

$$\lim_{h \rightarrow 0} \frac{1}{h} P\{N(h) \geq 2\} = \lim_{h \rightarrow 0} \left[\frac{1 - \exp(-\lambda h)}{h} - \lambda \exp(-\lambda h) \right] = 0.$$

En resumen, un proceso de Poisson satisface las propiedades *i*)-*iv*) en el Teorema 1.

Ahora se probará que estas condiciones son suficientes para que $N(t)$, $t \geq 0$, sea un proceso de Poisson. Supóngase entonces que el proceso $N(t)$, $t \geq 0$, satisface las propiedades *i*)-*iv*) del Teorema 1. Obsérvese que sólo es necesario probar (1.2). Defínase

$$P_n(t) := P\{N(t) = n\}, \quad t \geq 0, n \in \mathbb{N}_0.$$

Primero se mostrará que

$$P'_0(t) = -\lambda P_0(t), \quad \forall t \geq 0.$$

Para probar lo anterior, note que por la hipótesis independencia y estacionariedad de los incrementos, se tiene que

$$\begin{aligned} P_0(t+h) &= P\{N(t+h) = 0\} \\ &= P\{N(t) = 0, N(t+h) - N(t) = 0\} \\ &= P\{N(t) = 0\} P\{N(t+h) - N(t) = 0\} \\ &= P_0(t) (1 - \lambda h + o(h)), \end{aligned}$$

donde la última ecuación se deriva de las propiedades *iii*) y *iv*). Entonces,

$$\frac{P_0(t+h) - P_0(t)}{h} = -\lambda P_0(t) + \frac{o(h)}{h}.$$

Ahora, si $h \rightarrow 0$ se obtiene

$$P'_0(t) = -\lambda P_0(t),$$

o equivalentemente

$$\frac{P'_0(t)}{P_0(t)} = -\lambda,$$

lo cual implica, por integración, que

$$\log P_0(t) = -\lambda t + c,$$

de donde se deduce que

$$P_0(t) = k \exp(-\lambda t),$$

donde k es una constante. Puesto que $P_0(0) = P\{N(0) = 0\} = 1$, se concluye que

$$P_0(t) = \exp(-\lambda t), \quad \forall t \geq 0. \quad (1.3)$$

Similarmente, para $n > 0$, se tiene que

$$\begin{aligned} P_n(t+h) &= P\{N(t+h) = n\} \\ &= P\{N(t) = n, N(t+h) - N(t) = 0\} \\ &\quad + P\{N(t) = n-1, N(t+h) - N(t) = 1\} \\ &\quad + \sum_{k=2}^n P\{N(t) = n-k, N(t+h) - N(t) = k\}. \end{aligned}$$

Ahora, de la estacionariedad del proceso resulta que

$$\begin{aligned} \sum_{k=2}^n P\{N(t) = n-k, N(t+h) - N(t) = k\} &\leq \sum_{k=2}^n P\{N(t+h) - N(t) = k\} \\ &\leq P\{N(h) \geq 2\} = o(h). \end{aligned}$$

Por otra parte, de la hipótesis de incrementos independientes, se tiene que

$$\begin{aligned} P_n(t+h) &= P_n(t)P_0(h) + P_{n-1}(t)P_1(h) \\ &= (1-\lambda h)P_n(t) + \lambda h P_{n-1}(t) + o(h), \end{aligned}$$

lo cual implica que

$$\frac{P_n(t+h) - P_n(t)}{h} = -\lambda P_n(t) + \lambda P_{n-1}(t) + \frac{o(h)}{h}.$$

Si $h \rightarrow 0$, se obtiene

$$P_n'(t) = -\lambda P_n(t) + \lambda P_{n-1}(t),$$

ó equivalentemente

$$\exp(\lambda t) (P_n'(t) + \lambda P_n(t)) = \lambda \exp(\lambda t) P_{n-1}(t),$$

lo cual a su vez puede re-escribirse como

$$\frac{d}{dt} [\exp(\lambda t) P_n(t)] = \lambda \exp(\lambda t) P_{n-1}(t). \quad (1.4)$$

Esta última ecuación y (1.3) implican que

$$\frac{d}{dt} [\exp(\lambda t) P_1(t)] = \lambda,$$

entonces

$$P_1(t) = (\lambda t + c) \exp(-\lambda t),$$

dado que $P_0(0) = 0$, se concluye que

$$P_1(t) = \lambda t \exp(-\lambda t).$$

Se usará inducción matemática para demostrar que

$$P_n(t) = \frac{(\lambda t)^n}{n!} \exp(-\lambda t), \quad \forall t \geq 0, n \in \mathbb{N}_0. \quad (1.5)$$

De hecho, note que (1.3) es la base de inducción. Ahora supóngase que (1.5) se cumple para $n - 1$. Entonces, de (1.4), se sigue que

$$\frac{d}{dt} [\exp(\lambda t) P_n(t)] = \frac{\lambda^n t^{n-1}}{(n-1)!},$$

lo cual implica que

$$\exp(\lambda t) P_n(t) = \frac{(\lambda t)^n}{n!} + c.$$

Puesto que $P_n(0) = 0$, de esta última ecuación se concluye que

$$P_n(t) = \frac{(\lambda t)^n}{n!} \exp(-\lambda t), \quad \forall t \geq 0, n \in \mathbb{N}_0.$$

Finalmente, note que de la estacionariedad de los incrementos se tiene que

$$P\{N(s, t+s) = n\} = P_n(t) = \exp(-\lambda t) \frac{(\lambda t)^n}{n!}, \quad \forall t \geq 0, n \in \mathbb{N}_0,$$

lo cual prueba el resultado deseado. ■

1.3. Distribuciones de tiempos de espera y entre arribos

Sea $N(t)$, $t \geq 0$, un proceso de Poisson y T_n , $n \in \mathbb{N}_0$, el tiempo de espera para la ocurrencia del n -ésimo evento. Nótese que el tiempo de espera entre la n -ésima y la $(n-1)$ -ésima ocurrencia del evento es

$$S_n := T_n - T_{n-1}, \quad n \in \mathbb{N}.$$

Puesto que $T_0 = 0$, resulta que $S_1 = T_1$. Claramente

$$T_n = S_1 + \cdots + S_n, \quad \forall n \in \mathbb{N}_0.$$

Debido a que usualmente se piensa en un proceso de Poisson como un modelo para los arribos de clientes a una estación de servicio (como en el Ejemplo 1 i)) a las variables S_n , $n \in \mathbb{N}$, se les llama *tiempos entre arribos*.

A continuación se encontrará la distribución de los tiempos entre arribos S_n y de los tiempos de espera T_n , $n \in \mathbb{N}_0$. Para hacer esto, primero note que el evento $\{S_1 > t\}$ ocurre si y sólo si no ocurren eventos en el intervalo $(0, t]$, entonces

$$P\{S_1 > t\} = P\{N(t) = 0\} = \exp(-\lambda t), \quad \forall t > 0.$$

Por lo tanto, S_1 tiene una distribución exponencial con media $1/\lambda$.

Ahora obsérvese que

$$\begin{aligned} P\{S_2 > t | S_1 = s\} &= P\{N(s, t] = 0 | S_1 = s\} \\ &= P\{N(s, t] = 0 | N(0, s] = 1\} \\ &= P\{N(s, t] = 0\} \\ &= \exp(-\lambda t). \end{aligned}$$

Entonces,

$$\begin{aligned} P\{S_2 > t\} &= E[P\{S_2 > t | S_1\}] \\ &= \int_0^\infty P\{S_2 > t | S_1 = s\} \lambda \exp(-\lambda s) ds \\ &= \exp(-\lambda t). \end{aligned}$$

Repitiendo el mismo argumento se obtiene el siguiente resultado.

Teorema 2. Si $N(t)$, $t \geq 0$, es un proceso de Poisson, entonces los tiempos entre arribos S_n , $n \in \mathbb{N}$, son variables aleatorias independientes e idénticamente distribuidas con distribución exponencial con media $1/\lambda$.

La propiedad anterior de los tiempos entre arribos caracteriza al proceso de Poisson. Este resultado se enuncia a continuación y su demostración puede consultarse en Billingsley (1979; pág. 263).

Teorema 3. Sea $N(t)$, $t \geq 0$, un proceso de conteo. Si los tiempos entre arribos S_n , $n \in \mathbb{N}$, son variables aleatorias independientes e idénticamente distribuidas con distribución exponencial con parámetro $\lambda > 0$, entonces, $N(t)$, $t \geq 0$, es un proceso de Poisson con intensidad $\lambda > 0$.

Este teorema se utilizará en el Capítulo 4 para simular datos de un proceso de Poisson.

Teorema 4. Si $N(t)$, $t \geq 0$, es un proceso de Poisson con intensidad $\lambda > 0$, entonces T_n tiene densidad gamma con parámetros n y λ .

Demostración. Sea F_{T_n} la distribución de T_n . Para encontrar la distribución F_{T_n} obsérvese que $N(t) \geq n$ si y sólo si $T_n \leq t$. Entonces,

$$\begin{aligned} F_{T_n}(t) &= P\{T_n \leq t\} = P\{N(t) \geq n\} \\ &= \sum_{j=n}^{\infty} \exp(-\lambda t) \frac{(\lambda t)^j}{j!} \\ &= 1 - \sum_{j=0}^{n-1} \exp(-\lambda t) \frac{(\lambda t)^j}{j!}. \end{aligned}$$

Denótese por $f_{T_n}(\cdot)$ la densidad de T_n . Al derivar ambos lados de igualdad anterior se obtiene

$$\begin{aligned} f_{T_n}(t) &= -\sum_{j=0}^{n-1} \lambda \exp(-\lambda t) \frac{(\lambda t)^j}{j!} + \sum_{j=n}^{\infty} \lambda \exp(-\lambda t) \frac{(\lambda t)^{j-1}}{(j-1)!} \\ &= \lambda \exp(-\lambda t) \frac{(\lambda t)^{n-1}}{(n-1)!}. \end{aligned}$$

Por lo tanto, T_n tiene densidad gamma con parámetros n y λ . ■

1.4. Otras propiedades del proceso de Poisson

Considérese un proceso de Poisson $N(t)$, $t \geq 0$, con intensidad λ . Supóngase que cada vez que ocurre un evento éste se clasifica como tipo I ó tipo II con probabilidad $p \in (0, 1)$ y probabilidad $q := 1 - p$, respectivamente. Supóngase además que la clasificación de cada evento es independiente de los restantes. Por ejemplo, los clientes que llegan a una tienda de acuerdo con un proceso de Poisson con intensidad λ se clasifican como “hombre” (tipo I) o “mujer” (tipo II).

Denote por $N_1(t)$ y $N_2(t)$ al número eventos del tipo I y tipo II que ocurren en el intervalo de tiempo $(0, t]$, respectivamente. Claramente, $N(t) = N_1(t) + N_2(t)$.

Teorema 5. *Los procesos $N_1(t)$ y $N_2(t)$, $t \geq 0$, son procesos de Poisson con intensidades λp y λq , respectivamente. Además, estos procesos son independientes.*

Demostración. Primero se encontrará la distribución conjunta

$$P\{N_1(t) = n, N_2(t) = m\}.$$

Para hacer esto, se condiciona con respecto a $N(t)$ para obtener

$$P\{N_1(t) = n, N_2(t) = m\} = \sum_{k=0}^{\infty} P\{N_1(t) = n, N_2(t) = m | N(t) = k\} P\{N(t) = k\}.$$

Claramente,

$$P\{N_1(t) = n, N_2(t) = m | N(t) = k\} = 0 \quad \text{si } k \neq n + m.$$

En el caso contrario, esto es, si $k = n + m$, se tiene que

$$\begin{aligned} P\{N_1(t) = n, N_2(t) = m | N(t) = k\} &= C_n^{n+m} p^n q^m \exp(-\lambda t) \frac{(\lambda t)^{n+m}}{(n+m)!} \\ &= \frac{(\lambda t p)^n (\lambda t q)^m}{n! m!} \exp(-\lambda t). \end{aligned} \quad (1.6)$$

Como consecuencia, la distribución marginal de $N_1(t)$ se obtiene de la siguiente manera:

$$\begin{aligned} P\{N_1(t) = n\} &= \sum_{m=0}^{\infty} P\{N_1(t) = n, N_2(t) = m\} \\ &= \exp(-\lambda tp) \frac{(\lambda tp)^n}{n!} \sum_{m=0}^{\infty} \exp(-\lambda tq) \frac{(\lambda tq)^m}{m!} \\ &= \exp(-\lambda tp) \frac{(\lambda tp)^n}{n!}. \end{aligned}$$

De forma similar se obtiene

$$P\{N_2(t) = m\} = \exp(-\lambda tq) \frac{(\lambda tq)^m}{m!}.$$

Además, combinando las dos igualdades anteriores con (1.6) se obtiene que

$$P\{N_1(t) = n, N_2(t) = m\} = P\{N_1(t) = n\} P\{N_2(t) = m\}.$$

Por lo tanto, los procesos $N_1(t)$ y $N_2(t)$, $t \geq 0$, son procesos de Poisson independientes con intensidades λp y λq , respectivamente. ■

Capítulo 2

Procesos markovianos de salto

Los procesos que a continuación se estudiarán evolucionan en tiempo continuo. Si el espacio de estados E es continuo—por ejemplo, un intervalo—los procesos en tiempo continuo pueden exhibir comportamientos muy complejos que pueden incluir desplazamientos continuos, saltos o combinaciones de ellos. Sin embargo, si el espacio de estados E es discreto—es decir, finito o numerable—los desplazamientos o cambios del proceso sólo pueden darse por saltos, razón por la cual se les conoce como *procesos de salto*.

En este capítulo se estudian una clase de procesos en tiempo continuo conocidos como *procesos markovianos de salto*, es decir, procesos de salto que tienen la propiedad de pérdida de memoria. A esta clase de procesos también se les conoce como *cadena de Markov en tiempo continuo*.

2.1. Construcción de los procesos de salto

Sea $X(t)$, $t \in T = [0, \infty)$, un proceso estocástico que modela un sistema cuyo espacio de estados es un conjunto discreto E . Suponga que el sistema inicia en el estado $x_0 \in E$, es decir, $X(0) = x_0$, donde permanece un tiempo aleatorio τ_1 . Si $\tau_1 = \infty$, entonces el sistema nunca abandona x_0 , es decir, $X(t) = x_0$ para todo $t \in T$; si τ_1 es finito, entonces el sistema se mueve o salta a un nuevo estado $x_1 \neq x_0$ en el instante τ_1 . Si esto último es el caso, el sistema permanece en x_1 un tiempo aleatorio τ_2 , y se repite el proceso anterior indefinidamente.

Definición 3. El *tiempo del primer salto* del proceso $X(t)$, $t \in T$, se define como

$$\tau_1 := \inf\{t > 0 : X(t) \neq X(0)\}$$

donde se usa la convención $\inf \emptyset = \infty$. Para $k \geq 2$, el *tiempo del k -ésimo salto* se define recursivamente como

$$\tau_k := \inf\{t > \tau_{k-1} : X(t) \neq X(\tau_{k-1})\}.$$

Notese que si $\tau_k = \infty$, entonces $\tau_m = \infty$ para todo $m > k$.

El proceso $X(t)$, $t \in T$, puede expresarse en términos de los tiempos de salto de la siguiente manera:

$$X(t) = x_{k-1} := X(\tau_{k-1}), \quad \tau_{k-1} \leq t < \tau_k, \quad (2.1)$$

con la convención $\tau_0 = 0$.

Definición 4. Sea un estado $x \in E$ y supóngase que $X(t) = x$ para algún $t \geq 0$. Se dirá que x es un **estado absorbente** si $X(\tau_n) = x$ y $\tau_{n+1} = \infty$ para algún $n \in \mathbb{N}_0$. En caso contrario, se dirá que x es un estado **no-absorbente**.

A primera vista podría parecer que (2.1) define $X(t)$ para toda $t \geq 0$. Pero este no es necesariamente el caso, como lo muestra el siguiente ejemplo. Supóngase que una pelota rebota en el suelo una y otra vez indefinidamente y que $X(t)$ es el número de rebotes de la pelota ocurridos hasta el tiempo t inclusive. Supóngase además que el tiempo entre el n -ésimo rebote y el $(n+1)$ -ésimo rebote es 2^{-n} . Entonces, para todo $n \in \mathbb{N}$,

$$\begin{aligned}\tau_n &= 1 + \frac{1}{2} + \cdots + \frac{1}{2^{n-1}} = 2 - \frac{1}{2^{n-1}} \\ x_n &= X(\tau_n) = n\end{aligned}$$

Nótese que $\tau_n < 2$, para todo $n \in \mathbb{N}$, y también que $\tau_n \rightarrow 2$ cuando $n \rightarrow \infty$. Entonces, (2.1) define $X(t)$ solo para $0 \leq t < 2$. Además, obsérvese que el número de rebotes hasta el tiempo $t = 2$ es infinito, lo que indica que sería apropiado definir $X(t) = \infty$ para $t \geq 2$.

Definición 5. El proceso $X(t)$, $t \geq 0$, es **no-explosivo** si

$$P_x \left\{ \lim_{n \rightarrow \infty} \tau_n = \infty \right\} = 1, \quad \forall x \in E.$$

En caso contrario, se dice que es **explosivo**, esto es, si

$$P_x \left\{ \lim_{n \rightarrow \infty} \tau_n < \infty \right\} > 0$$

para algún estado $x \in E$.

Nótese que si el proceso $X(t)$ es no-explosivo, entonces (2.1) define a $X(t)$ para toda $t \geq 0$. En lo que resta de trabajo se supondrá que el proceso $X(t)$, $t \geq 0$, es no-explosivo.

A continuación se especificará una estructura probabilística para los procesos de salto. Sea $Q = \{Q_{xy}\}_{x,y \in E}$ una matriz de transición sobre E tal que $Q_{xx} = 0$ para todo $x \in E$, y $F_x(\cdot)$ una función de distribución concentrada en los reales no-negativos para cada estado no-absorbente x . Supóngase que $X(0) = x$ donde x es un estado no-absorbente. Entonces, permanece en dicho estado un tiempo aleatorio no-negativo τ_1 cuya función de distribución está dada como

$$P_x \{ \tau_1 \leq t \} = F_x(t), \quad \forall t \geq 0.$$

En el tiempo τ_1 el sistema salta a un nuevo estado $X(\tau_1) = y \neq x$ de acuerdo a la distribución conjunta

$$P_x \{ \tau_1 \leq t, X(\tau_1) = y \} = F_x(t) Q_{xy}, \quad \forall x, y \in E, t \geq 0.$$

Si el estado y es absorbente, el sistema permanecerá por siempre en dicho estado y no experimentará nuevas transiciones. Si, por el contrario, el estado y es no-absorbente, el sistema permanecerá un tiempo aleatorio τ_2 y saltará a un nuevo estado $X(\tau_2) = z$ de acuerdo a la distribución

$$P_x \{ \tau_1 \leq t, X(\tau_1) = y, \tau_2 - \tau_1 \leq s, X(\tau_2) = z \} = F_x(t) Q_{xy} F_y(s) Q_{yz}.$$

El proceso anterior se repite indefinidamente.

Para cada $x, y \in E$, denótese por $P_{xy}(t)$ la probabilidad de el proceso que inicia en el estado x esté en el estado y en el tiempo $t > 0$, es decir,

$$P_{xy}(t) = P_x \{ X(t) = y \}.$$

Por conveniencia notacional defínase $P_{xy}(0) := \delta_{xy}$ donde δ_{xy} es la delta de Dirac, esto es $\delta_{xy} = 1$ si $x = y$, y $\delta_{xy} = 0$ si $x \neq y$. Puesto que el proceso es no-explosivo, se tiene que

$$\sum_y P_{xy}(t) = 1, \quad \forall x \in E.$$

A la matriz $P(t) = \{P_{xy}(t)\}$ se le llama matriz de transición en el tiempo t y a sus elementos se les llama funciones o probabilidades de transición.

A la distribución

$$\pi_0(x) := P \{ X(0) = x \}, \quad x \in E,$$

se le llama la distribución inicial del proceso. Entonces

$$P \{ X(t) = y \} = \sum_x \pi_0(x) P_{xy}(t), \quad \forall t \geq 0.$$

Definición 6. Un **proceso markoviano de saltos** $X(t), t \geq 0$, es un proceso de saltos que satisface la propiedad de Markov

$$P \{ X(t) = y | X(s_1) = x_1, \dots, X(s_{n-2}) = x_{n-2}, X(s) = x \} = P_{xy}(t - s)$$

para todo $0 \leq s_1 \leq \dots \leq s_{n-2} \leq s \leq t$ y $x_1, \dots, x_{n-2}, x, y \in E$.

Se puede demostrar que un proceso de salto es markoviano si y sólo si

$$P_x \{ \tau_1 > t + s | \tau_1 > s \} = P_x \{ \tau_1 > t \}, \quad \forall s, t \geq 0,$$

para todo estado no-absorbente $x \in E$. Note que la condición anterior es equivalente a la condición

$$\frac{1 - F_x(t + s)}{1 - F_x(s)} = 1 - F_x(t), \quad \forall s, t \geq 0,$$

la cual, a su vez, es equivalente a que la función de distribución F_x sea una distribución exponencial con parámetro $q_x > 0$ y su demostración se puede consultar en

Small (2007; pág. 38). Resumiendo, se tiene entonces que un proceso de salto es markoviano si y sólo si para cada estado no-absorbente $x \in E$ la función de distribución $F_x(t)$ tiene densidad

$$f_x(t) = \begin{cases} q_x \exp(-q_x t), & t \geq 0, \\ 0, & t < 0. \end{cases}$$

Obsérvese que

$$P_x \{\tau_1 \geq t\} = \int_t^\infty q_x \exp(-q_x s) ds = \exp(-q_x t), \quad \forall t \geq 0.$$

Nuevamente, por conveniencia notacional, si x es un estado absorbente defínase $q_x := 0$.

Sea $X(t)$, $t \geq 0$, un proceso markoviano de saltos no-explosivo. De la propiedad de Markov se obtiene directamente que

$$\begin{aligned} P\{X(t_0) = x_0, \dots, X(t_n) = x_n\} \\ = \pi_0(x_0) P_{x_0 x_1}(t_1 - t_0) \cdots P_{x_{n-1} x_n}(t_n - t_{n-1}), \end{aligned} \quad (2.2)$$

para todo $0 \leq t_0 \leq \dots \leq t_n$ y $x_0, \dots, x_n \in E$. En particular, para $s \geq 0$ y $t \geq 0$, se tiene

$$P_x \{X(t) = z, X(t+s) = y\} = P_{xz}(t) P_{zy}(s), \quad \forall x, y, z \in E.$$

Por otra parte, se tiene que

$$P_{xy}(t+s) = \sum_z P_x \{X(t) = z, X(t+s) = y\},$$

lo cual combinado con la igualdad anterior implica que

$$P_{xy}(t+s) = \sum_z P_{xz}(t) P_{zy}(s), \quad \forall s, t \geq 0.$$

A esta ecuación se le conoce como la *ecuación de Chapman-Kolmogorov*.

Teorema 6. Para todo $x, y \in E$, la función de transición $P_{xy}(t)$ satisface la ecuación integral

$$P_{xy}(t) = \delta_{xy} \exp(-q_x t) + \int_0^t q_x \exp(-q_x s) \sum_{z \neq y} Q_{xz} P_{zy}(t-s) ds, \quad \forall t \geq 0. \quad (2.3)$$

Demostración. Primero obsérvese que si x es un estado absorbente, (2.3) se reduce a

$$P_{xy}(t) = \delta_{xy}, \quad \forall t \geq 0,$$

la cual obviamente se cumple.

Supóngase ahora que x es un estado no-absorbente. Si $X(0) = x$, el evento

$$\{\tau_1 \leq t, X(\tau_1) = z, X(t) = y\},$$

ocurre si y sólo si el primer salto ocurre en algún tiempo $s \leq t$ y el proceso salta a z , para posteriormente moverse de z a y en las $t - s$ unidades de tiempo restantes. Entonces,

$$P_x \{ \tau_1 \leq t, X(\tau_1) = z, X(t) = y \} = \int_0^t q_x \exp(-q_x s) Q_{xz} P_{zy}(t-s) ds.$$

Por otra parte, note que

$$\begin{aligned} P_x \{ \tau_1 \leq t, X(t) = y \} &= \sum_{z \neq x} P_x \{ \tau_1 \leq t, X(\tau_1) = z, X(t) = y \} \\ &= \int_0^t q_x \exp(-q_x s) \left[\sum_{z \neq y} Q_{xz} P_{zy}(t-s) \right] ds. \end{aligned}$$

Obsérvese también que

$$\begin{aligned} P_x \{ \tau_1 > t, X(t) = y \} &= \delta_{xy} P_x \{ \tau_1 > t \} \\ &= \delta_{xy} \exp(-q_x t). \end{aligned}$$

En consecuencia, se tiene que

$$\begin{aligned} P_{xy}(t) &= P_x \{ X(t) = y \} \\ &= P_x \{ \tau_1 > t, X(t) = y \} + P_x \{ \tau_1 \leq t, X(t) = y \} \\ &= \delta_{xy} \exp(-q_x t) + \int_0^t q_x \exp(-q_x s) \sum_{z \neq y} Q_{xz} P_{zy}(t-s) ds, \end{aligned}$$

lo cual prueba el resultado deseado. ■

Teorema 7. Para cada par de estados $x, y \in E$, la función de transición $P_{xy}(t)$ es diferenciable y además satisface la **ecuación hacia atrás de Kolmogorov**

$$P'_{xy}(t) = -q_x P_{xy}(t) + q_x \sum_{z \neq x} Q_{xz} P_{zy}(t), \quad \forall t \geq 0. \quad (2.4)$$

Demostración. Haciendo el cambio de variable $u = t - s$ en la integral que aparece en (2.3), se obtiene

$$P_{xy}(t) = \delta_{xy} \exp(-q_x t) + q_x \exp(-q_x t) \int_0^t \exp(q_x u) \sum_{z \neq x} Q_{xz} P_{zy}(u) du, \quad \forall t \geq 0. \quad (2.5)$$

Se sigue de (2.5) que $P_{xy}(t)$ es continua para $t \geq 0$. Entonces, el integrando que aparece en esta ecuación es una función continua, lo cual implica que $P_{xy}(t)$ es diferenciable. Finalmente, con cálculos directos se obtiene (2.4). ■

Definición 7. A las cantidades

$$q_{xy} := P'_{xy}(0), \quad x, y \in E,$$

se les llama **parámetros infinitesimales** del proceso markoviano de salto.

Para obtener una forma más concreta de los parámetros infinitesimales note que de (2.4) se obtiene

$$\begin{aligned} P'_{xy}(0) &= -q_x P_{xy}(0) + q_x \sum_{z \neq x} Q_{xz} P_{zy}(0) \\ &= -q_x \delta_{xy} + q_x \sum_{z \neq x} Q_{xz} \delta_{zy} \\ &= -q_x \delta_{xy} + q_x Q_{xz}, \end{aligned}$$

lo cual implica que

$$q_{xy} = \begin{cases} -q_x & \text{si } y = x, \\ q_x Q_{xy} & \text{si } y \neq x. \end{cases} \quad (2.6)$$

Por otra parte, note que

$$\sum_{y \neq x} q_{xy} = \sum_{y \neq x} q_x Q_{xy} = -q_{xx} = q_x, \quad \forall x \in E, \quad (2.7)$$

y que la ecuación hacia atrás de Kolmogorov (2.4) se puede re-escribir como

$$P'_{xy}(t) = \sum_z q_{xz} P_{zy}(t), \quad \forall t \geq 0.$$

Si el espacio de estados E es finito, se puede diferenciar la ecuación de Chapman-Kolmogorov con respecto a s para obtener

$$P'_{xy}(t+s) = \sum_z P_{xz}(t) P'_{zy}(s), \quad \forall s, t \geq 0. \quad (2.8)$$

En particular,

$$P'_{xy}(t) = \sum_z P_{xz}(t) P'_{zy}(0), \quad \forall t \geq 0,$$

o equivalentemente

$$P'_{xy}(t) = \sum_z P_{xz}(t) q_{zy}, \quad \forall t \geq 0. \quad (2.9)$$

A la ecuación (2.9) se le conoce como la *ecuación hacia adelante de Kolmogorov*. Se puede demostrar que (2.8) y (2.9) también son válidas si E es infinito, véase Brémaud (1999; pág. 340).

En la siguiente sección se describen algunos ejemplos que muestran que las ecuaciones hacia atrás y hacia adelante pueden ser usadas para encontrar una fórmula explícita para las funciones de transición $P_{xy}(t)$.

2.2. Procesos de nacimiento y muerte

Un *proceso de nacimiento y muerte* es un proceso markoviano de salto con espacios de estados $E = \{0, 1, \dots, d\}$ o $E = \{0, 1, \dots\}$ cuyos parámetros infinitesimales satisfacen la condición

$$q_{xy} = 0 \quad \text{si } |y - x| > 1.$$

A las cantidades $\lambda_x := q_{x,x+1}$, $x \in E$, se les llama *intensidad de nacimiento*, y a las cantidades $\mu_x := q_{x,x-1}$ para $x \geq 1$ y $\mu_0 := 0$ se les llama *intensidad de muerte*. Si $E = \{0, 1, \dots, d\}$ se agrega la condición $\lambda_k = 0$ para todo $k \geq d$.

De la ecuación (2.7) se tiene

$$-q_{xx} = q_x = q_{x,x+1} + q_{x,x-1},$$

de modo que

$$q_{xx} = -(\lambda_x + \mu_x) \quad \text{o} \quad q_x = \lambda_x + \mu_x, \quad (2.10)$$

lo cual implica que

$$Q_{xy} = \begin{cases} \frac{\mu_x}{\lambda_x + \mu_x} & y = x - 1, \\ \frac{\lambda_x}{\lambda_x + \mu_x} & y = x + 1, \\ 0 & \text{en otro caso.} \end{cases}$$

Entonces, un proceso de nacimiento y muerte que se encuentra en el estado x solo puede saltar hacia el estado $x - 1$ o hacia $x + 1$. De esto se deduce que un estado x es absorbente si y sólo si $\lambda_x = \mu_x = 0$.

Un proceso de nacimiento y muerte es un *proceso de nacimiento* si $\mu_x = 0$ para todo $x \in E$. Si $\lambda_x = 0$ para todo $x \in E$ se dice que es un *proceso de muerte*. Un proceso de nacimiento puede moverse sólo a la derecha, mientras que un proceso de muerte puede moverse sólo a la izquierda.

En los ejemplos que se discutirán más adelante se supone que las tasas de nacimiento tienen un "crecimiento lineal", es decir, que existen constantes positivas A y B tales que

$$\lambda_x \leq A + Bx, \quad \forall x \in E.$$

Proposición 2. Sean ξ_i , para $i = 1, \dots, n$, variables aleatorias independientes con distribución exponencial con parámetro α_i , para $i = 1, \dots, n$. Entonces:

- i) $X = \min \{\xi_1, \dots, \xi_n\}$ tiene distribución exponencial con parámetro $\alpha = \alpha_1 + \dots + \alpha_n$.

$$ii) P\{X = \xi_k\} = \frac{\alpha_k}{\alpha_1 + \dots + \alpha_n} \text{ para cada } k = 1, \dots, n.$$

$$iii) P\{\xi_i = \xi_j\} = 0 \text{ para todo } i, j = 1, \dots, n.$$

Demostración. Para probar la parte *i*) observe que

$$\begin{aligned} P\{X > t\} &= P\{\xi_1 > t, \dots, \xi_n > t\} \\ &= P\{\xi_1 > t\}, \dots, P\{\xi_n > t\} \\ &= \exp(-\alpha_1 t) \cdots \exp(-\alpha_n t) \\ &= \exp[-(\alpha_1 + \dots + \alpha_n)t], \end{aligned}$$

lo cual prueba el resultado deseado.

Para probar *ii*) defínase

$$\eta_k = \min\{\xi_j : j \neq k\}.$$

De la parte *i*), se tiene que η_k tiene una distribución exponencial con parámetro

$$\beta_k = \sum_{j \neq k} \alpha_j.$$

Note además que ξ_k y η_k son independientes. Entonces,

$$\begin{aligned} P\{\xi_k = \min(\xi_1, \dots, \xi_n)\} &= P\{\xi_k \leq \eta_k\} \\ &= \int_0^\infty \left[\int_x^\infty \alpha_k \exp(-\alpha_k x) \beta_k \exp(-\beta_k y) dy \right] dx \\ &= \int_0^\infty \alpha_k \exp(-\alpha_k x) \exp(-\beta_k x) dx \\ &= \frac{\alpha_k}{\alpha_k + \beta_k} \\ &= \frac{\alpha_k}{\alpha_1 + \dots + \alpha_n}. \end{aligned}$$

La parte *iii*) es directa puesto que

$$P\{\xi_i = \xi_j\} = \int \int_{\{(x,y):x=y\}} f(x,y) dx dy = 0,$$

donde f es la densidad conjunta de ξ_i y ξ_j . ■

2.2.1. Procesos de ramificación

Cada partícula de cierta colección da lugar a generaciones sucesivas de partículas de forma independiente de las restantes. El mecanismo de “reproducción” o “ramificación” es el siguiente: cada partícula, desde el momento en que aparece, espera un tiempo aleatorio con distribución exponencial de parámetro q para dividirse en dos

partículas idénticas con probabilidad p o desaparecer con probabilidad $1 - p$.

Al proceso $X(t)$, $0 \leq t < \infty$, que cuenta el número de partículas presentes en el tiempo t se le llama proceso de ramificación.

Suponga que $X(0) = x$ y denote por ξ_1, \dots, ξ_x a los tiempos en que estas partículas se dividen o desaparecen. Las variables ξ_1, \dots, ξ_x son independientes y tiene distribución exponencial con parámetro q , y por lo tanto, $\tau_1 = \min\{\xi_1, \dots, \xi_x\}$ tiene distribución exponencial con parámetro $q_x = xq$. Puesto que cualquiera de las partículas tiene probabilidad p de dividirse en dos partículas y probabilidad $1 - p$ de desaparecer, para cada $x \geq 1$, se tiene que

$$Q_{x,x+1} = p \quad \text{y} \quad Q_{x,x-1} = 1 - p.$$

Claramente, el estado 0 es un estado absorbente. Por otra parte, puesto que $\lambda_x = q_x Q_{x,x+1}$ y $\mu_x = q_x Q_{x,x-1}$, se concluye que

$$\lambda_x = xqp \quad \text{y} \quad \mu_x = xq(1 - p), \quad x \geq 0.$$

Por la ecuación (2.10) y la definición de λ_x y μ_x , la ecuaciones hacia atrás y adelante para el proceso de nacimiento y muerte son

$$P'_{xy}(t) = \mu_x P_{x-1,y}(t) - (\lambda_x + \mu_x) P_{xy}(t) + \lambda_x P_{x+1,y}(t), \quad \forall t \geq 0, \quad (2.11)$$

y

$$P'_{xy}(t) = \lambda_{y-1} P_{x,y-1}(t) - (\lambda_y + \mu_y) P_{xy}(t) + \mu_{y+1} P_{x,y+1}(t), \quad \forall t \geq 0. \quad (2.12)$$

En la ecuación (2.12) se toma $\lambda_{-1} := 0$; si $E = \{0, \dots, d\}$ para $d < \infty$, se toma $\mu_{d+1} := 0$.

Se resolverán las ecuaciones hacia atrás y adelante en casos particulares. Para ello se utilizará el siguiente resultado. Si

$$f'(t) = -\alpha f(t) + g(t), \quad \forall t \geq 0, \quad (2.13)$$

entonces

$$f(t) = f(0) \exp(-\alpha t) + \int_0^t \exp[-\alpha(t-s)] g(s) ds, \quad \forall t \geq 0. \quad (2.14)$$

La demostración de este resultado es sencilla. Se multiplica (2.13) por $\exp(\alpha t)$ y se re-escribe la ecuación resultante como

$$\frac{d}{dt} [\exp(\alpha t) f(t)] = \exp(\alpha t) g(t).$$

Ahora, se integra de 0 a t y se encuentra que

$$\exp(\alpha t) f(t) - f(0) = \int_0^t \exp(\alpha s) g(s) ds,$$

y, por lo tanto, la (2.14) se cumple.

2.2.2. Proceso de nacimiento y muerte con dos estados

Considérese un proceso de nacimiento y muerte con espacio de estados $E = \{0, 1\}$ y supóngase que ambos estados son no-absorbentes. Como $\mu_0 = \lambda_1 = 0$, el proceso queda determinado por los parámetros λ_0 y μ_1 . Para simplificar la notación, tomemos $\lambda = \lambda_0$ y $\mu = \mu_1$. Este proceso de nacimiento y muerte se puede interpretar como el modelo de un sistema (teléfono, computadora, máquina) en el que estado 1 indica que el sistema está activo, mientras que el estado 0 indica que el sistema está inactivo o fuera de servicio. Entonces, si el sistema inicia en un estado de inactividad, permanecerá en esta condición un tiempo aleatorio con distribución exponencial de parámetro λ , y si inicia en estado de actividad permanecerá en servicio un tiempo aleatorio con distribución exponencial de parámetro μ .

A continuación se encontrará la función de transición de este proceso resolviendo la ecuación hacia atrás. Tomando $y = 0$ en (2.11), se obtiene

$$P'_{00}(t) = -\lambda P_{00}(t) + \lambda P_{10}(t), \quad \forall t \geq 0. \quad (2.15)$$

y

$$P'_{10}(t) = \mu P_{00}(t) - \mu P_{10}(t), \quad \forall t \geq 0. \quad (2.16)$$

Restando la segunda ecuación de la primera, resulta

$$\frac{d}{dt}(P_{00}(t) - P_{10}(t)) = -(\lambda + \mu)(P_{00}(t) - P_{10}(t)).$$

De la ecuación (2.14), se tiene

$$\begin{aligned} P_{00}(t) - P_{10}(t) &= (P_{00}(0) - P_{10}(0)) \exp[-(\lambda + \mu)t] \\ &= \exp[-(\lambda + \mu)t]. \end{aligned} \quad (2.17)$$

Aquí se usaron las formulas $P_{00}(0) = 1$ y $P_{10}(0) = 0$. Ahora, se sigue de (2.15) que

$$\begin{aligned} P'_{00}(t) &= -\lambda(P_{00}(t) - P_{10}(t)) \\ &= -\lambda \exp[-(\lambda + \mu)t]. \end{aligned}$$

Entonces

$$\begin{aligned} P_{00}(t) &= P_{00}(0) + \int_0^t P'_{00}(s) ds \\ &= 1 - \int_0^t \lambda \exp[-(\lambda + \mu)s] ds \\ &= 1 - \frac{\lambda}{\lambda + \mu} \{1 - \exp[-(\lambda + \mu)t]\}, \end{aligned}$$

o equivalentemente,

$$P_{00}(t) = \frac{\mu}{\lambda + \mu} + \frac{\lambda}{\lambda + \mu} \exp[-(\lambda + \mu)t], \quad t \geq 0. \quad (2.18)$$

Ahora, por (2.17), $P_{10}(t) = P_{00}(t) - \exp[-(\lambda + \mu)t]$, y por lo tanto

$$P_{10}(t) = \frac{\mu}{\lambda + \mu} - \frac{\mu}{\lambda + \mu} \exp[-(\lambda + \mu)t], \quad t \geq 0. \quad (2.19)$$

Tomando $y = 1$ en la ecuación hacia atrás, o bien, restando $P_{00}(t)$ y $P_{10}(t)$ de la unidad, se concluye que

$$P_{01}(t) = \frac{\lambda}{\lambda + \mu} - \frac{\lambda}{\lambda + \mu} \exp[-(\lambda + \mu)t], \quad \forall t \geq 0. \quad (2.20)$$

y

$$P_{11}(t) = \frac{\lambda}{\lambda + \mu} + \frac{\mu}{\lambda + \mu} \exp[-(\lambda + \mu)t], \quad \forall t \geq 0. \quad (2.21)$$

De (2.18)-(2.21), obsérvese que

$$\lim_{t \rightarrow \infty} P_{xy}(t) = \pi(y),$$

donde

$$\pi(0) = \frac{\mu}{\lambda + \mu} \quad \text{y} \quad \pi(1) = \frac{\lambda}{\lambda + \mu}. \quad (2.22)$$

Si π_0 es la distribución inicial del proceso, entonces por (2.18) y (2.19)

$$\begin{aligned} P\{X(t) = 0\} &= \pi_0(0)P_{00}(t) + (1 - \pi_0(0))P_{10}(t) \\ &= \frac{\mu}{\lambda + \mu} + \left(\pi_0(0) - \frac{\mu}{\lambda + \mu}\right) \exp[-(\lambda + \mu)t], \quad t \geq 0. \end{aligned}$$

Similarmente,

$$P\{X(t) = 1\} = \frac{\lambda}{\lambda + \mu} + \left(\pi_0(1) - \frac{\lambda}{\lambda + \mu}\right) \exp[-(\lambda + \mu)t], \quad \forall t \geq 0.$$

Así que $P\{X(t) = 0\}$ y $P\{X(t) = 1\}$ son independientes de t si y sólo si π_0 es la distribución π dada por (2.22).

2.2.3. Proceso de Poisson

En esta subsección, se mostrará como el proceso de Poisson es un proceso de salto.

Considérese un proceso de nacimiento $X(t)$, $0 \leq t < \infty$, con espacio de estados E igual a los enteros no-negativos tal que

$$\lambda_x = \lambda, \quad \forall x \geq 0.$$

Un proceso de nacimiento puede moverse sólo a la derecha, de modo que

$$P_{xy}(t) = 0, \quad t \geq 0 \text{ si } y < x. \quad (2.23)$$

Además, $P_{xx}(t) = P_x(\tau_1 > t)$, de modo que

$$P_{xx}(t) = \exp(-\lambda t), \quad \forall t \geq 0. \quad (2.24)$$

La ecuación hacia adelante para $y \neq 0$ es

$$P'_{xy}(t) = \lambda P_{x,y-1}(t) - \lambda P_{xy}(t), \quad \forall t \geq 0.$$

De (2.14) se deduce que

$$P_{xy}(t) = \exp(-\lambda t)P_{xy}(0) + \lambda \int_0^t \exp[-\lambda(t-s)]P_{x,y-1}(s) ds, \quad t \geq 0.$$

Como $P_{xy}(0) = \delta_{xy}$, se concluye que para $y > x$

$$P_{xy}(t) = \lambda \int_0^t \exp[-\lambda(t-s)]P_{x,y-1}(s) ds, \quad t \geq 0. \quad (2.25)$$

Se sigue de (2.24) y (2.25) que

$$\begin{aligned} P_{x,x+1}(t) &= \lambda \int_0^t \exp[-\lambda(t-s)] \exp(-\lambda s) ds \\ &= \lambda \exp(-\lambda t) \int_0^t ds = \lambda t \exp(-\lambda t), \end{aligned}$$

y usando (2.25) una vez más que

$$\begin{aligned} P_{x,x+2}(t) &= \lambda \int_0^t \exp[-\lambda(t-s)] \lambda s \exp(-\lambda s) ds \\ &= \lambda^2 \exp(-\lambda t) \int_0^t s ds = \frac{(\lambda t)^2}{2} \exp(-\lambda t). \end{aligned}$$

Por inducción se obtiene

$$P_{xy}(t) = \frac{(\lambda t)^{y-x} \exp(-\lambda t)}{(y-x)!}, \quad 0 \leq x \leq y \text{ y } t \geq 0. \quad (2.26)$$

Las fórmulas (2.23) y (2.24) implican que

$$P_{xy}(t) = P_{0,y-x}(t), \quad t \geq 0, \quad (2.27)$$

y que si $X(0) = x$, entonces $X(t) - x$ tiene una distribución de Poisson con parámetro λt .

En general, para $0 \leq s \leq t$, $X(t) - X(s)$ tiene distribución de Poisson con parámetro $\lambda(t-s)$. Porque si $0 \leq s \leq t$ y y es un entero no negativo, entonces

$$\begin{aligned} P\{X(t) - X(s) = y\} &= \sum_x P\{X(s) = x, X(t) = x + y\} \\ &= \sum_x P\{X(s) = x\} P_{x,x+y}(t-s) \\ &= \sum_x P\{X(s) = x\} P_{0y}(t-s) \\ &= P_{0y}(t-s) \\ &= \frac{[\lambda(t-s)]^y \exp[-\lambda(t-s)]}{y!}. \end{aligned}$$

Para $0 \leq t_1 \leq \dots \leq t_n$, las variables aleatorias $X(t_2) - X(t_1), \dots, X(t_n) - X(t_{n-1})$ son independientes. Ya que si z_1, \dots, z_{n-1} son enteros arbitrarios entonces por (2.2) y (2.27)

$$\begin{aligned} P\{X(t_2) - X(t_1) = z_1, \dots, X(t_n) - X(t_{n-1}) = z_{n-1}\} \\ = \sum_x P\{X(t_1) = x\} P_{0z_1}(t_2 - t_1) \cdots P_{0z_{n-1}}(t_n - t_{n-1}) \\ = P_{0z_1}(t_2 - t_1) \cdots P_{0z_{n-1}}(t_n - t_{n-1}) \\ = P\{X(t_2) - X(t_1) = z_1\} \cdots P\{X(t_n) - X(t_{n-1}) = z_{n-1}\}. \end{aligned}$$

Resumiendo, un proceso de nacimiento con intensidad de nacimiento constante, $\lambda_x = \lambda > 0$, $x \in E$, es un proceso de Poisson.

2.2.4. Proceso de nacimiento

Considérese un proceso de nacimiento $X(t)$, $0 \leq t < \infty$, en $\{1, 2, 3, \dots\}$. La ecuación hacia adelante de Kolmogorov (2.12) se reduce a

$$P'_{xy}(t) = \lambda_{y-1} P_{x,y-1}(t) - \lambda_y P_{xy}(t), \quad t \geq 0. \quad (2.28)$$

Ya que el proceso se mueve sólo a la derecha,

$$P'_{xy}(t) = 0, \quad y < x \text{ y } t \geq 0. \quad (2.29)$$

Se sigue de (2.28) y (2.29) que

$$P'_{xx}(t) = -\lambda_x P_{xx}(t).$$

Ya que $P_{xx}(0) = 1$ y $P_{xy}(0) = 0$ para $y > x$, se concluye de (2.14) que

$$P_{xx}(t) = \exp(-\lambda_x t), \quad t \geq 0. \quad (2.30)$$

y

$$P_{xy}(t) = \lambda_{y-1} \int_0^t \exp[-\lambda_y(t-s)] P_{x,y-1}(s) ds, \quad y > x \text{ y } t \geq 0. \quad (2.31)$$

Se puede usar (2.30) y (2.31) para encontrar $P_{xy}(t)$ recursivamente para $y > x$. En particular,

$$P_{x,x+1}(t) = \lambda_x \int_0^t \exp[-\lambda_{x+1}(t-s)] \exp(-\lambda_x s) ds,$$

y por lo tanto para $t \geq 0$

$$P_{x,x+1}(t) = \begin{cases} \frac{\lambda_x}{\lambda_{x+1} - \lambda_x} [\exp(-\lambda_x t) - \exp(-\lambda_{x+1} t)], & \lambda_{x+1} \neq \lambda_x, \quad \lambda_x t \exp(-\lambda_x t), \\ \lambda_x t \exp(-\lambda_x t), & \lambda_{x+1} = \lambda_x. \end{cases} \quad (2.32)$$

Ejemplo 3. Proceso de nacimiento lineal. *Considérese un proceso de nacimiento en $\{1, 2, 3, \dots\}$ con intensidad de nacimiento*

$$\lambda_x = x\lambda, \quad x \geq 0,$$

para alguna constante λ positiva (el proceso de ramificación $p = 1$ es de esta forma). A continuación se encontrará $P_{xy}(t)$.

Como se señaló anteriormente, $P_{xy}(t) = 0$ para $y < x$ y

$$P_{xx}(t) = \exp(-\lambda_x t) = \exp(-x\lambda t).$$

Nótese que de (2.32) que

$$P_{x,x+1}(t) = x \exp(-x\lambda t) [1 - \exp(-\lambda t)].$$

Para calcular $P_{x,x+2}(t)$ tómese $y = x + 2$ en (2.31) para obtener

$$\begin{aligned} P_{x,x+2}(t) &= (x+1)x\lambda \int_0^t \exp[-(x+2)\lambda(t-s)] \exp(-x\lambda s) [1 - \exp(-\lambda s)] ds \\ &= (x+1)x\lambda \exp[-(x+2)\lambda t] \int_0^t \exp(2\lambda s) [1 - \exp(-\lambda s)] ds \\ &= (x+1)x\lambda \exp[-(x+2)\lambda t] \int_0^t \exp(\lambda s) [\exp(-\lambda s) - 1] ds \\ &= (x+1)x\lambda \exp[-(x+2)\lambda t] \frac{[\exp(-\lambda t) - 1]^2}{2\lambda} \\ &= \binom{x+1}{2} \exp(-x\lambda t) [1 - \exp(-\lambda t)]^2. \end{aligned}$$

Se puede mostrar por inducción que

$$P_{xy}(t) = \binom{y-1}{y-x} \exp(-x\lambda t) (1 - \exp(-\lambda t))^{y-x}, \quad y \geq x, y, t \geq 0. \quad (2.33)$$

Capítulo 3

Inferencia estadística

En este capítulo se presentan conceptos básicos del enfoque de verosimilitud como: la función de verosimilitud, la función de verosimilitud relativa, regiones de verosimilitud y regiones de verosimilitud-confianza. Todos estos conceptos teóricos serán aplicados más adelante cuando sea tratado uno de los problemas centrales de este trabajo, hacer inferencia en los procesos markovianos de salto.

3.1. La función de verosimilitud

Supóngase que se ha formulado un modelo de probabilidad para un fenómeno aleatorio de interés y que dicho modelo involucra un vector de parámetros desconocidos $\theta = (\theta_1, \theta_2, \dots, \theta_m) \in \Theta \subset \mathbb{R}^m$ donde Θ es la región de valores posibles que puede tomar θ . Además, supóngase que el resultado de observar el fenómeno aleatorio de interés es un conjunto de datos que puede considerarse como un evento observado E_{obs} en el espacio muestral del modelo de probabilidad. La probabilidad del evento observado E_{obs} se puede determinar del modelo, y en general, será una función del vector de parámetros desconocidos θ , $P(E_{obs}; \theta)$. La función de verosimilitud de θ se define como la probabilidad del evento observado E_{obs} pero vista como función de θ ,

$$L(\theta; E_{obs}) = P(E_{obs}; \theta). \quad (3.1)$$

La verosimilitud en (3.1) se definió para el caso general, ya que un evento observado E_{obs} se puede asociar a un vector de variables aleatorias continuas o discretas $X = (X_1, X_2, \dots, X_n)$, que pueden ser independientes e idénticamente distribuidas o dependientes con alguna función de probabilidad conjunta paramétrica. Para los fines de esta tesis se considerará que $X = (X_1, X_2, \dots, X_n)$ es un vector de variables aleatorias discretas.

Definición 8. *Supóngase que $X = (X_1, X_2, \dots, X_n)$ es un vector de variables aleatorias discretas con función de probabilidad conjunta $P(x; \theta) : \mathbb{R}^n \rightarrow [0, 1]$ y $x = (x_1, x_2, \dots, x_n)$ es un valor observado de X , entonces la función de verosimilitud $L(\theta; x) : \Theta \rightarrow [0, 1]$ se define como*

$$L(\theta; x) = P(x; \theta) = P(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n; \theta).$$

Cuando $X = (X_1, X_2, \dots, X_n)$ es un vector de variables aleatorias discretas independientes e idénticamente distribuidas con función de probabilidad $P(x'; \theta)$ y $x = (x_1, x_2, \dots, x_n)$ es una muestra observada de X , entonces

$$\begin{aligned} P(X = x; \theta) &= P(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n; \theta) \\ &= \prod_{i=1}^n P(X_i = x_i; \theta) \\ &= \prod_{i=1}^n P(x_i; \theta). \end{aligned}$$

Por lo tanto, en este caso la función de verosimilitud del parametro θ es

$$L(\theta; x) = \prod_{i=1}^n P(x_i; \theta). \quad (3.2)$$

Nótese que por medio de la verosimilitud se puede combinar datos de experimentos distintos de una manera muy sencilla. La probabilidad conjunta de eventos independientes es el producto de sus probabilidades individuales. Así, la verosimilitud combinará conjuntos de datos que provienen de eventos distintos e independientes por medio del producto de las verosimilitudes individuales basadas en cada uno de estos conjuntos de datos.

Es importante mencionar que la verosimilitud no es una función de probabilidad. La probabilidad es una función de conjuntos que tiene como dominio una σ -álgebra y como contradominio el intervalo $[0, 1]$. En cambio, la verosimilitud es una función puntual cuyo dominio es el espacio parametral y su contradominio es el intervalo $[0, 1]$.

La función de verosimilitud juega un papel fundamental en la inferencia estadística. Su rol principal es inferir qué valores del vector de parámetros $\theta = (\theta_1, \theta_2, \dots, \theta_m)$ son razonables para la función de probabilidad $P(x; \theta)$ que haya sido elegida para el fenómeno aleatorio de interés a partir de un evento observado (datos). Esto es particularmente relevante después de un experimento, cuando ya fueron observados los datos.

A continuación se mostrará un ejemplo de una función de verosimilitud para el caso de una distribución Poisson con parámetro λ .

Ejemplo 4. *Supóngase que $X = (X_1, X_2, \dots, X_n)$ es un vector de variables aleatorias independientes idénticamente distribuidas (v.a.i.i.d.) con función de probabilidad Poisson de parámetro desconocido λ ,*

$$P(k; \lambda) = \frac{\lambda^k}{k!} \exp(-\lambda),$$

con $\lambda \geq 0$. Sea $x = (x_1, x_2, \dots, x_n)$ una muestra observada de X . Entonces, la probabilidad de observar $X = x$ es

$$\begin{aligned} P(X = x; \lambda) &= \prod_{i=1}^n P(x_i; \lambda) \\ &= \prod_{i=1}^n \frac{\lambda^{x_i}}{x_i!} \exp(-\lambda) \\ &= C \lambda^t \exp(-n\lambda), \end{aligned}$$

donde $C = \prod_{i=1}^n 1/x_i!$ es una constante que no depende de λ y $t = \sum_{i=1}^n x_i$ es la suma de las observaciones. Por lo tanto, de la Definición 1 se sigue que la función de verosimilitud del parámetro λ es

$$L(\lambda; x) \propto \lambda^t \exp(-n\lambda). \quad (3.3)$$

3.1.1. Plausibilidad

La función de verosimilitud $L(\theta; x)$ nos permite ordenar la credibilidad o plausibilidad entre los valores de θ a la luz de la muestra observada. Si $L(\theta_1; x) > L(\theta_2; x)$ entonces de (3.1) se sigue que $P(x; \theta_1) > P(x; \theta_2)$. Es decir, el evento observado x es más probable cuando el parámetro θ toma el valor θ_1 que cuando toma el valor θ_2 . Así, el cociente de verosimilitudes

$$\frac{L(\theta_1; x)}{L(\theta_2; x)} = \frac{P(x; \theta_1)}{P(x; \theta_2)},$$

es una medida de plausibilidad de θ_1 relativa a θ_2 con base en la muestra observada x . El cociente $L(\theta_1; x)/L(\theta_2; x) = k$ significa que el valor θ_1 es k veces más plausible que el valor de θ_2 en el sentido de que θ_1 hace a la muestra observada k veces más probable de lo que hace θ_2 .

3.1.2. Estimador de máxima verosimilitud (emv)

La función de verosimilitud $L(\theta; x)$ es una medida de plausibilidad de diferentes valores posibles de θ . Así, un estimador puntual razonable de θ es aquel valor de θ que maximiza $L(\theta; x)$.

Definición 9. El estimador de máxima verosimilitud de θ es cualquier valor $\hat{\theta}$ en el espacio parametral, $\hat{\theta} \in \Theta$, que cumple que

$$L(\hat{\theta}; x) = \sup_{\theta \in \Theta} L(\theta; x).$$

Nótese que $L(\theta; x) = P(x; \theta) \leq 1$ debido a que $P(x; \theta)$ es una probabilidad como función de θ . Así, el supremo de $L(\theta; x)$ en $[0, 1]$ existe y es finito. Sin embargo, puede ocurrir que el emv no existe, o si existe, puede ser que no sea único. De aquí en adelante se supondrá que existe un único valor $\hat{\theta} \in \Theta$ que maximiza a la función de

verosimilitud $L(\theta; x)$. Es decir, se supondrá que el emv existe y es único.

El emv del parámetro θ es el valor más plausible de θ . Es decir, el emv $\hat{\theta}$ es el valor de θ que explica mejor a la muestra observada en el sentido de que maximiza su probabilidad bajo el modelo de probabilidad propuesto para el fenómeno aleatorio de interés.

Para encontrar el emv de θ es necesario maximizar la función de verosimilitud $L(\theta; x)$. Sin embargo, usualmente $L(\theta; x)$ es el producto de varios términos y resulta matemáticamente conveniente trabajar con el logaritmo natural de la función de verosimilitud.

Definición 10. La función log-verosimilitud se define como

$$l(\theta; x) = \ln [L(\theta; x)].$$

El valor $\hat{\theta}$ que maximiza $L(\theta; x)$ también maximiza $l(\theta; x)$ puesto que el logaritmo natural es una transformación monótona creciente.

A continuación se muestra el uso de la log-verosimilitud junto con un procedimiento típico de cálculo matemático para obtener el emv del parámetro λ del modelo Poisson, bajo el contexto descrito en el Ejemplo 4.

Ejemplo 5. La función log-verosimilitud correspondiente a la verosimilitud del parámetro λ dada en la ecuación (3.3), véase Ejemplo 4, es

$$l(\lambda; x) = t \ln(\lambda) - n\lambda + C', \quad (3.4)$$

donde $C' = \ln C = \sum_{i=1}^n \ln(1/x_i!)$ es una constante arbitraria que no depende del parámetro λ . Para obtener el emv de λ , primero se procederá a encontrar los puntos críticos. Sin pérdida de generalidad, se supondrá que la suma de las observaciones es mayor que cero ($t > 0$).

Nótese que la primera derivada de (3.4) es

$$l'(\lambda; x) = \frac{t}{\lambda} - n.$$

Así, la solución de $l'(\lambda; x) = 0$ para λ produce un único punto crítico

$$\lambda = \frac{t}{n} = \bar{x}, \quad (3.5)$$

donde \bar{x} es el promedio de las observaciones.

La segunda derivada de (3.4) es estrictamente negativa como función de λ ,

$$l''(\lambda; x) = -\frac{t}{\lambda^2} < 0.$$

Entonces, el punto crítico dado en (3.5) es un máximo global. Por lo tanto, el emv de λ es $\hat{\lambda} = \bar{x}$.

No siempre es posible obtener una expresión matemática cerrada para el emv de los parámetros del modelo elegido para el fenómeno aleatorio de interés. Muchas veces, la estructura matemática del modelo de probabilidad produce una verosimilitud difícil de manipular para obtener dicha expresión cerrada para el emv. En estos casos, el problema de calcular el emv se convierte en un problema de optimización numérica de la verosimilitud.

3.2. Invarianza

En algunas ocasiones se utilizan reparametrizaciones de un modelo de probabilidad por conveniencia matemática, conveniencia computacional, interés en un parámetro que es función de otros que aparecen en el modelo, etc. Para estas situaciones la invarianza funcional es una propiedad muy conveniente de la verosimilitud. En términos del cálculo del emv, significa que el emv de una transformación $g(\theta)$ es la función evaluada en el emv.

Teorema 8. Si $g(\theta) : \Theta \rightarrow \Delta$ es cualquier función y $\hat{\theta}$ es el emv de θ entonces el emv de $\delta = g(\theta)$ es $\hat{\delta} = g(\hat{\theta})$.

Casella (2002, pág. 320) presenta la demostración del teorema anterior.

En términos de plausibilidad, cualquier declaración cuantitativa acerca de θ implica la misma declaración cuantitativa acerca de cualquier función uno a uno de θ . Una función $g(\theta)$ se llama función uno a uno si $g(\theta_1) = g(\theta_2)$ implica que $\theta_1 = \theta_2$.

Un ejemplo simple de la propiedad de invarianza de la verosimilitud es el siguiente.

Ejemplo 6. Si $\theta > 0$ y $\delta = \ln(\theta)$, entonces la verosimilitud del nuevo parámetro δ es $L(\delta; x) = L(\theta = \exp(\delta); x)$. Como consecuencia se tiene que el emv de δ es $\hat{\delta} = \ln(\hat{\theta})$.

3.3. Función de verosimilitud relativa

La función de verosimilitud relativa es una normalización o estandarización de la función de verosimilitud que permite ordenar la plausibilidad de los valores de θ .

Definición 11. La función de verosimilitud relativa de θ se define como

$$R(\theta; x) = \frac{L(\theta; x)}{\sup_{\theta \in \Theta} L(\theta; x)}. \quad (3.6)$$

Nótese que el supremo de $L(\theta; x)$ en $[0, 1]$ existe y es finito, entonces la función de verosimilitud relativa está bien definida y se puede emplear para hacer declaraciones cuantitativas acerca del grado sobre el cual, valores del parámetro describen mejor el fenómeno aleatorio de interés, con base en la muestra observada. Nótese que cuando

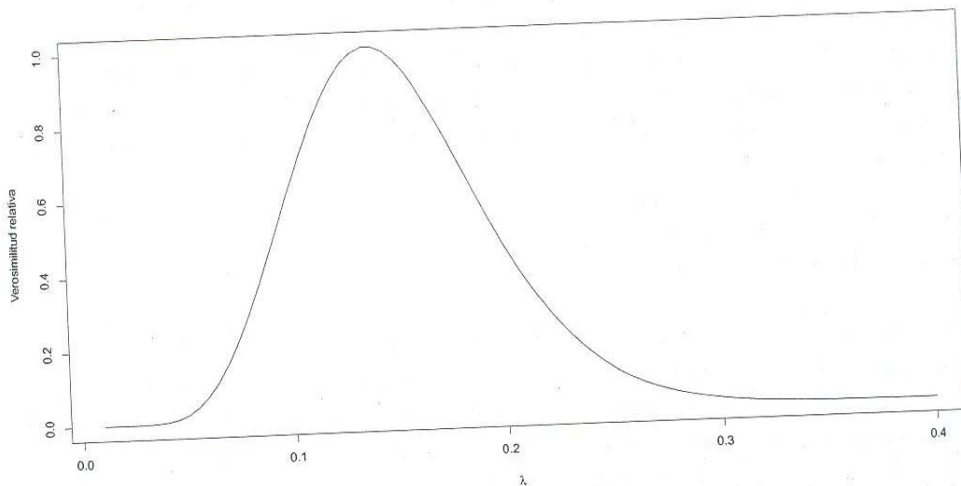


Figura 3.1: Función de verosimilitud relativa de la distribución de Poisson con parámetro $\lambda_0 = 0.1$ y $n = 10$.

el emv del parámetro θ existe y es único entonces la verosimilitud relativa de θ , $R(\theta; x)$ se puede expresar como

$$R(\theta; x) = \frac{L(\theta; x)}{L(\hat{\theta}; x)},$$

donde $\hat{\theta}$ es el emv del parámetro θ .

La función de verosimilitud relativa de θ permite medir la plausibilidad de cualquier valor específico de θ con respecto al emv $\hat{\theta}$ de θ . Valores de θ con $R(\theta; x)$ cercanos a 1 son muy plausibles mientras que valores cercanos a 0 son poco plausibles a la luz de los datos.

Ejemplo 7. La función de verosimilitud relativa del parámetro λ del modelo de Poisson, bajo el contexto descrito en el Ejemplo 4, es

$$\begin{aligned} R(\lambda; x) &= \frac{L(\lambda; x)}{L(\hat{\lambda}; x)} \\ &= \frac{\lambda^t \exp(-n\lambda)}{\hat{\lambda}^t \exp(-n\hat{\lambda})} \\ &= \left(\frac{\lambda}{\hat{\lambda}}\right)^t \exp[-n(\lambda - \hat{\lambda})], \end{aligned}$$

donde $t = \sum_{i=1}^n x_i$ y $\hat{\lambda} = \bar{x} = t/n$. La Figura 3.1 muestra la gráfica de $R(\lambda; x)$ donde $x = (x_1, x_2, \dots, x_n)$ es una muestra simulada de tamaño $n = 10$ de una distribución de Poisson con $\lambda_0 = 0.1$.

3.4. Regiones de verosimilitud

En la sección anterior, mediante la verosimilitud relativa, se obtuvo la plausibilidad de un cierto valor específico de θ con respecto al emv de θ , $\hat{\theta}$, dada la muestra observada. Sin embargo, en algunas ocasiones puede ser de interés especificar una región o intervalo para θ , donde cada valor de θ dentro de esta región o intervalo tenga al menos una plausibilidad especificada. Esta es una manera, diferente a la estimación puntual, de hacer inferencias sobre el parámetro θ usando el enfoque de verosimilitud.

Definición 12. Una región de verosimilitud de nivel c para θ , C , se define como

$$C(c) = \{\theta : R(\theta; x) \geq c, 0 \leq c \leq 1\}. \quad (3.7)$$

Cada valor de θ dentro de $C(c)$ tiene verosimilitud relativa mayor o igual a c y todo valor de θ fuera tiene verosimilitud relativa menor que c . Por lo tanto, $C(c)$ separa los valores plausibles de θ de los no plausibles a un nivel c .

Variando c de 0 a 1 se obtiene una familia jerarquizada y anidada de regiones, el cual es equivalente a la gráfica de $R(\theta; x)$. Nótese que el emv $\hat{\theta}$ está contenido en todas las regiones de verosimilitud. Además, cuando c tiende a uno la región $C(c)$ converge al emv de θ , $\hat{\theta}$.

En particular cuando θ es de dimensión uno, una región de verosimilitud para un determinado nivel de plausibilidad c puede ser un intervalo o la unión de varios intervalos disjuntos, esto dependerá de la forma que tome la función de verosimilitud. En particular, cuando una región de verosimilitud es un intervalo se le llama intervalo de verosimilitud.

3.4.1. Intervalos de verosimilitud

Definición 13. Un intervalo de verosimilitud de nivel c es una región de verosimilitud de nivel c de la forma $IV(c) = [A(c; x), B(c; x)]$, donde $A(c; x)$ y $B(c; x)$ son extremos del intervalo.

Generalmente los extremos de un intervalo de verosimilitud de nivel c , $A = A(c; x)$ y $B = B(c; x)$, cumplen que $R(A; x) = R(B; x) = c$. En estos casos, A y B pueden ser obtenidos trazando una línea horizontal en la gráfica de $R(\theta; x)$ a una distancia c paralela al eje cartesiano correspondiente a θ , y calculando los valores de θ asociados a la intersección de la línea con $R(\theta; x)$.

Un intervalo de verosimilitud por sí solo resulta poco informativo y en consecuencia insuficiente para mostrar el cambio en la plausibilidad de los valores de θ en el intervalo; por tal razón en muchos casos se recomienda utilizar distintos intervalos de verosimilitud de nivel c e indicar en ellos al emv $\hat{\theta}$ de θ . Esto permite identificar posibles asimetrías de la función de verosimilitud. En lo posible, se debe también

graficar y analizar la función de verosimilitud relativa completa.

A continuación se verá que es posible asociar una confianza deseada a los intervalos de verosimilitud, y que dicha confianza depende de la elección del nivel de plausibilidad c . En particular, niveles de plausibilidad de $c = 0.25$, 0.15 y 0.036 se encuentran asociados con niveles del 90 %, 95 % y 99 % de confianza (aproximadamente).

3.5. Regiones de verosimilitud-confianza

Sea $x = (x_1, x_2, \dots, x_n)$ una muestra observada de un vector de variables aleatorias $X = (X_1, X_2, \dots, X_n)$ con función de probabilidad conjunta $P(x; \theta)$, donde $\theta = (\theta_1, \theta_2, \dots, \theta_m) \in \Theta \subset \mathbb{R}^m$ es un vector de parámetros desconocidos. Si el valor verdadero de θ es θ_0 , entonces se puede calcular, a partir de la muestra observada x , una región C para ese valor verdadero θ_0 . Así, la región C cambia cuando varía la muestra. Debido a que la región C es función de las variables aleatorias, este puede algunas veces incluir y en otras no al valor verdadero θ_0 .

A la probabilidad de que una región C incluya o cubra al verdadero valor del parámetro θ_0 se le llama probabilidad de cobertura, $PC(\theta_0)$, y se define de la siguiente manera.

Definición 14. Sea $X = (X_1, X_2, \dots, X_n)$ un vector de variables aleatorias con función de probabilidad conjunta $P(x; \theta)$. La probabilidad de cobertura de una región $C = C(X)$ para $\theta = \theta_0$ se define como

$$PC(\theta_0) = P\{\theta_0 \in C; \theta = \theta_0\}.$$

Se puede interpretar a $PC(\theta_0)$ como la fracción de veces que la región C incluye o contiene al verdadero valor del parámetro θ , θ_0 , en un número muy grande de repeticiones de la muestra con θ fijo en θ_0 .

Una región C se llama región de confianza cuando su probabilidad de cobertura $PC(\theta_0)$ es la misma para todo valor del parámetro θ_0 . Una definición formal de región de confianza es la siguiente.

Definición 15. Sea $X = (X_1, X_2, \dots, X_n)$ un vector de variables aleatorias con función de probabilidad conjunta $P(x; \theta)$. Una región de confianza para θ_0 es una región $C = C(X)$ para $\theta = \theta_0$ cuya probabilidad de cobertura no depende de θ_0 .

Nótese que un valor específico de $\theta = \theta_0$ está en una región de verosimilitud $C(c)$ si y sólo si

$$R(\theta_0; x) \geq c,$$

o equivalentemente

$$-2 \ln R(\theta_0; x) \leq -2 \ln c.$$

Por lo tanto, la probabilidad de cobertura de $C(c)$ es

$$\begin{aligned} PC(\theta_0) &= P\{\theta_0 \in C(c); \theta_0\} \\ &= P\{D_n \leq -2 \ln c; \theta_0\}, \end{aligned}$$

donde $D_n = -2 \log R(\theta_0; x)$ es la estadística de la razón de verosimilitud para un θ fijo en θ_0 .

Existe teoría asintótica que da una buena aproximación a la distribución de D_n . Bajo ciertas condiciones de regularidad, la estadística de la razón de verosimilitud $D_n = -2 \ln R(\theta_0; x)$ converge en distribución a una ji-cuadrada con m grados de libertad para todo $\theta_0 \in \Theta \subset \mathbb{R}^m$. Es decir,

$$\lim_{n \rightarrow +\infty} P\{D_n \leq d\} = P\{\chi_m^2 \leq d\},$$

para todo d positivo. Para mayores detalles y la demostración, véase Serfling (1980; pág. 155-156).

Definición 16. Una región de verosimilitud-confianza para cualquier $\theta_0 \in \Theta \subset \mathbb{R}^m$ es una región de verosimilitud que tiene probabilidad de cobertura aproximada $(1 - \alpha)$ para θ_0 .

Entonces, si $d = -2 \ln c$ se tiene que $c = \exp(-d/2)$. Por lo tanto, cuando $d = q_{1-\alpha}$, donde $q_{1-\alpha}$ es el cuantil $(1 - \alpha)$ de una distribución ji-cuadrada con m grados de libertad se tiene la región de verosimilitud de nivel c , $C(c)$, tiene una probabilidad de cobertura del $100(1 - \alpha)\%$, donde $c = \exp(-q_{1-\alpha}/2)$. Así, $C(c)$ es una región de confianza para el parámetro θ y toma el nombre de región de verosimilitud-confianza para θ .

Los valores 2.706, 3.841 y 6.635 son los cuantiles 0.90, 0.95 y 0.99 de una distribución ji-cuadrada con un grado de libertad $m = 1$, respectivamente. Así, cuando θ es un parámetro escalar las regiones de verosimilitud (generalmente intervalos de verosimilitud) con $c = 0.25, 0.15$ y 0.036 tienen una probabilidad de cobertura aproximada del 90%, 95% y 99% respectivamente.

Ejemplo 8. En la Figura 3.2 se muestra la gráfica de la función de verosimilitud relativa del parámetro λ de la distribución de Poisson, dada en el Ejemplo 4, con base en una muestra de tamaño $n = 10$ y con $\lambda_0 = 0.1$. Además, en esta figura se marca la ubicación del valor verdadero del parámetro λ , $\lambda_0 = 0.1$, y tres intervalos de verosimilitud-confianza para λ de niveles de verosimilitud 0.25, 0.15 y 0.036. Nótese que estos intervalos tienen una confianza aproximada del 90%, 95% y 99%, respectivamente. Dichos intervalos son presentados en la Tabla 3.1 junto con el valor del emv del parámetro λ .

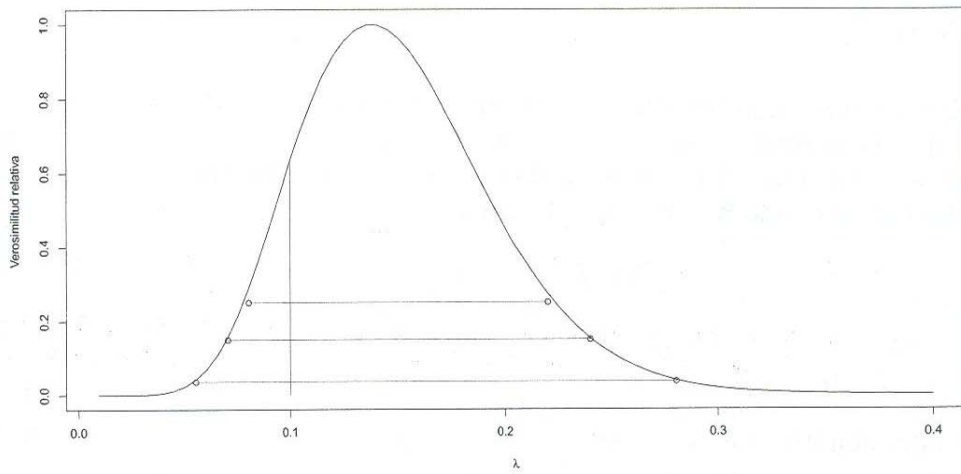


Figura 3.2: Función de verosimilitud relativa e intervalos de verosimilitud-confianza de la distribución de Poisson con parámetro λ (nivel de plausibilidad 0.25, 0.15 y 0.036) correspondientes a una muestra de tamaño $n = 10$ y con media $\lambda_0 = 0.1$.

Nivel de plausibilidad c	Nivel de confianza $100 * (1 - \alpha) \%$	Límite inferior del IV(c)	$\hat{\lambda}$	Límite superior del IV(c)
0.25	90 %	0.08	0.14	0.22
0.15	95 %	0.07	0.14	0.24
0.036	99 %	0.06	0.14	0.28

Tabla 3.1: Resumen de inferencias de verosimilitud para el parámetro λ de una distribución de Poisson.

Capítulo 4

Enfoque de verosimilitud para hacer inferencias en procesos markovianos de salto

En este capítulo se aplicará el enfoque de verosimilitud en dos procesos markovianos, el proceso de Poisson homogéneo y el proceso de nacimiento lineal, descritos en el Capítulo 1 y 2. En ambos casos se considerarán escenarios observacionales tanto sin censura como con censura. Además, se presentará un estudio de simulación para explorar el comportamiento de las frecuencias de cobertura de los intervalos de verosimilitud-confianza del parámetro asociado a los procesos bajo estudio; así como la longitud media de estos intervalos.

4.1. Inferencias en un proceso de Poisson homogéneo: Caso sin censura

Supóngase que se tiene un proceso de Poisson homogéneo $N(t)$, $t \geq 0$, con intensidad λ . Sean $t_1 < t_2 < \dots < t_n$ tiempos observados de ocurrencia de eventos (saltos) del proceso $N(t)$ a lo largo del tiempo. Es decir, se conoce el tiempo en el cual ocurrió cada salto. Sea $x_i = N(t_i) - N(t_{i-1})$, para $i = 1, 2, \dots, n$, donde $t_0 = 0$. Nótese que en este caso $x_i = 1$, para todo $i = 1, 2, \dots, n$. Entonces, la probabilidad conjunta de haber observado

$$E_{obs} = \{N(t_1) - N(t_0) = x_1, \dots, N(t_n) - N(t_{n-1}) = x_n\} \quad (4.1)$$

es

$$\begin{aligned} P\{E_{obs}; \lambda\} &= \prod_{i=1}^n \frac{[(t_i - t_{i-1}) \lambda]^{x_i}}{x_i!} \exp[-\lambda(t_i - t_{i-1})] \\ &= \left[\prod_{i=1}^n \frac{(t_i - t_{i-1})^{x_i}}{x_i!} \right] \lambda^{\sum_{i=1}^n x_i} \exp\left[-\lambda \sum_{i=1}^n (t_i - t_{i-1})\right] \\ &= C_1 \lambda^n \exp[-(\lambda T)], \end{aligned} \quad (4.2)$$

donde $C_1 = \prod_{i=1}^n (t_i - t_{i-1})^{x_i} / x_i!$ es una constante que no depende de λ , $t_i - t_{i-1}$ son los tiempos entre eventos consecutivos, para $i = 1, 2, \dots, n$, y $T = \sum_{i=1}^n (t_i - t_{i-1})$ la suma de estos tiempos. Nótese que en este caso $T = t_n$, el tiempo observado del último evento.

La función de verosimilitud del parámetro λ , correspondiente al proceso de Poisson $N(t)$ descrito anteriormente y al evento observado E_{obs} dado en (4.1), es

$$L(\lambda; \{x_i, t_i\}_{i=1}^n) \propto \lambda^n \exp[-(\lambda T)]. \quad (4.3)$$

La función de verosimilitud relativa de λ es

$$R(\lambda; \{x_i, t_i\}_{i=1}^n) = \frac{L(\lambda; \{x_i, t_i\}_{i=1}^n)}{L(\hat{\lambda}; \{x_i, t_i\}_{i=1}^n)} = \left(\frac{\lambda}{\hat{\lambda}}\right)^n \exp[-(\lambda - \hat{\lambda})T], \quad (4.4)$$

donde $\hat{\lambda} = n/T$ es el emv de λ . Para detalles sobre el cálculo del emv de λ , $\hat{\lambda}$, véase el Ejemplo 5 del Capítulo 3.

A continuación se ejemplifica el enfoque de verosimilitud, descrito en el Capítulo 3, para hacer inferencia sobre el parámetro λ con base en un juego de datos simulados.

La Figura 4.1 (a) muestra la gráfica de un proceso de Poisson homogéneo simulado con un parámetro de intensidad $\lambda_0 = 0.2$. Los $n = 20$ tiempos simulados de ocurrencia de cada evento o salto son mostrados (en orden creciente) en la Tabla 4.1. La Figura 4.1 (b) muestra la gráfica de la función de verosimilitud relativa del parámetro de intensidad λ , dada en (4.4), con base en los datos de la Tabla 4.1. Además, en esta figura se marca la ubicación del valor verdadero de λ , $\lambda_0 = 0.2$, y tres intervalos de verosimilitud-confianza para λ de niveles de verosimilitud 0.25, 0.15 y 0.036. Nótese que estos intervalos tienen una confianza aproximada del 90%, 95% y 99%, respectivamente. En la Figura 4.1 (b) se observa que la función de verosimilitud relativa de λ es ligeramente asimétrica (cola pesada a la derecha) con respecto al emv de λ , $\hat{\lambda} = 0.2093$. Además, se observa que λ_0 tiene una alta plausibilidad, $R(\lambda_0 = 0.2) = 0.9799$, y que los tres intervalos de verosimilitud-confianza cubren a λ_0 . Dichos intervalos son presentados en la Tabla 4.2 junto con el valor de emv de λ .

Salto	Tiempo	Salto	Tiempo	Salto	Tiempo	Salto	Tiempo
1	1.8829	6	19.1321	11	53.9793	16	75.8884
2	3.0772	7	23.2274	12	58.5594	17	78.0691
3	8.4854	8	28.7228	13	61.4574	18	83.2168
4	9.5987	9	42.6482	14	66.7989	19	84.5378
5	12.6897	10	51.6652	15	72.1686	20	95.5657

Tabla 4.1: Tiempos simulados (ordenados) de ocurrencia de eventos o saltos de un proceso Poisson de parámetro de intensidad $\lambda_0 = 0.2$.

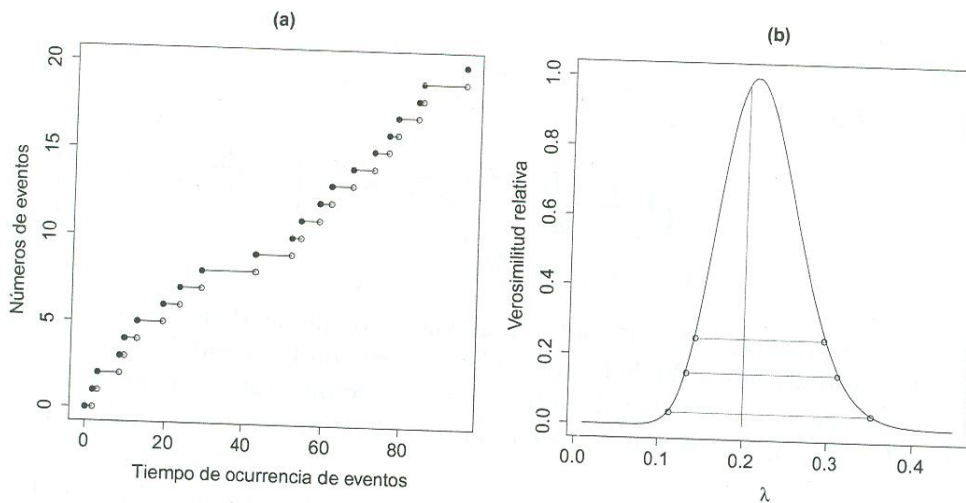


Figura 4.1: (a) Simulación de un proceso de Poisson con parámetro de intensidad $\lambda_0 = 0.2$. (b) Función de verosimilitud relativa e intervalos de verosimilitud-confianza para λ (nivel de plausibilidad 0.25, 0.15 y 0.036) correspondientes a los datos simulados mostrados en (a) y presentados en la Tabla 4.1.

Nivel de plausibilidad c	Nivel de confianza $100 * (1 - \alpha) \%$	Límite inferior del IV(c)	$\hat{\lambda}$	Límite superior del IV(c)
0.25	90 %	0.1413	0.2093	0.2962
0.15	95 %	0.1312	0.2093	0.3138
0.036	99 %	0.1112	0.2093	0.3538

Tabla 4.2: Resumen de inferencias de verosimilitud para el parámetro de intensidad λ de un proceso de Poisson homogéneo: Se usaron los datos de la Tabla 4.1, simulados con $\lambda_0 = 0.2$.

4.2. Inferencias en un proceso de Poisson homogéneo: Caso con censura

Supóngase que se tiene un proceso de Poisson homogéneo $N(t)$, $t \geq 0$, con intensidad λ . Sean $T_0 < T_1 < T_2 < \dots < T_k$ tiempos fijos definidos antes de observar el proceso $N(t)$. Sea $y_i = N(T_i) - N(T_{i-1})$, para $i = 1, 2, \dots, k$, donde $T_0 = 0$. Es decir, y_i es el número de saltos que han ocurrido entre los tiempos fijos T_{i-1} y T_i . Nótese que a diferencia del caso anterior donde se conoce que hubo un evento o salto en cada tiempo observado, aquí solo se conoce el número de eventos o saltos que ocurrieron en los intervalos de tiempo $[T_0, T_1], [T_1, T_2], \dots, [T_{k-1}, T_k]$, respectivamente. Es decir, y_i no necesariamente vale 1 para todo $i = 1, \dots, k$. Entonces, la probabilidad conjunta de haber observado

$$E_{obs} = \{N(T_1) - N(T_0) = y_1, \dots, N(T_k) - N(T_{k-1}) = y_k\} \quad (4.5)$$

es

$$\begin{aligned}
 P\{E_{obs}; \lambda\} &= \prod_{i=1}^k \frac{[(T_i - T_{i-1})\lambda]^{y_i}}{y_i!} \exp[-\lambda(T_i - T_{i-1})] \\
 &= \left[\prod_{i=1}^k \frac{(T_i - T_{i-1})^{y_i}}{y_i!} \right] \lambda^{\sum_{i=1}^k y_i} \exp\left[-\lambda \sum_{i=1}^k (T_i - T_{i-1})\right] \\
 &= C_2 \lambda^m \exp[-(\lambda S)],
 \end{aligned} \tag{4.6}$$

donde $C_2 = \prod_{i=1}^k (T_i - T_{i-1})^{y_i} / y_i!$ es una constante que no depende de λ , $T_i - T_{i-1}$ es la longitud del intervalo de tiempo donde ocurrieron los y_i saltos, para $i = 1, \dots, k$, $m = \sum_{i=1}^k y_i$ es la cantidad de saltos que ocurrieron hasta el tiempo T_k y $S = \sum_{i=1}^k (T_i - T_{i-1})$. Nótese que en este caso $S = T_k$, el último tiempo fijo.

La función de verosimilitud del parámetro λ , correspondiente al proceso de Poisson $N(t)$ descrito anteriormente y al evento observado E_{obs} dado en (4.5), es

$$L(\lambda; \{y_i, T_i\}_{i=1}^k) \propto \lambda^m \exp[-(\lambda S)]. \tag{4.7}$$

La función de verosimilitud relativa de λ es

$$R(\lambda; \{y_i, T_i\}_{i=1}^k) = \frac{L(\lambda; \{y_i, T_i\}_{i=1}^k)}{L(\hat{\lambda}; \{y_i, T_i\}_{i=1}^k)} = \left(\frac{\lambda}{\hat{\lambda}}\right)^m \exp[-(\lambda - \hat{\lambda})S], \tag{4.8}$$

donde $\hat{\lambda} = m/S$ es el emv de λ . Para encontrar $\hat{\lambda}$ se procede de manera similar al Ejemplo 5 del Capítulo 3. Nótese que en este caso la función de verosimilitud relativa de λ , dada en la ecuación (4.8), solo depende del tiempo fijo máximo $S = T_k$ y del número de saltos del proceso hasta T_k , m . Por lo tanto, las inferencias de verosimilitud sobre λ solo dependerán de estas cantidades.

A continuación se ejemplifica el enfoque de verosimilitud, descrito en el Capítulo 3, para hacer inferencia sobre el parámetro λ con base en una especificación particular de tiempos fijos de censura T_1, T_2, \dots, T_k para los datos simulados de la Tabla 4.1. Obsérvese que en este caso solo se requeriría especificar el tiempo máximo T_k ; sin embargo, se especificarán tiempos de censura T_1, T_2, \dots, T_{k-1} con el objetivo de proporcionar información del proceso en estos tiempos.

Sean $T_1 = 10, T_2 = 20, T_3 = 30, T_4 = 40$ los $k = 4$ tiempos fijos definidos antes de observar los tiempos simulados de salto del proceso Poisson homogéneo $N(t)$ con parámetro de intensidad $\lambda_0 = 0.2$, dados en la Tabla 4.1. El número de saltos y_1, y_2, y_3, y_4 que ocurrieron en los intervalos de tiempo $[T_0, T_1], (T_1, T_2], (T_2, T_3]$ y $(T_3, T_4]$, respectivamente, son mostrados en la Tabla 4.3. La Figura 4.2 muestra la gráfica de la función de verosimilitud relativa del parámetro de intensidad λ , dada en (4.8), con base en los datos de la Tabla 4.3. Además, en esta figura se marca la ubicación del valor verdadero de λ , $\lambda_0 = 0.2$, y tres intervalos de verosimilitud-confianza para λ de niveles de verosimilitud 0.25, 0.15 y 0.036 asociados a una

confianza aproximada del 90 %, 95 % y 99 %, respectivamente. En la Figura 4.2 se observa que la función de verosimilitud relativa de λ es asimétrica (cola pesada a la derecha) con respecto al emv de λ , $\hat{\lambda} = 0.2$, el cual fue incidentemente igual que $\lambda_0 = 0.2$. Nótese que la verosimilitud relativa de λ mostrada en la Figura 4.2, calculada bajo el escenario de censura descrito aquí, es mucho más asimétrica que la verosimilitud relativa de λ mostrada en la Figura 4.1 (b), calculada bajo el escenario sin censura. De hecho los intervalos de verosimilitud-confianza de λ son más anchos en el caso de censura que en el de sin censura (véase Tabla 4.2 y Tabla 4.4). Esto podría deberse principalmente a que en el caso de censura se tiene obviamente una cantidad menor de información que en el caso sin censura. Por otro lado, curiosamente la estimación puntual de λ fue mejor en el caso de censura que en el de sin censura, para las muestras observadas. Además como $\hat{\lambda} = \lambda_0 = 0.2$ entonces λ_0 tiene la más alta plausibilidad, $R(\lambda_0 = 0.2) = 1$, y por lo tanto los tres intervalos de verosimilitud-confianza, al igual que en el caso sin censura, cubren a λ_0 .

i	Intervalo de tiempo	Número de saltos (y_i)
1	0 – 10	4
2	10 – 20	2
3	20 – 30	2
4	30 – 40	0

Tabla 4.3: Datos censurados por intervalo correspondientes a los tiempos simulados de salto del proceso de Poisson homogéneo $N(t)$ con parámetro de intensidad $\lambda_0 = 0.2$, dados en la Tabla 4.1.

Nivel de plausibilidad c	Nivel de confianza $100*(1-\alpha) \%$	Límite inferior del IV(c)	$\hat{\lambda}$	Límite superior del IV(c)
0.25	90 %	0.1050	0.2	0.3412
0.15	95 %	0.0925	0.2	0.3700
0.036	99 %	0.0688	0.2	0.4412

Tabla 4.4: Resumen de inferencias de verosimilitud para el parámetro de intensidad λ de un proceso de Poisson homogéneo: Se usaron los datos de la Tabla 4.3, simulados con $\lambda_0 = 0.2$.

4.3. Inferencias en un proceso de nacimiento lineal: Caso sin censura

Supóngase que se tiene un proceso de nacimiento $X(t)$, $t \geq 0$, conjunto de estados $E = \{0, 1, 2, \dots\}$ e intensidad de nacimiento $\lambda_x = x\lambda$, $x \in E$. Sean $t_1 < t_2 < \dots < t_n$ tiempos observados de ocurrencia de eventos (saltos) del proceso $X(t)$ a lo largo del

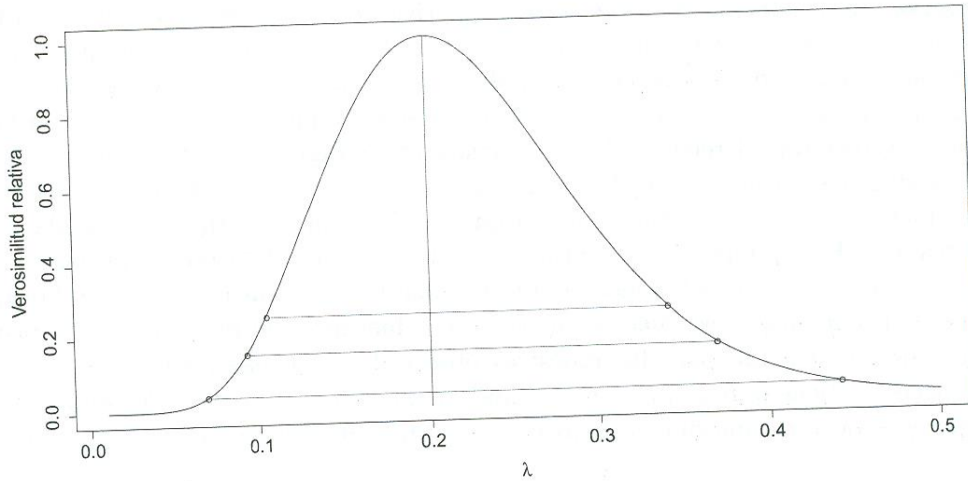


Figura 4.2: Función de verosimilitud relativa e intervalos de verosimilitud-confianza para λ (nivel de plausibilidad 0.25, 0.15 y 0.036) correspondientes a los datos simulados presentados en Tabla 4.3.

tiempo. Es decir, se conoce el tiempo en el cual ocurrió cada salto. Sea $x_i = X(t_i)$, para $i = 0, 1, \dots, n$, donde $t_0 = 0$ y $X(t_0) = x_0$, $x_0 \in E$. Nótese que en este caso $x_i = x_{i-1} + 1$, para todo $i = 1, 2, \dots, n$. Entonces, la probabilidad conjunta de haber observado

$$E_{obs} = \{X(t_0) = x_0, X(t_1) = x_1, \dots, X(t_n) = x_n\} \quad (4.9)$$

es

$$\begin{aligned} P\{E_{obs}; \lambda\} &= \prod_{i=1}^n x_{i-1} \exp[-x_{i-1}\lambda(t_i - t_{i-1})] \{1 - \exp[-\lambda(t_i - t_{i-1})]\}^{(x_i - x_{i-1})} \\ &= \left(\prod_{i=1}^n x_{i-1}\right) \exp\left[-\lambda \sum_{i=1}^n x_{i-1}(t_i - t_{i-1})\right] \prod_{i=1}^n \{1 - \exp[-\lambda(t_i - t_{i-1})]\}^{(x_i - x_{i-1})} \\ &= C_3 \exp\left[-\lambda \sum_{i=1}^n x_{i-1}(t_i - t_{i-1})\right] \prod_{i=1}^n \{1 - \exp[-\lambda(t_i - t_{i-1})]\}, \end{aligned} \quad (4.10)$$

donde $C_3 = \prod_{i=1}^n x_{i-1}$ es una constante que no depende de λ .

La función de verosimilitud del parámetro λ , correspondiente al proceso de nacimiento $X(t)$ descrito anteriormente y al evento observado E_{obs} dado en (4.9), es

$$\begin{aligned} L(\lambda; \{x_i, t_i\}_{i=1}^n) \\ \propto \exp\left[-\lambda \sum_{i=1}^n x_{i-1}(t_i - t_{i-1})\right] \prod_{i=1}^n \{1 - \exp[-\lambda(t_i - t_{i-1})]\}. \end{aligned} \quad (4.11)$$

La función de verosimilitud relativa de λ es

$$R(\lambda; \{x_i, t_i\}_{i=1}^k) = \frac{L(\lambda; \{x_i, t_i\}_{i=1}^k)}{L(\hat{\lambda}; \{x_i, t_i\}_{i=1}^k)} \\ = \exp \left[(\hat{\lambda} - \lambda) \sum_{i=1}^n x_{i-1}(t_i - t_{i-1}) \right] \prod_{i=1}^n \frac{1 - \exp[-\lambda(t_i - t_{i-1})]}{1 - \exp[-\hat{\lambda}(t_i - t_{i-1})]}, \quad (4.12)$$

donde $\hat{\lambda}$ es el emv de λ . Nótese que en este caso no es fácil obtener una expresión matemática cerrada para $\hat{\lambda}$. Esta es una situación donde resulta conveniente calcular $\hat{\lambda}$ maximizando numéricamente la función de verosimilitud dada en (4.11) o equivalentemente minimizando el negativo de la log-verosimilitud.

A continuación se ejemplifica el enfoque de verosimilitud, descrito en el Capítulo 3, para hacer inferencia sobre el parámetro λ con base en un juego de datos simulados.

La Figura 4.3 (a) muestra la gráfica de un proceso de nacimiento lineal $X(t)$ simulado con un parámetro de intensidad $\lambda_x = x\lambda_0$, donde $\lambda_0 = 0.2$ y estado inicial $X(0) = x_0 = 1$. Los $n = 20$ tiempos simulados de ocurrencia de cada evento o salto son mostrados (en orden creciente) en la Tabla 4.5. La Figura 4.3 (b) muestra la gráfica de la función de verosimilitud relativa del parámetro λ , con base en los datos de la Tabla 4.5. Además, en esta figura se marca la ubicación del valor verdadero de λ , $\lambda_0 = 0.2$, y tres intervalos de verosimilitud-confianza para λ de niveles de verosimilitud 0.25, 0.15 y 0.036. Nótese que estos intervalos tienen una confianza aproximada del 90 %, 95 % y 99 %, respectivamente. En la Figura 4.3 (b) se observa que la función de verosimilitud relativa de λ es ligeramente asimétrica (cola pesada a la derecha) con respecto al emv de λ , $\hat{\lambda} = 0.217$. Además, se observa que λ_0 tiene una alta plausibilidad, $R(\lambda_0 = 0.2) = 0.9372$, y que los tres intervalos de verosimilitud-confianza cubren a λ_0 . Dichos intervalos son presentados en la Tabla 4.6 junto con el valor de emv de λ .

Salto	Tiempo	Salto	Tiempo	Salto	Tiempo	Salto	Tiempo
1	0.2678	6	4.1752	11	7.8989	16	9.2316
2	1.4026	7	6.0789	12	8.1927	17	9.4833
3	2.1477	8	6.4083	13	8.7509	18	9.8878
4	3.0047	9	7.1059	14	9.0114	19	9.9938
5	3.6878	10	7.4431	15	9.1820	20	10.5204

Tabla 4.5: Tiempos simulados (ordenados) de ocurrencia de eventos o saltos de un proceso de nacimiento lineal $X(t)$ con parámetro de intensidad $\lambda_x = x\lambda_0$, donde $\lambda_0 = 0.2$, y estado inicial $X(0) = x_0 = 1$.

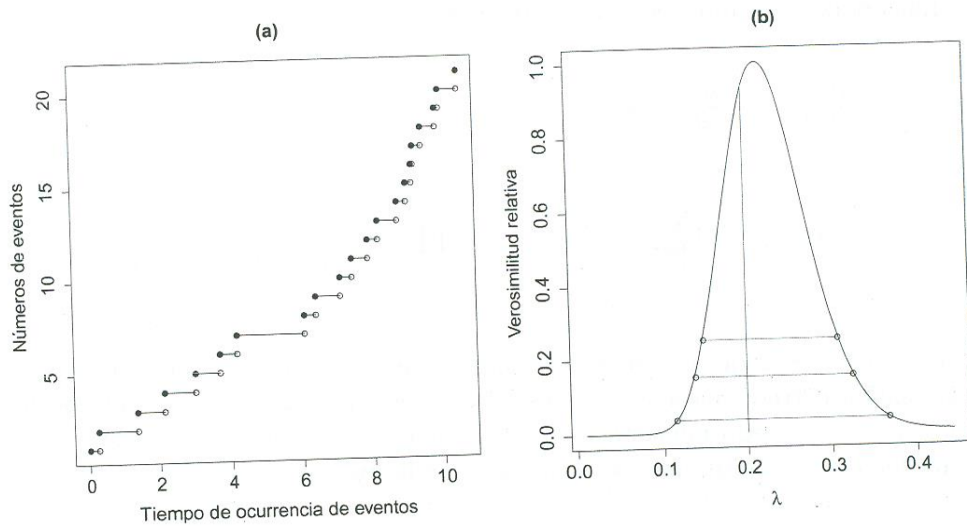


Figura 4.3: (a) Simulación de un proceso de nacimiento lineal con parámetro de intensidad $\lambda_x = x\lambda_0$, donde $\lambda_0 = 0.2$. (b) Función de verosimilitud relativa e intervalos de verosimilitud-confianza para λ (nivel de plausibilidad 0.25, 0.15 y 0.036) correspondientes a los datos simulados mostrados en (a) y presentados en la Tabla 4.5.

Nivel de plausibilidad c	Nivel de confianza $100*(1-\alpha)\%$	Límite inferior del IV(c)	$\hat{\lambda}$	Límite superior del IV(c)
0.25	90 %	0.1475	0.217	0.3075
0.15	95 %	0.1375	0.217	0.3250
0.036	99 %	0.1150	0.217	0.3675

Tabla 4.6: Resumen de inferencias de verosimilitud para el parámetro λ de un proceso de nacimiento lineal: Se usaron los datos de la Tabla 4.5, simulados con $\lambda_0 = 0.2$.

4.4. Inferencias en un proceso de nacimiento lineal: Caso con censura

Supóngase que se tiene un proceso de nacimiento $X(t)$, $t \geq 0$, conjunto de estados $E = \{0, 1, 2, \dots\}$ e intensidad de nacimiento $\lambda_x = x\lambda$, $x \in E$. Sean $T_0 < T_1 < T_2 < \dots < T_k$ tiempos fijos definidos antes de observar el proceso $X(t)$. Sea $y_i = X(T_i)$, para $i = 0, 1, \dots, k$, donde $T_0 = 0$ y $X(T_0) = y_0$, $y_0 \in E$. Es decir, y_i es el número de saltos que han ocurrido hasta el tiempo fijo T_i más el valor del estado inicial del proceso y_0 . Nótese que a diferencia del caso anterior donde se conoce que hubo un evento o salto en cada tiempo observado, aquí solo se conoce el número de eventos o saltos que ocurrieron hasta cada uno de los tiempos fijos $T_0, T_1, T_2, \dots, T_k$. Entonces,

la probabilidad conjunta de haber observado

$$E_{obs} = \{X(T_0) = y_0, X(T_1) = y_1, \dots, X(T_k) = y_k\} \quad (4.13)$$

es

$$\begin{aligned} P\{E_{obs}; \lambda\} &= \prod_{i=1}^k C_{y_i - y_{i-1}}^{y_i - 1} \exp[-y_{i-1} \lambda (T_i - T_{i-1})] \{1 - \exp[-\lambda (T_i - T_{i-1})]\}^{(y_i - y_{i-1})} \\ &= C_4 \exp\left[-\lambda \sum_{i=1}^k y_{i-1} (T_i - T_{i-1})\right] \prod_{i=1}^k \{1 - \exp[-\lambda (T_i - T_{i-1})]\}^{(y_i - y_{i-1})}, \end{aligned} \quad (4.14)$$

donde $C_4 = \prod_{i=1}^k C_{y_i - y_{i-1}}^{y_i - 1}$ es una constante que no depende de λ .

La función de verosimilitud del parámetro λ , correspondiente al proceso de nacimiento lineal $X(t)$ descrito anteriormente y al evento observado E_{obs} dado en (4.13), es

$$\begin{aligned} L(\lambda; \{y_i, T_i\}_{i=1}^k) &\propto \exp\left[-\lambda \sum_{i=1}^k y_{i-1} (T_i - T_{i-1})\right] \prod_{i=1}^k \{1 - \exp[-\lambda (T_i - T_{i-1})]\}^{(y_i - y_{i-1})}. \end{aligned} \quad (4.15)$$

La función de verosimilitud relativa de λ es

$$\begin{aligned} R(\lambda; \{y_i, T_i\}_{i=1}^k) &= \frac{L(\lambda; \{y_i, T_i\}_{i=1}^k)}{L(\hat{\lambda}; \{y_i, T_i\}_{i=1}^k)} \\ &= \exp\left[(\hat{\lambda} - \lambda) \sum_{i=1}^k y_{i-1} (T_i - T_{i-1})\right] \prod_{i=1}^k \left\{ \frac{1 - \exp[-\lambda (T_i - T_{i-1})]}{1 - \exp[-\hat{\lambda} (T_i - T_{i-1})]} \right\}^{(y_i - y_{i-1})}, \end{aligned} \quad (4.16)$$

donde $\hat{\lambda}$ es el emv de λ . Nótese que en este caso no es fácil obtener una expresión matemática cerrada para $\hat{\lambda}$, entonces se procede de manera similar al caso anterior (Sección 4.3) maximizando numéricamente la función de verosimilitud dada en (4.15) o equivalentemente minimizando el negativo de la log-verosimilitud.

A continuación se ejemplifica el enfoque de verosimilitud, descrito en el Capítulo 3, para hacer inferencia sobre el parámetro λ con base en una especificación particular de tiempos fijos de censura T_1, T_2, \dots, T_k para los datos simulados de la Tabla 4.5.

Sean $T_1 = 2, T_2 = 4, T_3 = 6, T_4 = 8$ los $k = 4$ tiempos fijos definidos antes de observar los tiempos simulados de salto del proceso de nacimiento lineal $X(t)$

con parámetro $\lambda_0 = 0.2$, dados en la Tabla 4.5. El número de saltos y_1, y_2, y_3, y_4 que ocurrieron hasta los tiempos fijos T_1, T_2, T_3, T_4 (más y_0) son mostrados en la Tabla 4.7. La Figura 4.4 muestra la gráfica de la función de verosimilitud relativa del parámetro λ , dada en (4.16), con base en los datos de la Tabla 4.7. Además, en esta figura se marca la ubicación del valor verdadero de λ , $\lambda_0 = 0.2$, y tres intervalos de verosimilitud-confianza para λ de niveles de verosimilitud 0.25, 0.15 y 0.036 asociados a una confianza aproximada del 90 %, 95 % y 99 %, respectivamente. En la Figura 4.4 se observa que la función de verosimilitud relativa de λ es asimétrica (cola pesada a la derecha) con respecto al emv de λ , $\hat{\lambda} = 0.3095$. Nótese que la verosimilitud relativa de λ mostrada en la Figura 4.4, calculada bajo el escenario de censura descrito aquí, es mucho más asimétrica que la verosimilitud relativa de λ mostrada en la Figura 4.3 (b), calculada bajo el escenario sin censura. De hecho los intervalos de verosimilitud-confianza de λ son más anchos en el caso de censura que en el de sin censura (véase Tabla 4.6 y Tabla 4.8). Además se observa que el emv de λ , $\hat{\lambda} = 0.3095$, se encuentra un poco alejado del valor verdadero de λ , $\lambda_0 = 0.2$, y que los tres intervalos de verosimilitud-confianza cubren a λ_0 , ya que λ_0 tiene una plausibilidad de $R(\lambda_0 = 0.2) = 0.3788$ la cual es menor que en el caso sin censura. Dichos intervalos son presentados en la Tabla 4.8 junto con el valor de emv de λ .

i	Tiempos fijos (T_i)	Datos (y_i)
0	0	1
1	2	3
2	4	6
3	6	7
4	8	12

Tabla 4.7: Datos censurados correspondientes a los tiempos simulados de salto del proceso de nacimiento lineal $X(t)$ con parámetro $\lambda_0 = 0.2$, dados en la Tabla 4.5.

Nivel de plausibilidad c	Nivel de confianza $100*(1-\alpha)$ %	Límite inferior del IV(c)	$\hat{\lambda}$	Límite superior del IV(c)
0.25	90 %	0.1825	0.3095	0.4862
0.15	95 %	0.1662	0.3095	0.5225
0.036	99 %	0.1312	0.3095	0.6088

Tabla 4.8: Resumen de inferencias de verosimilitud para el parámetro λ de un proceso de nacimiento lineal: Se usaron los datos de la Tabla 4.7, simulados con $\lambda_0 = 0.2$.

4.5. Estudio de simulación

En las secciones anteriores se aplicó el enfoque de verosimilitud para hacer inferencias sobre el parámetro λ , asociado al proceso de Poisson homogéneo y al proceso

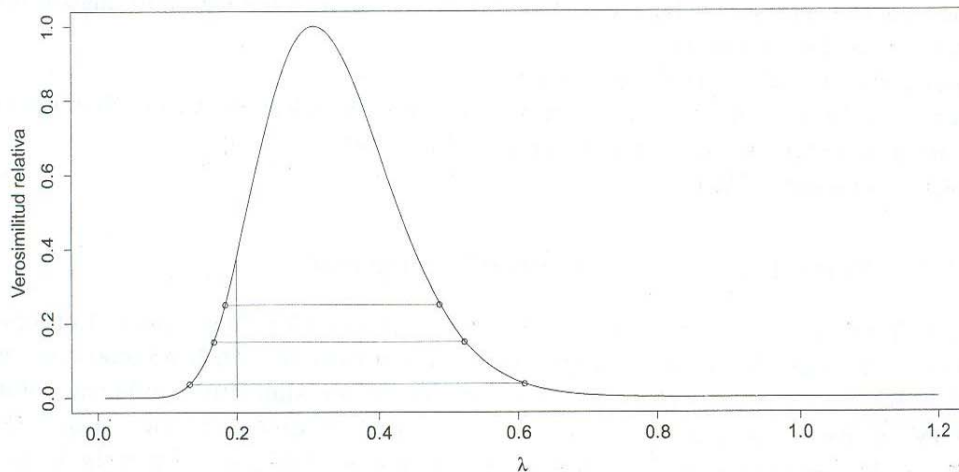


Figura 4.4: Función de verosimilitud relativa e intervalos de verosimilitud-confianza para λ (nivel de plausibilidad 0.25, 0.15 y 0.036) correspondientes a los datos simulados presentados en Tabla 4.7.

de nacimiento lineal, tanto para escenarios observacionales sin censura como con censura. En todos los casos se consideraron datos particulares, simulados con el objetivo de ejemplificar el enfoque de verosimilitud. En esta sección se explorará, usando datos simulados, el comportamiento de las frecuencias de cobertura de los intervalos de verosimilitud-confianza del parámetro λ ; así como la longitud media de estos intervalos.

Los resultados de la simulación se presentarán en dos partes, una correspondiente a el proceso de Poisson homogéneo y otra a el proceso de nacimiento lineal. En ambas partes se consideraron escenarios observacionales sin censura y con censura; así como los siguientes valores para λ_0 : 0.05, 0.2 y 1. En el proceso de Poisson homogéneo sin censura se especificaron los siguientes tamaños de muestra n : 20, 35 y 50. Para el caso del proceso de nacimiento lineal fueron: 40, 100, 400 y 500. Este último proceso de salto inició en $x_0 = 1$, es decir, $X(0) = x_0 = 1$. Por otro lado, en el caso del proceso de Poisson homogéneo con censura se determinó el tiempo máximo T_k (ya que en este caso las inferencias solo dependen de T_k y del número de saltos hasta T_k) como base en un porcentaje del valor esperado del tiempo de ocurrencia de n saltos: $T'_k = 0.65(n/\lambda_0)$ y $T''_k = 0.5(n/\lambda_0)$. En el caso del proceso de nacimiento lineal con censura se utilizó el mismo criterio anterior para fijar el tiempo máximo T_k . Luego, para determinar los otros tiempos fijos de censura se hizo una partición equidistante del intervalo $[T_0, T_k]$ en $k = 4$ subintervalos $[T_0, T_1]$, $(T_1, T_2]$, $(T_2, T_3]$ y $(T_3, T_4]$, donde $T_0 = 0$ y $X(T_0) = x_0 = 1$.

En este estudio de simulación se consideraron intervalos de verosimilitud-confianza con una probabilidad de cobertura nominal $1 - \alpha = 0.95$ correspondiente a un nivel

de verosimilitud de $c = 0.15$. La frecuencia de cobertura y la longitud media de los intervalos se estimó con base en $M = 10000$ repeticiones del siguiente algoritmo:

Paso 1: simular la muestra.

Paso 2: calcular el intervalo de verosimilitud.

Paso 3: registrar si el intervalo cubre o no al valor verdadero de λ (λ_0). Si el intervalo cubre a λ_0 entonces calcular la longitud del intervalo.

Paso 4: regresar al Paso 1.

4.5.1. Parte 1: proceso de Poisson homogéneo

En la Tabla 4.9 se muestran los resultados del estudio simulación para el proceso de Poisson homogéneo. Se observa que tanto para el caso sin censura como con censura, la frecuencia de cobertura de los intervalos de verosimilitud-confianza, columna FCIV, es aproximadamente el valor de la probabilidad de cobertura nominal (0.95), para todos los valores de λ_0 y escenarios explorados. Además, para cada λ_0 fijo, en el caso sin censura se observó que la estimación puntual de la longitud media del intervalo de verosimilitud-confianza, columna LMIV, decrece conforme aumenta el tamaño de la muestra. Esta misma situación ocurre en el caso con censura cuando se incrementa el tiempo máximo T_k . Es decir, cuando se observó más tiempo el proceso. Por último se observa que para cada valor de λ_0 , la estimación puntual de la longitud media del intervalo de verosimilitud-confianza obtenida en los escenarios sin censura es menor que la correspondiente longitud media estimada en el caso con censura.

λ_0	Sin censura			Con censura					
	n	FCIV	LMIV	T'_k	FCIV	LMIV	T''_k	FCIV	LMIV
0.05	20	0.9480	0.0454	260	0.9488	0.0555	200	0.9456	0.0639
0.2	20	0.9500	0.1823	65	0.9485	0.2225	50	0.9433	0.2558
1	20	0.9430	0.9093	13	0.9436	1.1107	10	0.9377	1.2762
0.05	35	0.9485	0.0337	455	0.9389	0.0415	350	0.9433	0.0477
0.2	35	0.9463	0.1350	113.75	0.9419	0.1662	87.50	0.9481	0.1907
1	35	0.9473	0.6748	22.75	0.9418	0.8312	17.50	0.9463	0.9546
0.05	50	0.9451	0.0281	650	0.9422	0.0346	500	0.9542	0.0395
0.2	50	0.9465	0.1122	162.50	0.9448	0.1385	125	0.9539	0.1582
1	50	0.9523	0.5596	32.50	0.9487	0.6915	25	0.9526	0.7896

Tabla 4.9: Resumen de resultados del estudio de simulación para el caso del proceso de Poisson homogéneo.

4.5.2. Parte 2: proceso de nacimiento lineal

En la Tabla 4.10 se muestran los resultados del estudio simulación para el proceso de nacimiento lineal. Se observa que para el caso sin censura, la frecuencia de cobertura de los intervalos de verosimilitud-confianza, columna FCIV, es aproximadamente

el valor de la probabilidad de cobertura nominal (0.95), en todos los escenarios explorados. En contraste, para el caso con censura la frecuencia de cobertura de los intervalos de verosimilitud-confianza se encuentra por debajo de la probabilidad de la cobertura nominal, en todos los escenarios explorados. Nótese que en este último caso la frecuencia de cobertura es aproximadamente la misma para T'_k y T''_k (la especificación previa de los tiempos fijos que resultan de las particiones de $[0, T'_k]$ y $[0, T''_k]$), y aumenta conforme se incrementa los valores de estas cantidades, para cada λ_0 fijo. Además, para cada λ_0 fijo, en el caso sin censura se observó que la estimación puntual de la longitud media del intervalo de verosimilitud-confianza, columna LMIV, decrece conforme aumenta el tamaño de la muestra. Esta misma situación ocurre en el caso con censura cuando se incrementa el tiempo máximo T_k . Es decir, cuando se observó más tiempo el proceso. Por último se observa que para cada valor de λ_0 , la estimación puntual de la longitud media del intervalo de verosimilitud-confianza obtenida en los escenarios sin censura es menor que la correspondiente longitud media estimada en el caso con censura.

λ_0	Sin censura			Con censura					
	n	FCIV	LMIV	T'_k	FCIV	LMIV	T''_k	FCIV	LMIV
0.05	40	0.9340	0.0305	51.64	0.9038	0.0612	43.03	0.8729	0.0745
0.2	40	0.9418	0.1220	13.98	0.8979	0.2218	10.75	0.8683	0.2964
1	40	0.9368	0.6086	2.79	0.9016	1.1142	2.15	0.8655	1.4872
0.05	100	0.9416	0.0193	67.56	0.9179	0.0438	51.97	0.9026	0.0605
0.2	100	0.9438	0.0772	16.89	0.9156	0.1761	12.99	0.9048	0.2433
1	100	0.9406	0.3864	3.38	0.9178	0.8762	2.60	0.9008	1.2095
0.05	400	0.9487	0.0097	85.44	0.9076	0.0303	65.72	0.9102	0.0453
0.2	400	0.9471	0.0388	21.36	0.9139	0.1217	16.43	0.9166	0.1813
1	400	0.9476	0.1941	4.27	0.9084	0.6119	3.27	0.9132	0.9103
0.05	700	0.9491	0.0073	92.70	0.9215	0.0262	71.30	0.9260	0.0406
0.2	700	0.9480	0.0294	23.17	0.9241	0.1056	17.83	0.9192	0.1623
1	700	0.9488	0.1470	4.63	0.9197	0.5279	3.57	0.9204	0.8121

Tabla 4.10: Resumen de resultados del estudio de simulación para el caso del proceso de nacimiento lineal.

Capítulo 5

Conclusiones

En esta tesis se mostró el uso del enfoque de verosimilitud para hacer inferencia estadística sobre los parámetros involucrados en las tasas de intensidad de dos procesos markovianos de salto, el proceso de Poisson y el proceso de nacimiento lineal, con base en escenarios observacionales tanto sin censura (información completa del proceso) como con censura (información parcial del proceso). Además en ambos escenarios se exploró, mediante estudios de simulación, la frecuencia de cobertura y la longitud media de los intervalos de verosimilitud-confianza.

En relación al caso de simulación del capítulo anterior se concluye que:

1. Proceso de Poisson

- Tanto para el caso sin censura como con censura, la frecuencia de cobertura de los intervalos de verosimilitud-confianza es aproximadamente el valor de la probabilidad de cobertura nominal (0.95), en todos los escenarios explorados. Es importante mencionar aquí que para el caso sin censura se requiere un tamaño de muestra pequeño ($n = 20$) para alcanzar de manera aproximada la cobertura nominal.
- En el caso sin censura, la estimación puntual de la longitud media del intervalo de verosimilitud-confianza decrece conforme aumenta el tamaño de la muestra, en todos los escenarios explorados. Esta misma situación ocurre en el caso con censura cuando se incrementa el tiempo máximo de observación.
- La estimación puntual de la longitud media del intervalo de verosimilitud-confianza obtenida en los escenarios sin censura es menor que la correspondiente longitud media estimada en el caso con censura.

2. Proceso de nacimiento lineal

- En el caso sin censura, la frecuencia de cobertura de los intervalos de verosimilitud-confianza es aproximadamente el valor de la probabilidad de cobertura nominal (0.95), en todos los escenarios explorados. Esto se cumple inclusive para un tamaño de muestra pequeño de $n = 20$.
- En el caso con censura, la manera de especificar los tiempos fijos que definen la censura puede afectar negativamente la frecuencia de cobertura de los intervalos de verosimilitud-confianza; es decir, pueden producir

frecuencias de cobertura que esten por debajo de la probabilidad de la cobertura nominal.

- En el caso sin censura, la estimación puntual de la longitud media del intervalo de verosimilitud-confianza decrece conforme aumenta el tamaño de la muestra.
- La estimación puntual de la longitud media del intervalo de verosimilitud-confianza obtenida en los escenarios sin censura es menor que la correspondiente longitud media estimada en el caso con censura.

Con base en lo presentado en este trabajo, se concluye que es importante explorar a través de simulaciones las propiedades estadísticas de las inferencias que se hagan sobre los parámetros involucrados en los procesos markovianos de salto, independientemente del metodo estadístico elegido para hacer inferencias.

Bibliografía

- [1] BAILEY, N.T.J. (1964). *The Elements of Stochastic Processes with Applications to the Natural Sciences*. New York: Jhon Wiley & Sons.
- [2] BILLINGSLEY, P. (1979). *Probability and Measure*. New York: Jhon Wiley & Sons.
- [3] BRÉMAUD, P. (1999). *Markov Chains Gibbs Fields, Monte Carlo Simulation, and Queues*. New York: Springer-Verlag.
- [4] CASELLA, G. y BERGER, R.L. (2002). *Statistical Inference*. California: Duxbury.
- [5] HOEL, P.G., PORT, S.C. y STONE, C.J. (1971). *Introduction to Probability Theory*. Boston: Houghton Mifflin Company.
- [6] HOEL, P.G., PORT, S.C. y STONE, C.J. (1972). *Introduction to Stochastic Processes*. Boston: Houghton Mifflin Company.
- [7] MONTOYA, J.A. (2008). *La Verosimilitud Perfil en la Inferencia Estadística*. Tesis de Doctorado en Ciencias con Especialidad en Probabilidad y Estadística, CIMAT, Guanajuato, Gto., México.
- [8] ROSS, S.M. (1996). *Stochastic Processes*. New York: Jhon Wiley & Sons.
- [9] SERFLING, R. (1980). *Aproximation Theorems of Mathematical Statistics*. New York: Jhon Wiley & Sons.
- [10] SMALL, C.G. (2007). *Functional Equations and How to Solve Them*. New York: Springer-Verlag.
- [11] SPROTT, D.A. (2000). *Statistical Inference in Science*. New York: Springer-Verlag.

