



"El saber de mis hijos  
hará mi grandeza"

---

# UNIVERSIDAD DE SONORA

DIVISIÓN DE CIENCIAS EXACTAS Y NATURALES

Programa de Posgrado en Matemáticas

Verosimilitud para Inferencia en Procesos de Poisson

## T E S I S

Que para obtener el grado académico de:

**Maestro en Ciencias**  
(Matemáticas)

Presenta:

María Elena Conde Hernández

Director de Tesis: Dr. Miguel Nakamura Savoy

Hermosillo, Sonora, México,      29 de Junio, 2009.

# 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

QA274.42

.C65

RIS, T1536

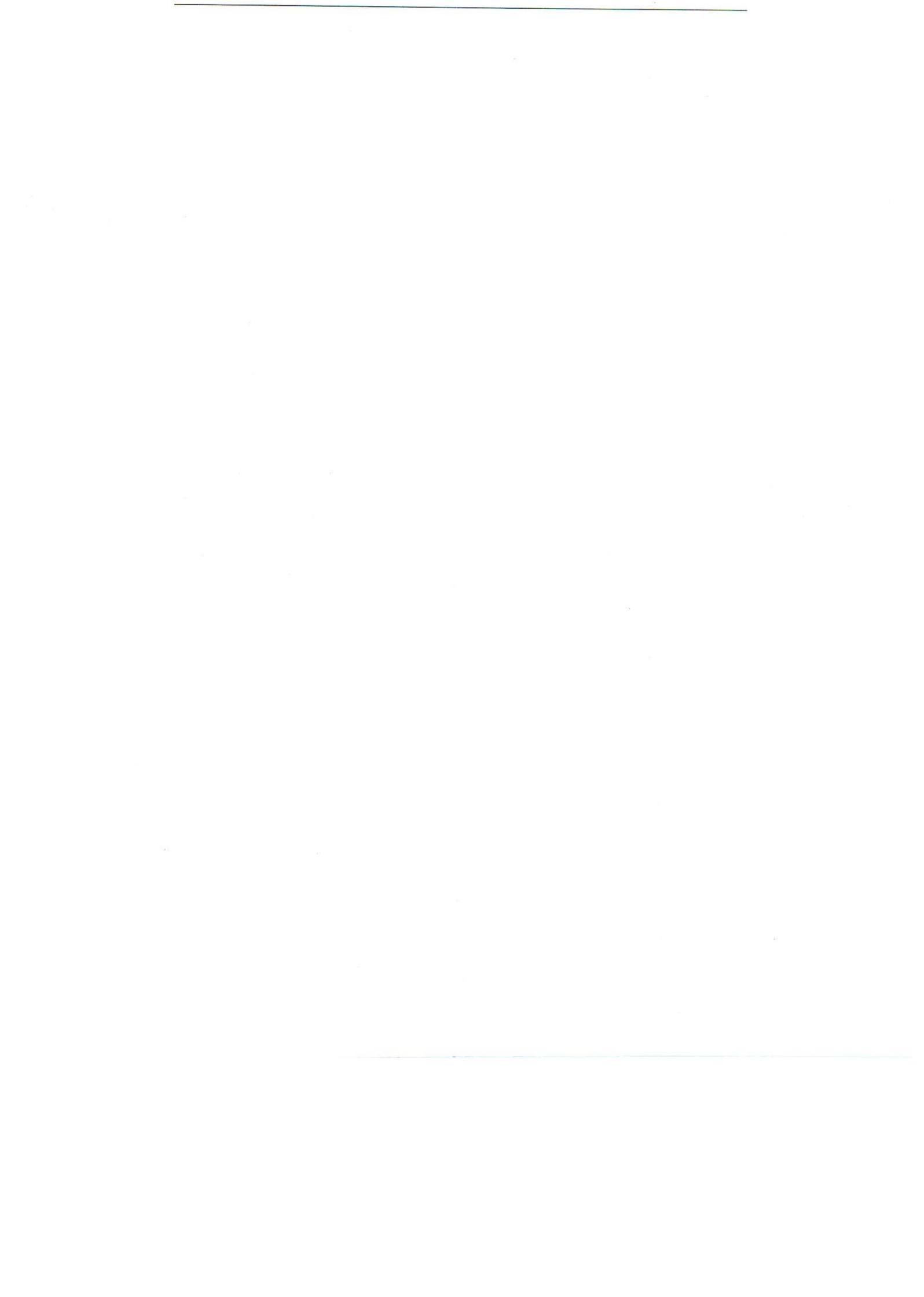
## SINODALES

Dr. Fernando Luque Vásquez  
Departamento de Matemáticas, Universidad de Sonora

Dr. José Arturo Montoya Laos  
Departamento de Matemáticas, Universidad de Sonora

Dr. Miguel Nakamura Savoy  
Centro de Investigación en Matemáticas, Guanajuato, Gto.

M. C. María Teresa Robles Alcaraz  
Departamento de Matemáticas, Universidad de Sonora



UNIVERSIDAD DE SONORA

Acta de Examen de Grado



EL SABER DE MIS HIJOS  
HARA MI GRANDEZA

En la ciudad de Hermosillo, Sonora, México, siendo las 12:00 horas del día 29 de Junio del 2009, se reunieron en:

El Auditorio del Departamento de Matemáticas  
de la Universidad de Sonora, los señores:

Dr. Fernando Luque Vásquez, Dr. José Arturo Montoya Laos  
Dr. Miguel Nakamura Savoy  
M.C. Ma. Teresa Robles Alcaraz

bajo la Presidencia del primero y fungiendo como Secretario el último, para efectuar el Examen de Grado a:

MARIA ELENA CONDE HERNANDEZ

de Maestro en Ciencias (Matemáticas)

quien presentó el trabajo de tesis

"Verosimilitud para Inferencia Estadística en Procesos de Poisson"

Acta No. 10  
Foja 10  
Libro 1  
Exp. No. 204290132

Los sinodales replicaron a la 1a sustentante y después de debatir entre sí reservada y libremente, l 1a declararon:

Aprobada por unanimidad

Acto continuo el Presidente del Jurado le hizo saber el resultado de su examen y para constancia se levanta la presente, firmando los que intervinieron.

  
Firma del sustentante

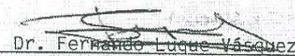
  
PRESIDENTE

Mo. Tania Robles A.  
SECRETARIO

  
VOCAL

Dr. Fernando Luque Vásquez Coordinador  
del Programa de: Posgrado en Matemáticas  
de la Universidad de Sonora, hace constar que las firmas que anteceden corresponden a  
los sinodales que intervinieron en el examen que contiene la presente acta.

Hermosillo, Sonora, a 29 de Junio del 2009.

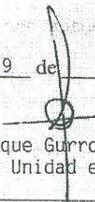
  
Dr. Fernando Luque Vásquez  
Coordinador del Programa



DIVISIÓN DE CIENCIAS  
EXACTAS Y NATURALES  
COORDINACIÓN  
POSGRADO EN MATEMÁTICAS

Ing. Enrique Gurrola Mac Secretario de Unidad en funciones de Vicerrector  
de la Unidad Regional Centro de la Universidad de Sonora, hace  
constar que: el Dr. Fernando Luque Vásquez  
es a la fecha de expedición de este documento, Coordinador del Programa arriba citado  
y suya la firma que aparece al calce del mismo.

Hermosillo, Sonora, a 29 de Junio del 2009.

  
Ing. Enrique Gurrola Mac  
Secretario de Unidad en funciones de



AL SABER DE MUY BUEN  
PARA EL SEÑOR  
UNIDAD REGIONAL  
CENTRO  
VICERRECTORIA

**VICERRECTOR** de la  
**UNIDAD REGIONAL CENTRO**

# Dedicatoria

A la memoria de mi querido padre Jesús Conde y de mi querido hermano Ricardo. Promesa cumplida. Los extraño enormemente y siempre los llevo en mis pensamientos.

A mis amores Gerardo y Edgar, por tantas horas que no pude estar con ellos a causa de este trabajo y que supieron sobrellevar con paciencia mis ausencias.

A mi esposo René por su apoyo incondicional.

A mi madre por sus oraciones y a mi hermano Edgar, que tanto quiero; por su apoyo estoy aquí. A mis hermanas Yadira y Bety, por todas las risas y alegrías, incluso en los momentos más difíciles.



# Agradecimientos

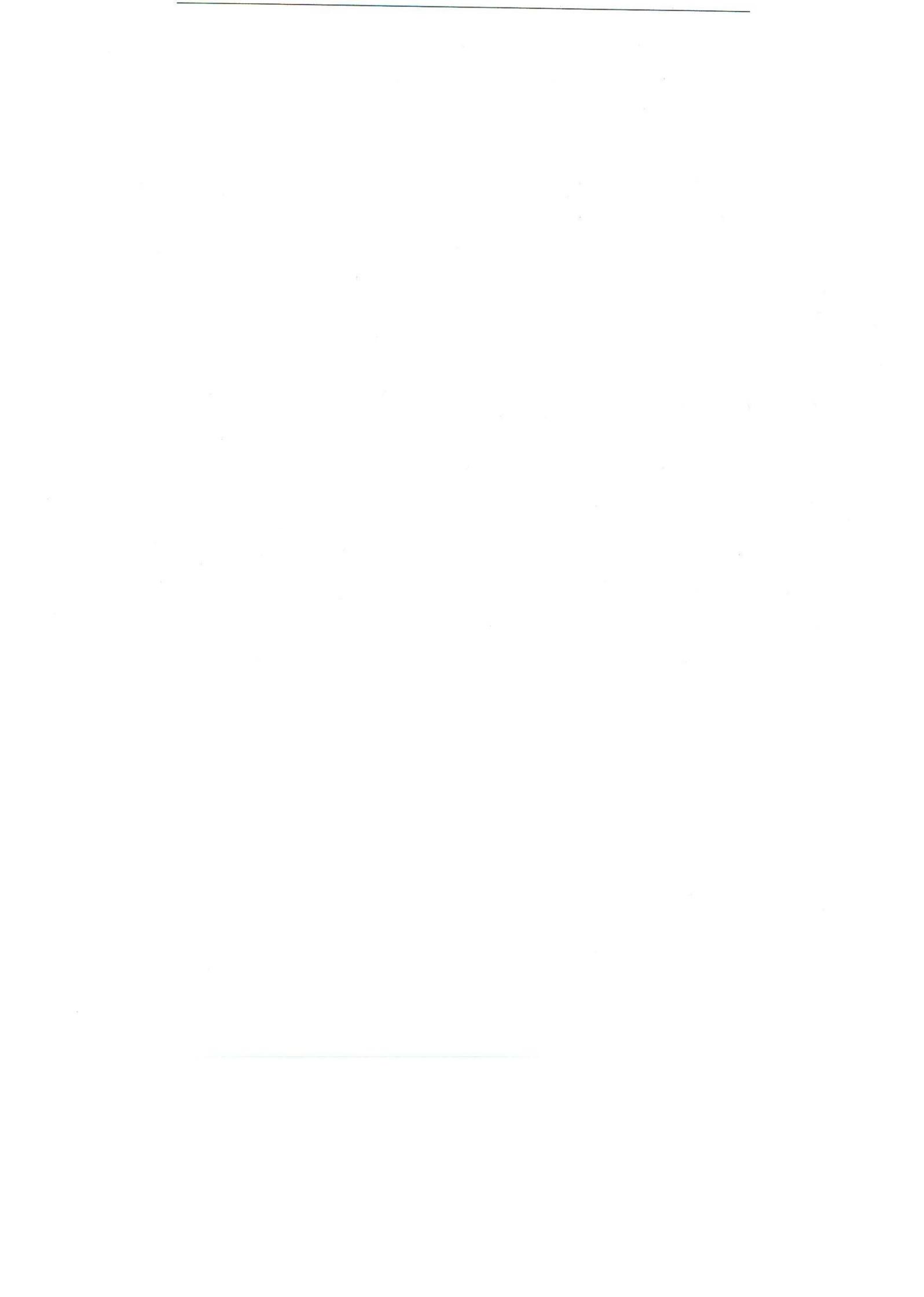
Agradezco con todo mi corazón a mi Dios que día a día me brinda fortaleza a través de los seres que más quiero y atesoro, quienes han estado a mi lado en silencio, brindando su apoyo incondicional, mis hijos y mi esposo.

Al Dr. Miguel Nakamura, con una gran admiración y respeto. Agradezco sus enseñanzas y su paciencia en la dirección de este trabajo.

A mis maestros asesores y amigos Tere, José Montoya y Gudelia. Mil gracias por compartir sus conocimientos conmigo. También agradezco su paciencia y el apoyo brindado durante todo este tiempo.

A mis maestros sinodales Dr. Fernando Luque Vázquez, M. C. María Teresa Robles Alcaraz y Dr. José Arturo Montoya. Gracias por sus aportaciones a este trabajo; realmente lograron un crecimiento en mi persona durante el seminario de tesis.

A mi querida amiga Cecilia. Gracias por no desfallecer en esas interminables noches de estudio conjunto y por los fines de semana de trabajo hombro con hombro.



# Contenido

Introducción	xiii
<b>1 Procesos de Poisson</b>	<b>1</b>
1.1 Proceso de Poisson espacial . . . . .	1
1.2 El Proceso de Poisson en la recta real . . . . .	5
1.3 Proceso de Poisson no homogéneo . . . . .	18
1.4 Comentarios finales . . . . .	20
<b>2 Inferencia estadística en procesos de Poisson</b>	<b>23</b>
2.1 Función de verosimilitud . . . . .	23
2.2 Inferencia en procesos de Poisson homogéneos . . . . .	35
2.3 Inferencia en procesos de Poisson no homogéneos . . . . .	40
<b>3 Simulación de procesos de Poisson y validación de probabilidades de cobertura</b>	<b>47</b>
3.1 Generación de un proceso de Poisson homogéneo . . . . .	47
3.2 Generación de un proceso de Poisson no homogéneo . . . . .	53
<b>4 Caso de estudio</b>	<b>65</b>
Conclusiones	75
Bibliografía	77



## Introducción

Muchos fenómenos físicos se modelan estocásticamente de la misma manera en el tiempo como en el espacio, tal es la distribución aleatoria de partículas en una pieza de material radiactivo, los tiempos de llegadas de llamadas telefónicas, el rompimiento de cromosomas bajo irradiación, los tiempos de fallas mecánicas de una máquina, la distribución de cierta especie arbolaria o el número de terremotos en una superficie, la distribución de galaxias en el sistema estelar, la distribución de centros de poblaciones de animales o epidemias, etc.

En este trabajo de tesis estudiaremos procesos que consisten en una sucesión de eventos que ocurren en el tiempo, y nuestro interés será contar el número de dichos eventos en un intervalo de tiempo dado, cada ocurrencia de un evento lo representaremos como un punto en el eje del tiempo. Pediremos a un nivel intuitivo que cumplan con las siguientes condiciones: La información con respecto al número de eventos, que ocurren en intervalos de tiempo que no se traslapan son estocásticamente independientes; las condiciones experimentales permanecen constantes en el tiempo, la ocurrencia de eventos en un intervalo solo depende de la longitud y no de la posición del intervalo, si ésta condición no se cumple se dice que el proceso es no homogéneo en el tiempo. Siempre se espera un tiempo positivo, pero finito entre cualesquiera dos eventos, y solo ocurren un número finito de eventos en un intervalo de tiempo finito. Bajo éstas condiciones resulta natural pensar que las trayectorias del proceso que cuenta el número de eventos, son no decrecientes, escalonadas con saltos de tamaño uno y acotadas en cada intervalo finito.

Por tanto, los fenómenos aleatorios mencionados con anterioridad pueden ser modelados como un proceso de Poisson con cierta tasa o intensidad de ocurrencia, que involucra el o los parámetros del proceso, la cual puede ser constante, es decir, ser un proceso de Poisson homogéneo, o ser no homogéneo, esto es, que la intensidad de ocurrencia esté en función del tiempo.

En esta tesis, se pretenden varios objetivos. El primero, la caracterización de la función de intensidad de un proceso de Poisson no homogéneo. Segundo, aplicar el enfoque de verosimilitud a procesos de Poisson, realizando inferencia sobre los parámetros de la función de intensidad mediante intervalos de verosimilitud. Tercero, simular procesos de Poisson tanto homogéneos como no homogéneos, para luego validar las propiedades frecuentistas de la verosimilitud en el cálculo de las probabilidades de cobertura de los intervalos de verosimilitud, como una aproximación a los intervalos de confianza. Cuarto y último objetivo, aplicar los resultados de probabilidad y estadística desarrollados en este trabajo, a un problema en meteorología mediante una variable de interés que denominaremos "*Rachas de calor*", que es determinada por las temperaturas máximas diarias que exceden los 31° C.

El interés en esta situación es seleccionar un modelo probabilístico para modelar

los eventos "*Rachas de calor*" registradas por una estación meteorológica en una región específica del centro del país. La hipótesis que aquí se sustenta es, que este tipo de fenómenos puede ser modelado como un proceso de Poisson cuyos parámetros varían con el tiempo de manera no homogénea con cierta función de intensidad proporcional a una gama.

Se han contemplado cuatro capítulos en la organización de este trabajo. En el primer capítulo se introduce el proceso de Poisson basándose principalmente en Resnick[22] y Billingsley[2]. El Capítulo 2, consiste de tres secciones; en la primera sección se revisan algunos conceptos básicos de la teoría de verosimilitud que serán utilizados a lo largo de esta tesis, basados en Kalbfleisch[11] y Sprott[25], estudiaremos la función de verosimilitud y sus propiedades, la obtención e interpretación de los intervalos de verosimilitud, validando las propiedades frecuentistas del método de verosimilitud en el cálculo de las probabilidades de cobertura de los intervalos de verosimilitud, así como la relación que hay entre éstos últimos y los intervalos de confianza. En las Secciones 2 y 3, de acuerdo a Basawa[1], trabajaremos con los planes de inspección de un proceso de Poisson homogéneo en diferentes situaciones de inspección, por ejemplo la observación de un proceso de Poisson en un intervalo de tiempo fijo, así como en un intervalo de tiempo aleatorio; para terminar con las propiedades de importancia de los procesos de Poisson no homogéneos, respectivamente.

El Capítulo 3 contiene el material operativo de simulación de datos, basado en las teorías de probabilidad y estadística desarrolladas en los capítulos precedentes. Mediante métodos numéricos, realizaremos la inferencia estadística bajo la teoría de verosimilitud a procesos de Poisson homogéneos como no homogéneos, tanto de un parámetro como de dos, con la finalidad de entender e ilustrar el comportamiento de los parámetros involucrados en la caracterización de la función de intensidad propuesta, así como algunas conclusiones resultado de su aplicación y la corroboración empírica de su funcionalidad. Al final, en el Capítulo 4, se estimula la aplicación de los conceptos y resultados obtenidos en los capítulos previos mediante un caso de estudio aplicado en meteorología, a través de la variable de interés "*Rachas de calor*".

Las simulaciones realizadas en este trabajo de tesis, la generación de procesos de Poisson, el cálculo de los estimadores, así como los intervalos de verosimilitud, probabilidades de cobertura y todas las figuras de las funciones mostradas aquí, se realizaron con los programas "R" versión 2.3.1 y matricial MATLAB versión 6.5, en una computadora personal Acer Aspire 3000.

# Capítulo 1

## Procesos de Poisson

Fundamentalmente en este capítulo se persigue el objetivo de deducir las condiciones, propiedades y características que establecerán las bases probabilísticas para nutrir de validez la hipótesis que se sustenta en el caso de estudio: los eventos “*Rachas de calor*”, pueden ser modelados como un proceso de Poisson no homogéneo. Para sustentar de manera intuitiva esta hipótesis, podemos pensar que la distribución del número de rachas de calor en intervalos de tiempo que no se traslapan son estocásticamente independientes, es decir, que el número de rachas que se presentan en un intervalo de tiempo no depende del número de rachas que se presentan en otro; el tiempo que transcurre entre racha y racha es estrictamente positivo (porque si no estaríamos hablando de una sola racha). Además es natural esperar que sólo una cantidad finita de rachas de calor se presenten en un intervalo finito de tiempo, por lo que de acuerdo a las condiciones intuitivas que ya hemos comentado, podríamos pensar que el proceso de Poisson es un modelo pertinente para las rachas de calor.

En este capítulo consideraremos un modelo probabilístico para la distribución aleatoria de partículas en el espacio y eventos en la recta real. En la primera sección presentaremos la definición del proceso de Poisson en el espacio; en la segunda, estudiaremos el caso particular del proceso de Poisson homogéneo en la recta real, donde trataremos diferentes caracterizaciones del mismo. Finalmente en las secciones tres y cuatro, obtendremos un proceso de Poisson no homogéneo a partir de un proceso homogéneo mediante una transformación para terminar con comentarios finales, esta presentación está basada en las referencias Billingsley[2], Cinlar[6], Davenport[8], Karlin[12], Karlin & Howard[13], Kulkarni[14], Parzen[18] y Resnick[22].

### 1.1 Proceso de Poisson espacial

El objetivo en esta sección es introducir un modelo para la distribución aleatoria de puntos en el espacio a través de una reseña de resultados del trabajo publicado en 1992 por Resnick[22] basado en procesos de Poisson, no sin antes establecer primeramente los elementos de donde partiremos.

Consideremos  $(\Omega, \mathcal{F}, P)$  un espacio de probabilidad. Sea  $(E, \xi)$  un espacio medible; donde  $E$  es el espacio euclidiano de estados donde viven los puntos aleatorios del modelo, el cual será usualmente un subconjunto de  $\mathbb{R}^d$ ,  $d \geq 1$ ; y sea  $\xi$  la  $\sigma$ -álgebra de Borel correspondiente. El propósito es modelar la distribución aleatoria de puntos en el espacio de estados  $E$ , e iniciaremos definiendo lo que es un proceso puntual.

Sea  $\{X_n\}_{n \geq 1}$  una sucesión de variables aleatorias

$$X_n : \Omega \rightarrow E;$$

que representan los puntos aleatorios en el espacio de estados  $E$ .

**Definición 1.1** Para cada conjunto  $A \in \xi$ , acotado,  $n \geq 1$  y  $\omega \in \Omega$  fijo, definimos la *medida discreta*  $\epsilon_{X_n(\cdot)} : \xi \rightarrow \mathbb{R}$ ,

$$\epsilon_{X_n(\omega)}(A) = \begin{cases} 1 & \text{si } X_n(\omega) \in A \\ 0 & \text{si } X_n(\omega) \notin A \end{cases}.$$

Entonces sumando sobre  $n$  obtenemos el número total de puntos aleatorios  $X_n$  que están en  $A$ , como lo expresa la siguiente definición.

**Definición 1.2** Una *medida de conteo*  $N$  en  $E$ , es una medida tal que para cada  $A \in \xi$  acotado,

$$N(A) := \sum_n \epsilon_{X_n}(A).$$

Notemos que la sucesión de puntos aleatorios  $\{X_n\}$  tiene como dominio el espacio muestral  $\Omega$ , entonces  $N$  es así una función con dominio  $\Omega$ . Por lo tanto, para  $\omega \in \Omega$ , escribimos

$$N(\omega, A) := \sum_n \epsilon_{X_n(\omega)}(A),$$

aunque por lo general suprimiremos en la notación la dependencia sobre  $\omega$ . Entonces para cada  $A \in \xi$  fijo;  $N(\cdot, A) = N(A)$  es una variable aleatoria que representa el número aleatorio de puntos  $\{X_n\}_{n \geq 1}$  que caen en  $A$ . Con estos recursos ya podemos establecer la siguiente definición.

**Definición 1.3** Un *proceso de conteo*  $N$  es una medida de conteo

$$N(\cdot) := \sum_n \epsilon_{X_n}(\cdot), \tag{1.1}$$

donde  $\{X_n\}_{n \geq 1}$  es el proceso de puntos aleatorios asociado al proceso de conteo; tal que para cualquier  $A \in \xi$  acotado,  $P[N(A) < \infty] = 1$ .

**Ejemplo 1.1** Consideremos un modelo de localización de terremotos y los tiempos en que éstos ocurren. El espacio de estados es  $E = [0, \infty) \times \mathbb{R}^2$ , si  $\tau_n$  es el tiempo en que ocurre el  $n$ -ésimo temblor, entonces el proceso puntual está expresado por:

$$N(\cdot) = \sum_n \epsilon_{(\tau_n, (L_{n1}, L_{n2}))}(\cdot),$$

donde  $(L_{n1}, L_{n2})$  representa las coordenadas latitud y longitud de la localización del temblor. Así para un tiempo  $t > 0$  y un conjunto  $B \in \beta(\mathbb{R}^2)$ ; el proceso puntual  $N([0, t] \times B)$  representa el número de temblores que ocurren en el intervalo de tiempo  $[0, t]$  localizados en la región  $B$ .

Habiendo ya considerado los elementos básicos y necesarios para la definición de lo que es un proceso puntual, es tiempo de establecer lo que es un proceso de Poisson.

Sea  $\mu$  una medida en  $(E, \xi)$ , finita para conjuntos acotados de  $\xi$ .

**Definición 1.4** *Un proceso de Poisson con media  $\mu$ , es un proceso de conteo  $N$  con un espacio de estados  $E$ , tal que:*

1. Para  $A \in \xi$ ;

$$P[N(A) = k] = \begin{cases} \frac{e^{-\mu(A)}(\mu(A))^k}{k!}, & \text{si } \mu(A) < \infty \\ 0, & \text{si } \mu(A) = \infty \end{cases}$$

2. Si  $A_1, A_2, \dots, A_n$  son subconjuntos ajenos en  $\xi$ , entonces  $N(A_1), \dots, N(A_n)$  son variables aleatorias independientes.

Por lo tanto,  $N$  es un proceso de Poisson si el número aleatorio de puntos en un conjunto  $A$  está distribuido como una variable aleatoria de Poisson con parámetro  $\mu(A)$ ; y el número de puntos en regiones ajenas son variables aleatorias independientes.

**Observación 1.2** *Si  $E = \mathbb{R}$ ; la Condición (2) de la definición anterior es conocida como "incrementos independientes", ya que para tiempos  $0 \leq t_0 < t_1 < t_2 < \dots < t_n < \dots < t_k$  tenemos que  $N((t_i, t_{i+1}])$ ,  $i = 1, 2, \dots, k - 1$  son variables aleatorias independientes.*

**Observación 1.3** *Cuando la medida  $\mu$  es un múltiplo de la medida de Lebesgue (es decir, la longitud, cuando  $E = [0, \infty)$  o  $\mathbb{R}$ , área, cuando  $E = \mathbb{R}^2$ , volumen, cuando  $E = \mathbb{R}^3$ , etc), llamamos al proceso de Poisson **homogéneo**. Así en el caso homogéneo, existe un parámetro  $\lambda > 0$  tal que para cualquier  $A \in \xi$ , tenemos que  $N(A)$  tiene una distribución de Poisson con media  $E[N(A)] = \lambda|A|$ , donde  $|A|$  es la medida de Lebesgue de  $A$ .*

**Observación 1.4** *Cuando  $E = [0, \infty)$  y  $N$  es un proceso de Poisson homogéneo, tenemos que  $E[N((0, t])] = \lambda|(0, t]| = \lambda t$ ; el parámetro  $\lambda$  puede ser interpretado como el número promedio de eventos ocurridos por unidad de tiempo.*

### 1.1.1 Transformaciones del proceso Poisson

Algunos resultados útiles surgen al considerar un proceso de Poisson bajo diferentes transformaciones, por ejemplo obtener un proceso de Poisson no homogéneo de un proceso homogéneo mediante el uso de una transformación. Las siguientes definiciones y resultados se pueden consultar en la Sección 4.2 de Resnick[22].

Sea  $T$  una transformación medible

$$T : E \rightarrow E',$$

con  $E$  y  $E'$  espacios euclidianos.

**Definición 1.5** Se define la función  $T^{-1}$  como un mapeo medible conjuntista de subconjuntos de  $E'$  en subconjuntos de  $E$ . Para  $A' \subset E'$  tenemos

$$T^{-1}(A') = \{e \in E : T(e) \in A'\}.$$

De esta manera  $T^{-1}(A')$  es la preimagen de  $A'$  bajo  $T$ ; esto es, el conjunto de puntos de  $E$  los cuales  $T$  mapea dentro de  $A'$ .

Sea  $N$  un proceso de Poisson con media  $\mu$ ,  $E$  el espacio de estados y  $\{X_n\}_{n \geq 1}$  el proceso de puntos asociado. Sea  $T : E \rightarrow E'$  una transformación medible; de las Definiciones 1.3 y 1.4 consideremos las medidas  $N$  y  $\mu$  en  $(E, \xi)$ , y definimos las medidas inducidas  $N'$  y  $\mu'$  en  $E'$  de la siguiente manera.

**Definición 1.6** Para  $A' \subset E'$  definimos:

$$N'(A') := N(T^{-1}(A')) \quad (1.2)$$

$$\mu'(A') := \mu(T^{-1}(A')). \quad (1.3)$$

Consideremos  $T^{-1}(A')$  en  $E$  y tomamos las medidas  $N, \mu$  respectivamente de la preimagen de  $A'$  bajo  $T$ . Además, si el proceso  $N$  tiene asociado el proceso puntual  $\{X_n\}_{n \geq 1}$  en  $E$ , entonces  $N'$  tiene asociado el proceso puntual  $\{X'_n\}_{n \geq 1} = \{T(X_n)\}$  en  $E'$ , así para  $A' \subset E'$  tenemos que:

$$\begin{aligned} N'(A') : N(T^{-1}(A')) &= \sum_n \epsilon_{X_n(T^{-1}(A'))} = \sum_n \mathbb{I}_{[X_n \in T^{-1}(A')]} \\ &= \sum_n \mathbb{I}_{[T(X_n) \in A']} = \sum_n \epsilon_{T(X_n)}(A'). \end{aligned}$$

El siguiente resultado asegura que si  $N$  es un proceso de Poisson con media  $\mu$  y proceso puntual  $\{X_n\}_{n \geq 1}$  en el espacio de estados  $E$ , entonces  $N' = N(T^{-1}(\cdot))$  es un proceso de Poisson con media  $\mu' = \mu(T^{-1}(\cdot))$  y proceso puntual  $\{T(X_n)\}$  en  $E'$ .

**Proposición 1.1** Sea  $T : E \rightarrow E'$  un mapeo medible tal que  $T^{-1}(B')$  es un conjunto acotado en  $E$  para cada subconjunto  $B'$  acotado en  $E'$ . Si  $N$  es un proceso de Poisson con media  $\mu$  y proceso puntual  $\{X_n\}_{n \geq 1}$ , entonces  $N' := N \circ T^{-1}$  es un proceso de Poisson con media  $\mu' = \mu \circ T^{-1}$  y proceso puntual  $\{T(X_n)\}$ .

**Demostración.** Para determinar que  $N'$  es un proceso de Poisson verificaremos que se cumplan las Condiciones (1) y (2) de la Definición 1.4.

Si  $N$  es un proceso de Poisson con media  $\mu$ , para  $B' \in E'$  tenemos de (1.2) que:

$$P[N'(B') = k] = P[N(T^{-1}(B')) = k] = \frac{e^{-\mu(T^{-1}(B'))} (\mu(T^{-1}(B')))^k}{k!};$$

de esta manera  $N'$  tiene distribución de Poisson con media  $\mu'(B') = \mu(T^{-1}(B'))$ .

Para verificar la Condición (2), demostraremos que las variables aleatorias

$$N'(B'_1), N'(B'_2), \dots, N'(B'_m),$$

son independientes, siempre que  $B'_1, B'_2, \dots, B'_m$  sean conjuntos ajenos en  $E'$ . Notemos de (1.2) que para cada  $i = 1, 2, \dots, m$ ;

$$N'(B'_i) = N(T^{-1}(B'_i)).$$

Además se cumple que,

$$T^{-1}(B'_i) \cap T^{-1}(B'_j) = \emptyset, \text{ para } i \neq j;$$

ya que si  $x \in [T^{-1}(B'_1) \cap T^{-1}(B'_2)]$ , entonces

$$x \in T^{-1}(B'_1) \text{ y } x \in T^{-1}(B'_2),$$

luego

$$T(x) \in B'_1 \text{ y } T(x) \in B'_2,$$

esto implica que

$$T(x) \in B'_1 \cap B'_2;$$

pero si  $B'_1 \cap B'_2 = \emptyset$ , entonces

$$T^{-1}(B'_1) \cap T^{-1}(B'_2) = \emptyset,$$

por lo tanto, se sigue que

$$(N'(B'_1), N'(B'_2), \dots, N'(B'_m)) = (N(T^{-1}(B'_1)), N(T^{-1}(B'_2)), \dots, N(T^{-1}(B'_m)));$$

son independientes por la Condición (2) de la Definición 1.4.

De esta manera  $N'$  cumple con las Condiciones (1) y (2) de la Definición 1.4, es decir,  $N'$  es un proceso de Poisson con media  $\mu'(\cdot) = \mu(T^{-1}(\cdot))$ . ■

## 1.2 El Proceso de Poisson en la recta real

El propósito de esta sección es presentar diferentes caracterizaciones del proceso de Poisson en la recta aprovechando la estructura de orden que presentan los reales, la cual nos permite introducir además de los procesos puntual y de conteo, el proceso que representa los tiempos entre eventos consecutivos. El contenido que desarrollaremos en esta sección está apoyado básicamente en las referencias Billingsley [2], Cinlar [6], Davenport[8] y Parzen [18].

En esta sección consideraremos el caso para  $d = 1$  de la Definición 1.4. Sea  $(\Omega, \mathcal{F}, P)$  un espacio de probabilidad,  $E = [0, \infty)$  el espacio de estados. Sea  $\{S_n\}_{n \geq 1}$ ,  $S_n : \Omega \rightarrow E$  un proceso de puntos aleatorios; en este caso  $S_n$  representa el "el tiempo en el que ocurren cada uno de los eventos en la línea del tiempo". Si dos o más puntos aleatorios no pueden ocurrir simultáneamente, entonces la sucesión de variables aleatorias  $\{S_n\}_{n \geq 1}$  es estrictamente creciente, y si sólo una cantidad finita de puntos ocurren en cada intervalo finito de tiempo, entonces  $\{S_n\}_{n \geq 1}$  satisface:

**Condición 0°.**  $S_0 = 0 < S_1 < S_2 < S_3 < \dots; \sup_n S_n = \infty$ .

Asimismo, para cada  $(a, b]$  se define la medida discreta

$$\epsilon_{S_n}((a, b]) = \begin{cases} 1 & \text{si } S_n \in (a, b] \\ 0 & \text{si } S_n \notin (a, b] \end{cases} .$$

Sumando sobre  $n$  obtenemos el proceso de conteo  $N$ , para cada intervalo  $(s, t]$  tenemos,

$$N((s, t]) = \sum_{n \geq 1} \epsilon_{S_n}((s, t]), \tag{1.4}$$

para abreviar si  $S = 0$ ; denotaremos por  $N(t)$  al proceso de conteo  $N((0, t])$ . Notemos que la Condición 0<sup>o</sup> asegura que  $N(t) < \infty$  en todo  $(0, t]$  finito.

**Condición 1<sup>o</sup>.** Un proceso de conteo  $N$  satisface la Condición 1<sup>o</sup> si cumple:

- (i)  $P[N(0) = 0] = 1$ .
- (ii) Para  $0 \leq t_0 < t_1 < t_2 < \dots < t_n$  tenemos que  $N((t_i, t_{i+1} ])$ ,  $i = 0, 1, 2, \dots, n - 1$  son variables aleatorias independientes.
- (iii) Para cada  $s, t \geq 0$  y  $\lambda > 0$

$$P[N((s, t]) = n] = \frac{e^{-\lambda(t-s)} \lambda(t-s)^n}{n!}; \quad n = 0, 1, 2, \dots$$

**Observación 1.5** La relación (iii) de la Condición 1<sup>o</sup> afirma que, el número de eventos o puntos que ocurren en un intervalo  $(s, t]$ , tiene distribución de Poisson con media  $\lambda(t - s)$ , que depende solamente de la longitud del intervalo  $(t - s)$ .

**Observación 1.6** De (1.4), para cada intervalo  $(s, t]$  se cumple:

$$\begin{aligned} N((s, t]) &:= \sum_{n \geq 1} \epsilon_{S_n}((s, t]) = \sum_{n \geq 1} \epsilon_{S_n}((0, t]) - \sum_{n \geq 1} \epsilon_{S_n}((0, s]) \\ &= N((0, t]) - N((0, s]) = N(t) - N(s); \end{aligned}$$

por lo que, el número de puntos que ocurren en el intervalo  $(s, t]$  es el incremento  $N(t) - N(s)$ .

**Observación 1.7** Para cada  $\omega \in \Omega$  fijo, tenemos las trayectorias del proceso de conteo  $t \rightarrow N(t, \omega)$  como función de  $t \in [0, \infty)$ . En particular  $N(0) = 0$ , además el intervalo  $(0, t]$  es cerrado por la derecha, entonces las trayectorias son continuas por la derecha, no decrecientes con saltos de tamaño uno en los tiempos  $t = S_n$ ;  $n = 1, 2, \dots$ . Ver Figura 1.1.

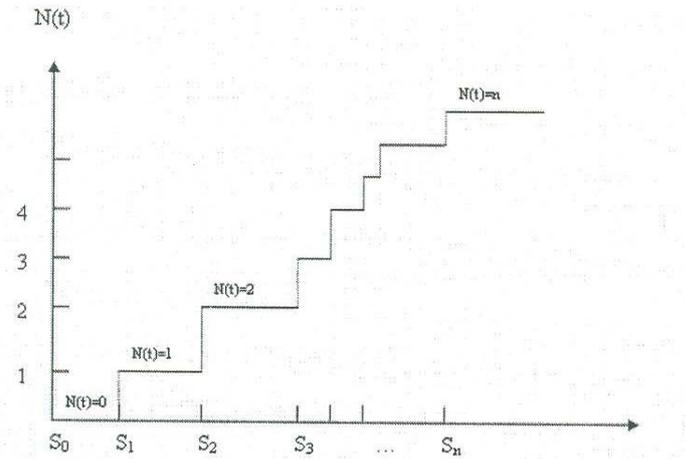


Fig. 1.1 Trayectorias del proceso  $N$  en la recta real.

De acuerdo a la Observación 1.2 de la sección previa, definimos.

**Definición 1.7** Una colección  $\{N(t)\}_{t \geq 0}$  de variables aleatorias que satisfacen la Condición 1ª, se le llama **proceso de Poisson homogéneo** con parámetro  $\lambda > 0$ .

De (1.4),  $N(t) = N((0, t]) = \sum_{n \geq 1} \epsilon_{S_n}((0, t])$ , representa el número de puntos  $S_n$  que caen en el intervalo  $(0, t]$ , entonces la variable  $N$  es el entero  $n$  más grande tal que  $S_n \leq t$ , es decir

$$N(t) = \max\{n \geq 0 : S_n \leq t\}, \quad t \geq 0. \quad (1.5)$$

$N(t) = 0$  si  $t < S_1$ ; en particular  $N(0) = 0$ .

Notemos de (1.4) que, en el intervalo  $(0, t]$  ocurren al menos  $n$  eventos si y solamente si el tiempo para que ocurra el  $n$ -ésimo evento es menor o igual a  $t$ . Entonces se cumple la siguiente igualdad de eventos que relaciona las variables  $N(t)$  y  $S_n$ :

$$[N(t) \geq n] = [S_n \leq t]. \quad (1.6)$$

Para demostrar la igualdad de eventos, sea  $\omega \in [N(t)(\omega) \geq n]$ , por (1.4) tenemos:

$$\begin{aligned} N(t)(\omega) &= \sum_{k \geq 1} \epsilon_{S_k(\omega)}((0, t]) \geq n \\ &= \sum_{k \geq 1} \mathbb{I}_{[S_k(\omega) \in ((0, t])]} \geq n; \end{aligned}$$

esto implica que al menos  $n$  eventos se encuentran en el intervalo  $(0, t]$ , es decir,  $S_n(\omega) \in ((0, t])$  por lo que  $S_n(\omega) \leq t$ , de esta manera queda demostrada la primera contención. Ahora consideremos  $\omega \in [S_n(\omega) \leq t]$ , esto implica que  $S_n(\omega) \in ((0, t])$ , luego los tiempos  $S_{n-1}(\omega), S_{n-2}(\omega), \dots, S_1(\omega) \in ((0, t])$ , es decir,

$$\sum_{k \geq 1} \mathbb{I}_{[S_k(\omega) \in ((0, t])]} \geq n = \sum_{k \geq 1} \epsilon_{S_k(\omega)}((0, t]) \geq n = [N(t) \geq n].$$

De la expresión (1.6) se sigue que

$$[N(t) = n] = [N(t) \geq n] - [N(t) \geq n + 1] = [S_n \leq t] - [S_{n+1} \leq t],$$

es decir,

$$[N(t) = n] = [S_n \leq t < S_{n+1}]; \quad n \geq 1. \quad (1.7)$$

Dado el proceso puntual  $\{S_n\}_{n \geq 1}$ , se define el proceso  $\{T_n\}_{n \geq 1}$  que llamaremos "tiempo entre eventos consecutivos", de la siguiente forma:

$$T_n := S_n - S_{n-1}; \quad n \geq 1; \quad (1.8)$$

por lo tanto

$$T_1 + T_2 + \dots + T_n = \sum_{k=1}^n T_k = S_1 + S_2 - S_1 + S_3 - S_2 + \dots + S_n - S_{n-1} = S_n.$$

Además bajo la Condición 0<sup>o</sup> se cumple:

$$\sum_{k=1}^{\infty} T_k = \lim_{n \rightarrow \infty} S_n = \infty \quad \text{y} \quad T_1 > 0, T_2 > 0, \dots \quad (1.9)$$

Así  $T_1 = S_1$  representa el tiempo entre 0 y el primer punto, para  $n > 1$ ,  $T_n$  es el tiempo entre el  $(n-1)$ -ésimo evento y el  $n$ -ésimo evento. Recíprocamente si tenemos el proceso  $\{T_n\}_{n \geq 1}$ , de los tiempos entre eventos consecutivos se define el proceso  $\{S_n\}_{n \geq 1}$ ,

$$\begin{aligned} S_0 &= 0; \\ S_n &= T_1 + T_2 + \dots + T_n; \quad n \geq 1, \end{aligned}$$

con la finalidad de estudiar las distribuciones de las variables aleatorias  $N(t)$  bajo condiciones en el proceso  $\{T_n\}_{n \geq 1}$ . Si tenemos un proceso de Poisson  $N$  y su correspondiente proceso de puntos  $\{S_n\}_{n \geq 1}$ , esto determina la distribución de las variables aleatorias  $\{T_n\}_{n \geq 1}$ , como lo establece el siguiente teorema, para mayor referencia al lector puede consultar la Sección 23 de Billingsley[2].

**Condición 2<sup>o</sup>.** Las variables aleatorias  $\{T_n\}_{n \geq 1}$  son independientes e idénticamente distribuidas con distribución exponencial con parámetro  $\lambda$ .

**Teorema 1.8** La Condición 1<sup>o</sup> es equivalente a la Condición 2<sup>o</sup> en presencia de la Condición 0<sup>o</sup>.

**Demostración. Condición 1<sup>o</sup> implica Condición 2<sup>o</sup>.**

Sea  $N$  un proceso de conteo y  $\{S_n\}_{n \geq 1}$  su correspondiente proceso de puntos, supongamos que se cumple la Condición 1<sup>o</sup>, supongamos además que el proceso  $\{S_n\}_{n \geq 1}$  satisface la Condición 0<sup>o</sup>. Entonces de (1.7) y de la relación (ii) de la Condición 1<sup>o</sup> tenemos

$$P[S_1 = T_1 > t] = P[N(t) = 0] = e^{-\lambda t},$$

así  $T_1$  es una variable aleatoria con distribución exponencial de parámetro  $\lambda$ .

Para encontrar la distribución conjunta de  $T_1$  y  $T_2$ , suponemos que  $0 \leq s_1 < t_1 < s_2 < t_2$ ,

$$\begin{aligned} P[s_1 < S_1 \leq t_1, s_2 < S_2 \leq t_2] &= P[N(s_1) = 0, N(t_1) - N(s_1) = 1, N(s_2) - N(t_1) = 0, N(t_2) - N(s_2) \geq 1] \\ &= e^{-\lambda s_1} \times \lambda(t_1 - s_1)e^{-\lambda(t_1 - s_1)} \times e^{-\lambda(s_2 - t_1)} \times (1 - \lambda e^{-\lambda(t_2 - s_2)}) \\ &= \lambda(t_1 - s_1)(e^{-\lambda s_2} - e^{-\lambda t_2}). \end{aligned}$$

De esta manera

$$P[s_1 < S_1 \leq t_1, s_2 < S_2 \leq t_2] = \iint_{\substack{s_1 < y_1 \leq t_1 \\ s_2 < y_2 \leq t_2}} \lambda^2 e^{-\lambda y_2} dy_1 dy_2.$$

De aquí que el vector  $(S_1, S_2)$  tiene densidad conjunta  $f(y_1, y_2) = \lambda^2 e^{-\lambda y_2}$  en  $[(y_1, y_2) : 0 < y_1 < y_2]$ .

Para el caso general en  $\mathbb{R}^k$  por argumentos similares, suponemos  $0 \leq s_1 < t_1 < s_2 < t_2 < \dots < s_k < t_k$ ; y obtenemos que el vector  $(S_1, S_2, \dots, S_k)$  tiene densidad conjunta  $f(y_1, y_2, \dots, y_k) = \lambda^k e^{-\lambda y_k}$  en el conjunto  $[(y_1, y_2, \dots, y_k) : 0 < y_1 < \dots < y_k]$ .

Sea  $g(y_1, y_2, \dots, y_k) = (x_1, x_2, \dots, x_k)$ ; definida por  $x_i = y_i - y_{i-1}$ ;  $i = 1, 2, \dots, k$ , por lo que, el vector  $(T_1, T_2, \dots, T_k) = h(S_1, S_2, \dots, S_k)$ , (de acuerdo al teorema de cambio de variable, ver página 225 de Billingsley[2]), tiene densidad

$$\begin{aligned} h(x_1, x_2, \dots, x_k) &= f(g^{-1}(x_1, x_2, \dots, x_k)) |J| = f\left(x_1, x_1 + x_2, \dots, \sum_{i=1}^k x_i\right) |J| \\ &= \lambda^k e^{-\lambda \sum_{i=1}^k x_i} = \prod_{i=1}^k \lambda e^{-\lambda x_i}; \end{aligned}$$

donde  $|J|$  es el Jacobiano de  $g^{-1}(x_1, x_2, \dots, x_k) = \left(x_1, x_1 + x_2, \dots, \sum_{i=1}^k x_i\right)$  que es igual a 1.

Por lo tanto,  $(T_1, T_2, \dots, T_k)$  son variables aleatorias independientes con densidad exponencial con parámetro  $\lambda$ .

**Condición 2<sup>o</sup> implica Condición 1<sup>o</sup>** (Para mayor referencia ver pág. 262 de Billingsley[2]).

Si  $\{T_i\}_{i=1}^n$  son variables aleatorias independientes e idénticamente distribuidas con densidad  $\exp(\lambda)$ , entonces  $S_n = \sum_{i=1}^n T_i$  tiene densidad *Gama* con parámetros  $n$  y  $\lambda$ . Por lo tanto,  $S_n$  tiene la función de distribución dada por:

$$G_n(x) := P[S_n \leq x] = 1 - e^{-\lambda x} \sum_{i=0}^{n-1} \frac{(\lambda x)^i}{i!}; \quad \text{para } x \geq 0.$$

De (1.6) tenemos,

$$P[N(t) \geq n] = P[S_n \leq t] = 1 - e^{-\lambda t} \sum_{i=0}^{n-1} \frac{(\lambda t)^i}{i!} = \sum_{i=n}^{\infty} e^{-\lambda t} \frac{(\lambda t)^i}{i!},$$

y por (1.7),

$$\begin{aligned} P[N(t) = n] &= P[S_n \leq t < S_{n+1}] = P[S_n \leq t] - P[S_{n+1} \leq t] \\ &= \sum_{i=n}^{\infty} e^{-\lambda t} \frac{(\lambda t)^i}{i!} - \sum_{i=n+1}^{\infty} e^{-\lambda t} \frac{(\lambda t)^i}{i!} \\ &= e^{-\lambda t} \frac{(\lambda t)^n}{n!}, \quad n = 0, 1, 2, \dots; \end{aligned}$$

de esta manera,  $N(t)$  tiene distribución de Poisson con media  $\lambda t$ .

Ahora verificaremos que el proceso  $\{N(t)\}_{t \geq 0}$  cumple con (ii) de la Condición 1ª, esto es, que  $N((t_i, t_{i+1} ])$ ,  $i = 0, 1, 2, \dots, n-1$  son variables aleatorias independientes. Sea  $t$  fijo, consideremos los eventos que ocurren después del tiempo  $t$ . Por las expresiones (1.5) y (1.7) tenemos

$$S_{N(t)} \leq t < S_{N(t)+1}.$$

Definimos la sucesión  $\{T_i^{(t)}\}_{i \geq 0}$ ,  $i \in \mathbb{N}$ ;

$$T_1^{(t)} = S_{N(t)+1} - t, \quad T_2^{(t)} = T_{N(t)+2}, \quad T_3^{(t)} = T_{N(t)+3}, \dots; \quad (1.10)$$

donde  $T_1^{(t)}$  representa el tiempo de espera para el primer evento después de  $t$ ;  $T_2^{(t)}$  es el tiempo de espera entre el primero y segundo evento después de  $t$ , y así sucesivamente. De esta manera la sucesión (1.10) define los tiempos de espera después de  $t$ .

Si  $N(t+s) - N(t) \geq m$ , tenemos  $N(t+s) \geq N(t) + m$  y entonces de (1.6) se tiene que la desigualdad anterior ocurre si y sólo si,

$$S_{N(t)+m} \leq t + s.$$

Luego por (1.10) tenemos

$$\underbrace{S_{N(t)+1} - t}_{T_1^{(t)}} + \underbrace{S_{N(t)+2} - S_{N(t)+1}}_{T_2^{(t)}} + \underbrace{S_{N(t)+3} - S_{N(t)+2}}_{T_3^{(t)}} + \dots + \underbrace{S_{N(t)+m} - S_{N(t)+(m-1)}}_{T_m^{(t)}} \leq s,$$

de esta manera obtenemos que:

$$N(t+s) - N(t) = \max[m : T_1^{(t)} + T_2^{(t)} + \dots + T_m^{(t)} \leq s]. \quad (1.11)$$

De aquí

$$[N(t+s) - N(t) = m] = [T_1^{(t)} + T_2^{(t)} + \dots + T_m^{(t)} \leq s < T_1^{(t)} + T_2^{(t)} + \dots + T_{m+1}^{(t)}].$$

Una comparación entre las expresiones (1.11) y (1.5), muestra que para  $t$  fijo las variables aleatorias  $N(t+s) - N(t)$  para  $s \geq 0$ , están definidas en términos de la sucesión (1.10), exactamente de la misma forma como las variables  $N(s)$  están definidas en términos de la sucesión original de tiempos de espera  $\{T_i\}_{i \geq 1}$ .

Sea  $n \in \mathbb{N}_0$ , y  $y > 0$ . De las expresiones (1.6) y (1.7) tenemos que

$$\begin{aligned} P[S_n \leq t < S_{n+1}, S_{n+1} - t > y] &= P[S_n \leq t, T_{n+1} > y + t - S_n] \\ &= E[I_{[S_n \leq t, T_{n+1} > y + t - S_n]}] = E[I_{[S_n \leq t]} I_{[T_{n+1} > y + t - S_n]}] \\ &= \int_{\tau} \leq t P[T_{n+1} > y + t - \tau] G_n(d\tau) = \int_{\tau} \leq t e^{-\lambda(y+t-\tau)} G_n(d\tau) \\ &= e^{-\lambda y} \int_{\tau} \leq t e^{-\lambda(t-\tau)} G_n(d\tau) = e^{-\lambda y} \int_{\tau \leq t} P[T_{n+1} > t - \tau] G_n(d\tau) \\ &= e^{-\lambda y} P[S_n \leq t, T_{n+1} > t - S_n] = e^{-\lambda y} P[S_n \leq t < S_{n+1}]; \end{aligned}$$

por lo que

$$P[S_n \leq t < S_{n+1}, S_{n+1} - t > y] = e^{-\lambda y} \underbrace{P[S_n \leq t < S_{n+1}]}_{[N(t)=n]}. \quad (1.12)$$

Además, por la independencia y la distribución exponencial de  $\{T_i\}_{i \geq 1}$ , tenemos que,

$$\begin{aligned} P[S_{n+1} - t > y_1, S_n \leq t < S_{n+1}, T_{n+2} > y_2, \dots, T_{n+j} > y_j] \\ &= P[S_{n+1} - t > y_1, S_n \leq t < S_{n+1}] \cdot e^{-\lambda y_2} \dots e^{-\lambda y_j} \\ &= P[N(t) = n] e^{-\lambda y_1} e^{-\lambda y_2} \dots e^{-\lambda y_j}. \end{aligned}$$

Donde la última igualdad se sigue de (1.12).

Si  $H = (y_1, \infty) \times (y_2, \infty) \times \dots \times (y_j, \infty)$ , entonces

$$\begin{aligned} P[N(t) = n, (T_1^{(t)}, T_2^{(t)}, \dots, T_j^{(t)}) \in H] \\ &= P[S_n \leq t < S_{n+1}, S_{n+1} - t > y_1, \dots, T_{n+j} > y_j] \\ &= P[S_n \leq t < S_{n+1}] e^{-\lambda y_1} e^{-\lambda y_2} \dots e^{-\lambda y_j} \end{aligned}$$

$$P[N(t) = n, (T_1^{(t)}, T_2^{(t)}, \dots, T_j^{(t)}) \in H] = P[N(t) = n] P[(T_1, T_2, \dots, T_j) \in H]. \quad (1.13)$$

Por el teorema de extensión, (1.13) se extiende a todo boreliano en  $\mathbb{R}^j$ , ver pág. 135 de Billingsley [2].

Ahora, el evento  $[N(s_i) = m_i, 1 \leq i \leq u]$  lo podemos escribir de la forma  $[(T_1, T_2, \dots, T_j) \in \hat{H}]$ , donde  $j = m_u + 1$  y  $\hat{H}$  es el conjunto de  $x = (x_1, x_2, \dots, x_j) \in \mathbb{R}^j$  para los cuales:

$$x_1 + x_2 + \dots + x_{m_i} \leq s_i \leq x_1 + x_2 + \dots + x_{m_i+1}, \quad 1 \leq i \leq u.$$

De esta manera:

$$\begin{aligned} [N(s_i) = m_i, 1 \leq i \leq u] &= [S_{m_i} \leq s_i < S_{m_i+1}; 1 \leq i \leq u] \\ &= [T_1 + T_2 + \dots + T_{m_i} \leq s_i \leq T_1 + T_2 + \dots + T_{m_i+1}, 1 \leq i \leq u] \\ &= [(T_1, T_2, \dots, T_j) \in \hat{H}]. \end{aligned}$$

Además se cumple

$$\begin{aligned} [(T_1^{(t)}, T_2^{(t)}, \dots, T_j^{(t)}) \in \hat{H}] &= [T_1^{(t)} + T_2^{(t)} + \dots + T_{m_i}^{(t)} \leq s_i < T_1^{(t)} + T_2^{(t)} + \dots + T_{m_i+1}^{(t)}] \\ &= [N(t + s_i) - N(t) = m_i, \quad 1 \leq i \leq u], \end{aligned}$$

por lo que

$$\begin{aligned} P[N(t) = n, N(t + s_i) - N(t) = m_i, \quad 1 \leq i \leq u] \\ &= P[N(t) = n, (T_1^{(t)}, T_2^{(t)}, \dots, T_j^{(t)}) \in \hat{H}] \\ &= P[N(t) = n]P[(T_1, T_2, \dots, T_j) \in H] \\ &= P[N(t) = n]P[N(s_i) = m_i, \quad 1 \leq i \leq u]; \end{aligned}$$

la segunda igualdad se sigue de (1.13). Por lo tanto se cumple:

$$\begin{aligned} P[N(t) = n, N(t + s_i) - N(t) = m_i, \quad 1 \leq i \leq u] \\ &= P[N(t) = n]P[N(s_i) = m_i, \quad 1 \leq i \leq u]. \end{aligned}$$

De esto se sigue, por inducción sobre  $k$ , que si  $0 = t_0 < t_1 < \dots < t_k$ , entonces

$$P[N(t_i) - N(t_{i-1}) = n_i, 1 \leq i \leq k] = \prod_{i=1}^k P[N(t_i - t_{i-1}) = n_i].$$

Por lo tanto los incrementos son independientes, es decir, se cumple (ii) de la Condición 1<sup>o</sup>. ■

Enseguida estudiaremos otras dos caracterizaciones del proceso de Poisson. Para mayor referencia al lector, puede consultar Billingsley[2], Davenport[8], Parzen[18] y Resnick[22].

**Observación 1.9** Una función  $f : \mathbb{R} \rightarrow \mathbb{R}$  es una función  $o(h)$  cuando  $h \rightarrow 0$ , en cuyo caso escribiremos  $f \in o(h)$ , si

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

Si  $f(h) = e^{\lambda h}$  para alguna constante  $\lambda$ , entonces  $f(h) = 1 + \lambda t + o(h)$ .

Las siguientes caracterizaciones del proceso de Poisson son esencialmente cualitativas.

**Condición 3<sup>o</sup>**

(i) Para  $0 < t_1 < \dots < t_k$ , los incrementos

$$N(t_1), N(t_2) - N(t_1), \dots, N(t_k) - N(t_{k-1});$$

son independendientes.

(ii) La distribución de  $N(t) - N(s)$  depende solamente de la diferencia  $t - s$  y  $N(0) = 0$  con probabilidad 1.

**Condición 4°.** Si  $0 < t_1 < \dots < t_k$ , y si  $n_1, n_2, \dots, n_k$  son enteros no negativos entonces

$$P[N(t_k + h) - N(t_k) = 1 | N(t_j) = n_j, j \leq k] = \lambda h + o(h); \quad (1.14)$$

y

$$P[N(t_k + h) - N(t_k) \geq 2 | N(t_j) = n_j, j \leq k] = o(h); \quad (1.15)$$

cuando  $h \downarrow 0$ . Además

$$\lim_{s \rightarrow t} P[N(s) \neq N(t)] = 0, \quad (1.16)$$

y  $N(0) = 0$  con probabilidad 1.

**Observación 1.10** Notemos que de (1.14) y (1.15) tenemos:

$$\begin{aligned} P[N(t_k + h) - N(t_k) = 0 | N(t_j) = n_j, j \leq k] \\ = 1 - P[N(t_k + h) - N(t_k) \geq 1 | N(t_j) = n_j, j \leq k]; \\ = 1 - \lambda h + o(h); \end{aligned}$$

sin pérdida de generalidades la suma de funciones  $o(h)$  es nuevamente  $o(h)$ , luego la probabilidad de que no ocurran eventos en cualquier intervalo de tiempo de longitud  $h$  se aproxima a 1 cuando la longitud del intervalo  $h$  se aproxima a cero.

**Teorema 1.11** Las Condiciones 1° hasta la 4° son equivalentes en presencia de la Condición 0°.

Solamente la Condición 3° no involucra una  $\lambda$ ; esto implica que 1°, 2° y 4° se cumplen para algún  $\lambda > 0$ . La Condición 3° en algunos casos hace que el modelo de Poisson sea bastante creíble; los incrementos serán esencialmente independientes,  $N(t_i + h) - N(t_i)$  depende de  $h$  si las condiciones no cambian en el tiempo.

La Condición 4° es "más cualitativa" que la 3°, ya que ésta involucra solamente la distribución condicional de  $N(t_k + h) - N(t_k)$  para valores pequeños de  $h$ . En el Teorema 1.8 se demostró  $1^\circ \iff 2^\circ$ , entonces para la demostración del Teorema 1.11 es suficiente verificar que:  $1^\circ \rightarrow 3^\circ \rightarrow 4^\circ \rightarrow 1^\circ$ . La primera de ellas  $1^\circ \rightarrow 3^\circ$  es obvia, por lo que resta demostrar las implicaciones  $3^\circ \rightarrow 4^\circ$  y  $4^\circ \rightarrow 1^\circ$ .

**Demostración.** De  $3^\circ \rightarrow 4^\circ$ : De la hipótesis de incrementos independientes y de que la distribución de  $N(t + h) - N(t)$  es igual a la de  $N(h)$ , se sigue que para demostrar (1.14), (1.15) y (1.16) es suficiente verificar

$$P[N(h) = 1] = \lambda h + o(h);$$

y

$$P[N(h) \geq 2] = o(h).$$

Sea

$$U(t) = P[N(t) = 0],$$

por lo que

$$\begin{aligned} U(s+t) &= P[N(s) = 0, N(s+t) - N(s) = 0] \\ &= P[N(s) = 0]P[N(s+t) - N(s) = 0] \\ &= U(s)U(t); \end{aligned}$$

la segunda igualdad se cumple por (i) de la Condición 3<sup>o</sup> de incrementos independientes, entonces tenemos,

$$U(s+t) = U(s)U(t).$$

Por lo tanto  $U(t)$  es una función exponencial, (ver pág. 168 de Billingsley[2]), es decir,

$$U(t) = e^{-\lambda t}; \text{ para algún valor de } \lambda.$$

Además por la Condición 0<sup>o</sup>:  $[S_1 = \infty] = \cap_{n=1}^{\infty} [S_1 > t_n]$  a su vez tenemos  $0 = P[S_1 = \infty] = \cap_{n=1}^{\infty} P[S_1 > t_n] = \lim_{n \rightarrow \infty} P[S_1 > t_n] = 0$ . Por lo tanto  $P[S_1 > t] \rightarrow 0$ , cuando  $t \rightarrow \infty$ . De esta manera  $P[N(t) = 0] = P[T_1 > t] \rightarrow 0$ ; cuando  $t \rightarrow \infty$ , por lo que  $\lambda$  debe ser positivo.

Sea  $\phi(t) = P[N(t) \geq 2]$ , definimos el evento  $A_n$

Entonces para  $\omega$  en el conjunto donde se cumple la Condición 0<sup>o</sup>, existe  $n_0(\omega)$  tal que para  $n > n_0(\omega)$  implica que  $\omega \notin A_n$ ; ya que basta tomar  $n_0(\omega)$  de tal manera que  $1/n_0(\omega) < \min_{i \leq N_1(\omega)} T_i(\omega)$ . Si  $\omega \notin A_n$ ; esto implica que  $\omega \notin A_{n+1}$ ; esto se sigue ya que si  $1/n < \min_{i \leq N_1(\omega)} T_i(\omega)$ , entonces  $1/n + 1 < \min_{i \leq N_1(\omega)} T_i(\omega)$ , es decir  $\omega$  no está a partir de cierto índice en  $A_n$ , entonces  $P[\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_n] = P[\limsup_n A_n] = 0$ , además para cualquier sucesión de eventos por el Lema de Fatou se cumple

$$P\left(\liminf_n A_n\right) \leq \liminf_n P(A_n) \leq \limsup_n P(A_n) \leq P\left(\limsup_n A_n\right),$$

Por lo que

$$\liminf_n P(A_n) = \limsup_n P(A_n) = 0.$$

Por lo tanto,

$$\lim_{n \rightarrow \infty} P[A_n] = 0.$$

Luego (1.17) implica que

$$\begin{aligned} 1 - P(A_n) &= (1 - \phi(n^{-1}))^n \\ \log(1 - P(A_n)) &= n \log(1 - \phi(n^{-1})); \end{aligned}$$

si  $x = -\phi(n^{-1})$ , entonces de la desigualdad  $\log(1+x) \leq x$ , se sigue que

$$\log(1 - P(A_n)) \leq -n\phi(n^{-1});$$

luego

$$n\phi(n^{-1}) \leq -\log(1 - P[A_n]);$$

como  $P[A_n] \rightarrow 0$ , entonces

$$0 \leq n\phi(n^{-1}) \leq -\log(1 - P[A_n]) \rightarrow 0;$$

por lo que

$$n\phi(n^{-1}) \xrightarrow{n \rightarrow \infty} 0.$$

Se sigue

$$\frac{\phi(h)}{h} \xrightarrow{h \rightarrow 0} 0,$$

cuando  $h$  es de la forma  $h = 1/n$ . Entonces

$$\phi(h) = o(h); \quad \text{para } h = \frac{1}{n}.$$

Ahora lo demostraremos para toda  $h$ . Notemos que  $\phi(h)$  es no decreciente ya que si  $t_1 < t_2$  tenemos

$$P[N(t_1) \geq 2] \leq P[N(t_2) \geq 2];$$

entonces  $\phi(t_1) \leq \phi(t_2)$ . Además notemos que  $[1/h] < 1/h$  (donde  $[x]$  significa "la parte entera de  $x$ "), y  $1/h \leq 1 + [1/h]$ ; y como  $\phi(h)/h \leq 1/h$  entonces por las desigualdades anteriores y por ser  $\phi(h)$  no decreciente se cumple:

$$\phi(h) \leq \phi\left(\frac{1}{[h^{-1}]}\right);$$

por lo que

$$\frac{\phi(h)}{h} \leq \phi\left(\frac{1}{[h^{-1}]}\right) \left(\left[\frac{1}{n}\right] + 1\right) \xrightarrow{h \rightarrow 0} 0.$$

Por lo tanto,

$$\phi(h) = P[N(h) \geq 2] = o(h).$$

Por otra parte

$$\begin{aligned} P[N(h) = 1] &= 1 - P[N(h) = 0] - P[N(h) \geq 2] \\ &= 1 - e^{-\lambda h} + o(h) = \lambda h + o(h). \end{aligned}$$

De esta manera se completa la demostración.

**Demostración de  $4^\circ \rightarrow 1^\circ$ .**

Para  $k$  fija, sea el evento

$$A = [N(t_j) = n_j; \quad j \leq k],$$

y para  $t \geq 0$ ;

$$P_n(t) = P[N(t_k + t) - N(t_k) = n | A].$$

Por demostrar que

$$P_n(t) = e^{-\lambda t} \frac{(\lambda t)^n}{n!}; \quad n = 0, 1, 2, \dots \quad (1.18)$$

Para  $B$  y  $C$  eventos arbitrarios, se cumple  $|P(B) - P(C)| \leq P(B \Delta C)$ .

Por lo tanto:

$$\begin{aligned} & |P_n(t) - P_n(s)| \\ &= \left| \frac{P[N(t_k + t) - N(t_k) = n \cap A]}{P(A)} - \frac{P[N(t_k + s) - N(t_k) = n \cap A]}{P(A)} \right| \\ &\leq P^{-1}(A)P[N(t_k + s) \neq N(t_k + t)]. \end{aligned}$$

Por (1.16) se cumple

$$\lim_{t \rightarrow s} |P_n(t) - P_n(s)| = 0,$$

por lo tanto,  $P_n(t)$  es continua en  $t$ .

Sea  $D_t = N(t_k + t) - N(t_k)$ .

Si  $D_{t+h} = n$ , entonces  $D_t = m$  para algún  $m \leq n$ . Por las reglas de la probabilidad condicional

$$\begin{aligned} P_n(t+h) &= P[N(t_k + t + h) - N(t_k) = n | A] = \frac{P[N(t_k + t + h) - N(t_k) = n \cap A]}{P(A)} \\ &= \frac{P[D_{t+h} = n \cap A]}{P(A)} = P^{-1}(A) \sum_{m=0}^n P[D_{t+h} = n \cap A \cap (D_t = m)] \\ &= P^{-1}(A) \{ P[D_{t+h} - D_t = 0 \cap A \cap (D_t = n)] \\ &\quad + P[(D_{t+h} - D_t = 1) \cap A \cap (D_t = n - 1)] \\ &\quad + \sum_{m=0}^{n-2} P[(D_{t+h} - D_t = n - m) \cap A \cap (D_t = m)] \}; \end{aligned}$$

luego

$$\begin{aligned} P_n(t+h) &= P[D_{t+h} - D_t = 0 | A \cap (D_t = n)] \underbrace{P[A \cap (D_t = n)]}_{P^{-1}(A)} + \\ &\quad P[D_{t+h} - D_t = 1 | A \cap (D_t = n - 1)] \underbrace{P[A \cap (D_t = n - 1)]}_{P^{-1}(A)} + \\ &\quad \sum_{m=0}^{n-2} P[D_{t+h} - D_t = n - m | A \cap (D_t = m)] \underbrace{P[A \cap (D_t = m)]}_{P^{-1}(A)}. \end{aligned}$$

Pero los términos señalados son  $P_n(t)$ ,  $P_{n-1}(t)$  y  $P_m(t)$  respectivamente, entonces

$$\begin{aligned} P_n(t+h) &= P_n(t)P[D_{t+h} - D_t = 0 | A \cap (D_t = n)] + \\ &\quad P_{n-1}(t)P[D_{t+h} - D_t = 1 | A \cap (D_t = n - 1)] + \\ &\quad P_m(t) \sum_{m=0}^{n-2} P[D_{t+h} - D_t = n - m | A \cap (D_t = m)]. \end{aligned}$$

Por (1.15) el tercer sumando es una función  $o(h)$  para cada  $n$  fijo. Aplicando (1.14) y (1.15) nos queda

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

reordenando y dividiendo por  $h$  tenemos,

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

haciendo  $h \downarrow 0$ , nos da

$$P'_n(t) = -\lambda P_n(t) + \lambda P_{n-1}(t); \quad n \geq 1. \quad (1.19)$$

Para el caso en que  $n = 0$ ;  $P_{-1}(t)$  es considerado idénticamente 0, luego

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

dividiendo por  $P_0(t)$ ,

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

integrando tenemos

$$\log P_0(t) = -\lambda t + C;$$

esto implica

$$P_0(t) = e^{-\lambda t + C} = e^{-\lambda t} e^C;$$

tomando  $k = e^C$  nos queda,

$$P_0(t) = k e^{-\lambda t}.$$

Además

$$P_0(0) = P[N(0) = 0] = k e^0 = 1;$$

se sigue que  $k = 1$ , por lo tanto

$$P_0(t) = e^{-\lambda t}. \quad (1.20)$$

Análogamente para  $n > 0$ , de (1.19) tenemos

$$e^{\lambda t} (P'_n(t) + \lambda P_n(t)) = \lambda e^{\lambda t} P_{n-1}(t);$$

luego,

$$\frac{d}{dt} (e^{\lambda t} P_n(t)) = \lambda e^{\lambda t} P_{n-1}(t); \quad (1.21)$$

Integrando (1.21) tenemos la solución de (1.19) dada por:

$$P_n(t) = e^{-\lambda t} \left[ P_n(0) + \lambda \int_0^t P_{n-1}(s) e^{\lambda s} ds \right]. \quad (1.22)$$

Ya que

$$P_n(0) = P[D_0 = n|A] = \begin{cases} 1 & \text{si } n = 0 \\ 0 & \text{si } n \neq 0 \end{cases}$$

la demostración de  $P_n(t) = e^{-\lambda t} \frac{(\lambda t)^n}{n!}$ ;  $n = 0, 1, 2, \dots$  se verificará por inducción.

Para  $n = 1$  de (1.20) y (1.22) se tiene que,

$$P_1(t) = e^{-\lambda t} \left[ P_1(0) + \lambda \int_0^t P_0(s) e^{\lambda s} ds \right] = e^{-\lambda t} [\lambda t];$$

por lo que

$$P_1(t) = \lambda t e^{-\lambda t}.$$

Suponemos que se cumple para  $n - 1$ , es decir

$$P_{n-1}(t) = \frac{e^{-\lambda t} (\lambda t)^{n-1}}{(n-1)!};$$

sustituimos  $P_{n-1}(t)$  en (1.22),

$$\begin{aligned} P_n(t) &= e^{-\lambda t} \left[ P_n(0) + \lambda \int_0^t \frac{e^{-\lambda s} (\lambda s)^{n-1}}{(n-1)!} e^{\lambda s} ds \right] = e^{-\lambda t} \left[ P_n(0) + \frac{\lambda^n}{(n-1)!} \int_0^t s^{n-1} ds \right] \\ &= e^{-\lambda t} \left[ P_n(0) + \frac{\lambda^n}{(n-1)!} \cdot \frac{t^n}{n} \right]; \end{aligned}$$

de esta manera tenemos

$$P_n(t) = \frac{e^{-\lambda t} (\lambda t)^n}{n!}, \quad n = 0, 1, 2, \dots$$

Por lo tanto,  $P_n(t)$  tiene una distribución de Poisson con media  $\lambda t$ . ■

### 1.3 Proceso de Poisson no homogéneo

Una generalización del proceso de Poisson es permitir que el parámetro  $\lambda$  sea una función de  $t$ . Esto es, en muchas aplicaciones se consideran parámetros que varían con el tiempo en forma no homogénea, tales procesos son denominados procesos de Poisson no homogéneos o no estacionarios; para una mayor referencia ver Karlin and Howard[13] y Resnick[22].

En la Definición 1.4 de proceso de Poisson, consideremos el espacio de estados  $E = [0, \infty)$ ,  $\xi$  es la  $\sigma$ -álgebra generada por los intervalos de la forma  $(a, b]$  de  $E$ ,  $\mu$  una medida en  $(E, \xi)$ , tal que es finita para conjuntos en  $\xi$ ,

$$\mu(\cdot) : \xi \rightarrow \mathbb{R},$$

y absolutamente continua con densidad  $\lambda(t)$ ; esto es

$$\mu(t) = \int_0^t \lambda(t) dt,$$

donde  $\lambda(t)$  es la función de intensidad del proceso.

**Definición 1.8** *Un proceso de conteo  $\{N(t)\}_{t \geq 0}$  es un proceso de Poisson no homogéneo con función de intensidad  $\lambda(t)$  si*

(i)  $P[N(0) = 0] = 1$ .

(ii) Para  $0 \leq t_0 < t_1 < t_2 < \dots < t_k < \dots < t_n$  tenemos que  $N((t_i, t_{i+1}])$ ,  $i = 1, 2, \dots, n$  son variables aleatorias independientes.

(iii) Para cada  $t \geq 0$

$$P[N((0, t]) = n] = e^{-\{\mu(t)\}} \frac{(\mu(t))^n}{n!}; \quad \text{para } n \geq 0;$$

$$\text{donde } \mu(t) = \int_0^t \lambda(u) du.$$

**Observación 1.12** Si  $N(t)$  es un proceso de Poisson no homogéneo en la recta con función de intensidad  $\lambda(t)$ , entonces un incremento  $N(t) - N(s)$ , nos proporciona el número de eventos en un intervalo  $(s, t]$ , y los incrementos sobre intervalos ajenos son variables aleatorias independientes.

Este proceso puede ser obtenido como una transformación del proceso de Poisson homogéneo de la siguiente manera.

**Definición 1.9** (i) Se define la función  $m(t)$  como:

$$m(t) = \mu[0, t] = \int_0^t \lambda(s) ds;$$

(ii) Se define la función inversa por:

$$\bar{m}(t) := \inf\{u \in E : m(u) \geq t\}; \quad t \geq 0.$$

Si suponemos que  $m(\infty) = \infty$ , entonces el conjunto  $\{u : m(u) \geq t\}$  es no vacío para toda  $t$ , así

$$\bar{m}(t) : [0, \infty) \rightarrow [0, \infty),$$

como  $m$  es continua, tenemos que  $\bar{m}$  es estrictamente creciente. Si  $N = \sum_{n \geq 1} \epsilon_{S_n}$  es un proceso de Poisson homogéneo con media igual a uno, entonces considerando en la Proposición 1.4 la transformación  $T = \bar{m}$  se sigue que

$$N' := \sum_{n \geq 1} \epsilon_{\bar{m}(S_n)};$$

también es un proceso de Poisson con media dada por:

$$\begin{aligned} \mu'([0, t]) &= \mu(\bar{m}^{-1}(0, t]) = |\{x : \bar{m}(x) \leq t\}| \\ &= |\{x : x \leq m(t)\}| = |[0, m(t)]| = m(t) = \mu[0, t]; \end{aligned}$$

donde  $|[0, t]|$  representa la medida de Lebesgue.

De esta forma hemos obtenido un proceso de Poisson no homogéneo con media  $\mu(t) = \mu \int_0^t \lambda(s) ds$  a partir de un proceso de Poisson homogéneo.

**Observación 1.13** Lo anterior significa que, si  $\sum_{n=1}^{\infty} \epsilon_{s'_n}$  es un proceso de Poisson con media  $\mu$  en  $[0, \infty)$  absolutamente continua con densidad  $\lambda(s)$ , donde  $0 < s'_1 < s'_2 < \dots$ , es la sucesión de puntos asociado al proceso, entonces se cumple que la distribución del proceso de puntos  $\{S'_n\}_{n \geq 1}$  es la misma para el proceso de puntos  $\{\bar{m}(S_n)\}_{n \geq 1}$

$$\{S'_n\}_{n \geq 1} \stackrel{d}{=} \{\bar{m}(S_n), n \geq 1\};$$

donde  $\stackrel{d}{=}$  significa igualdad en distribución. Así que en  $R^\infty$ ,

$$\{m(S'_n)\}_{n \geq 1} \stackrel{d}{=} \{S_n\}_{n \geq 1}.$$

De aquí que,  $\sum_{n \geq 1} \epsilon_{m(S'_n)}$  y  $\sum_n \epsilon_{S_n}$  tienen la misma distribución, por lo tanto  $\sum_n \epsilon_{m(S'_n)}$  es un proceso de Poisson homogéneo con  $\lambda = 1$ .

**Observación 1.14** Al igual que en el caso homogéneo, las Condiciones 3<sup>o</sup> y 4<sup>o</sup> se cumplen para un proceso de Poisson no homogéneo. Esto es, podemos expresar un proceso de Poisson no homogéneo en términos de funciones  $o(h)$ .

## 1.4 Comentarios finales

En este capítulo se ha presentado el proceso de Poisson, tanto en el espacio como en la recta real. A menudo son más realistas los modelos basados en procesos de Poisson no homogéneos, en los que la tasa de ocurrencia de un evento es una función del tiempo. De acuerdo a este capítulo una forma alternativa de estudiar el comportamiento de las rachas de calor es estudiar el tiempo que transcurre entre dos rachas de calor consecutivas. Por lo que en el caso de estudio verificaremos, en primera instancia, si los tiempos entre rachas de calor consecutivas  $\{T_n\}_{n \geq 1}$ , siguen una distribución exponencial.

A la luz de los datos, una vez determinada la forma funcional de la función de distribución que sigue la variable aleatoria objeto de estudio, contamos con un conjunto de funciones de probabilidad que depende de parámetros desconocidos. Suponiendo que los datos han sido obtenidos de uno de los modelos de la familia, ¿qué puede concluirse sobre los valores de los parámetros desconocidos o acerca de los datos del mismo modelo? ¿son los datos consistentes con la familia de modelos de probabilidad? Esta situación se presenta con frecuencia debido a que es posible, a menudo, conocer la forma funcional de la distribución de probabilidad por consideraciones teóricas, quedando únicamente indeterminados los parámetros que determinan la función de distribución, jugando un papel fundamental la "Teoría de la Probabilidad" en las distintas formas funcionales de las distribuciones de probabilidad.

La diferencia entre un modelo probabilístico y un modelo estadístico es que en el primero la probabilidad está totalmente especificada, mientras que en el segundo la distribución de probabilidad es desconocida. Todo lo que se conoce son algunas características de esta distribución de probabilidad. El quehacer de la estadística es inferir un elemento de esa familia de modelos de probabilidad, que en la mayoría de

los casos es conocida excepto por un número limitado de parámetros desconocidos. Son precisamente la función de intensidad y los parámetros involucrados en ella sobre los cuales realizaremos inferencia. Estableceremos un rango de predicción sobre la función de intensidad para el comportamiento de futuras rachas de calor en la región, y estimaremos los parámetros desconocidos mediante el enfoque de verosimilitud, validando las propiedades frecuentistas del método a través de las probabilidades de cobertura real de los intervalos de verosimilitud, mismos que resultarán una aproximación a los intervalos de confianza seleccionados en el estudio del fenómeno aleatorio "*Rachas de calor*". Éste es precisamente el propósito del siguiente capítulo: realizar inferencia estadística en procesos de Poisson.



## Capítulo 2

### Inferencia estadística en procesos de Poisson

La inferencia estadística, en base a un conjunto de observaciones de una variable aleatoria de una población es un planteamiento inductivo que infiere acerca de la función de distribución que aceptablemente se ajuste a las observaciones y del parámetro de probabilidad de ocurrencia de la muestra.

En el presente capítulo describiremos una manera de realizar inferencia estadística sobre los parámetros de un proceso de Poisson basándonos en la teoría de verosimilitud y encaminar el propósito de este trabajo de tesis. En primer instancia determinaremos el estimador o estimadores de máxima verosimilitud en un proceso de Poisson, con la finalidad de realizar inferencia sobre los mismos, calculando, graficando e interpretando los intervalos de verosimilitud así como los de confianza, de cada uno de los parámetros, e intervalos que cubrirán el valor del parámetro con una cierta probabilidad. El siguiente compendio de resultados está basado en las referencias: Basawa[1], Canavos[4], Evans & Rosenthal[9], Gómez[10], Kalbfleisch[11], Pranab[19] y Sprott[25].

#### 2.1 Función de verosimilitud

La estructura inferencial de un modelo se especifica por los siguientes tres elementos: el espacio muestral, conjunto de todas las muestras aleatorias posibles  $X = \{X_1, X_2, \dots, X_n\}$ , el espacio parametral  $\Theta = \{\theta_1, \theta_2, \dots, \theta_k\}$  que es el conjunto de posibles valores que puede tomar el vector de parámetros y la función de probabilidad  $P(X = x; \theta)$  que depende del vector de parámetros desconocidos. Este es el modelo estadístico usual de un experimento; el conjunto de funciones de probabilidad con parámetros definidos en un espacio parametral,

$$\mathcal{F} = \{P(\cdot; \theta) : \theta \in \Theta \subseteq \mathbb{R}^k\}.$$

Una vez observada la muestra, el problema es inferir el elemento de la familia de distribuciones de probabilidad para  $X$ . Si tenemos una o varias variables aleatorias  $X \sim f(\cdot; \theta_0)$ , donde  $\theta_0$  es el valor verdadero; la estadística infiere sobre ese valor verdadero  $\theta_0$ . Pero ¿Cómo se debe proceder para estimar el valor de  $\theta_0$  con base en la muestra aleatoria? Un posible método es seleccionar el valor de  $\theta$  para el cual la probabilidad del valor observado es máxima; dicho método requiere de un concepto, el de función de verosimilitud. Una vez justificados los medios, es en este concepto que centraremos el desarrollo de este trabajo de tesis aplicado a procesos de Poisson

tanto homogéneos como no homogéneos respectivamente, con la finalidad de producir inferencia sobre los parámetros desconocidos estimados por máxima verosimilitud.

Considérese una muestra de variables aleatorias independientes e idénticamente distribuidas  $X = (X_1, X_2, \dots, X_n)$  con función de probabilidad  $P(X = x; \theta)$ , que depende de un número finito de parámetros reales desconocidos  $\theta = (\theta_1, \theta_2, \dots, \theta_k) \in \Theta$ , donde  $\theta$  puede ser un vector de parámetros.

**Definición 2.1** La *función de verosimilitud* de una muestra  $\underline{x} = (x_1, x_2, \dots, x_n)$ , denotada por  $L(\theta; \underline{x})$ , se define como una función del parámetro  $\theta$  que es proporcional a la probabilidad de la muestra observada  $P(X = \underline{x}; \theta)$ . La función  $L : \Theta \rightarrow \mathbb{R}^+ \cup \{0\}$  está dada por:

$$L(\theta; \underline{x}) \propto P(X = \underline{x}; \theta). \tag{2.1}$$

La función de verosimilitud tiene un rol muy importante dentro de la inferencia estadística, su principal objetivo es la de inferir acerca de los parámetros involucrados en la distribución que haya sido seleccionada con la finalidad de describir mejor el fenómeno aleatorio de interés a partir de la muestra observada. Es aquí donde hacemos hincapié del hecho de que ésto toma relevancia cuando el experimento ya ha sido ejecutado, esto es, después de que las variables aleatorias ya han sido observadas.

### Aproximación continua de la función de verosimilitud

En la realidad todo instrumento de medición tiene precisión finita y sólo pueden registrarse mediciones con un número finito de decimales, por lo que los datos observados siempre son discretos. Pero cuando  $X$  es una variable aleatoria continua, la observación  $X_i = x_i$  debe interpretarse como  $X_i \in (x_i - \varepsilon/2, x_i + \varepsilon/2)$ , donde  $\varepsilon$  representa la precisión del instrumento de medición. Entonces para una muestra de variables aleatorias independientes e idénticamente distribuidas  $X = (X_1, X_2, \dots, X_n)$  con función de densidad  $f_X(x; \theta)$ , la función de verosimilitud de  $\theta$  es proporcional a la probabilidad conjunta de la muestra observada,

$$\begin{aligned} L(\theta; \underline{x}) &\propto \prod_{i=1}^n P(x_i - \frac{\varepsilon}{2} \leq X_i \leq x_i + \frac{\varepsilon}{2}; \theta) \\ &= \prod_{i=1}^n \int_{x_i - \frac{\varepsilon}{2}}^{x_i + \frac{\varepsilon}{2}} f_X(x; \theta) dx. \end{aligned} \tag{2.2}$$

Si  $\varepsilon = 0$ , significa que el instrumento de medición tiene precisión infinita, esto es, que las observaciones pueden registrarse con un número infinito de decimales, lo cual es imposible en la realidad. Además para una variable continua se tiene que  $P(X = \underline{x}; \theta) = 0$  para toda  $\underline{x}$  y  $\theta$ , por lo que no es posible entonces definir la función de verosimilitud a partir de dichas probabilidades para este caso. Entonces se supone que la precisión del instrumento de medición es  $\varepsilon > 0$ , y bajo ciertas condiciones la función de densidad  $f(x; \theta)$  podrá usarse como una aproximación a la función de verosimilitud (2.2), referencias Montoya[16], Montoya, Díaz, Sprott[17].

Por el teorema del valor medio para integrales de funciones continuas, para algún punto  $x'$  en el intervalo cerrado  $[x_i - \varepsilon/2, x_i + \varepsilon/2]$ , la  $i$ -ésima integral en (2.2) es

$$[(x_i + \varepsilon/2) - (x_i - \varepsilon/2)]f_X(x'; \theta) = \int_{x_i - \varepsilon/2}^{x_i + \varepsilon/2} f_X(x; \theta) dx. \quad (2.1)$$

Si  $f_X(x; \theta)$  es aproximadamente constante en este intervalo para todo valor plausible de  $\theta$ , entonces  $f_X(x'; \theta) \approx f_X(x; \theta)$  en ese mismo intervalo. Si esta aproximación es adecuada para algunos o todos los  $i \in \{1, 2, \dots, n\}$  y si  $\varepsilon$  no depende de  $\theta$ , entonces las correspondientes probabilidades en (2.2) pueden reemplazarse por la función de densidad evaluada en el valor observado  $x_i$ . Esto es,

$$L(\theta; \underline{x}) \propto \prod_{i=1}^n P(x_i - \frac{\varepsilon}{2} \leq X_i \leq x_i + \frac{\varepsilon}{2}; \theta) \approx \prod_{i=1}^n \varepsilon f_X(x_i; \theta) \propto \prod_{i=1}^n f_X(x_i; \theta).$$

El propósito principal de la aproximación continua es la conveniencia matemática de facilitar los cálculos mediante derivadas e integrales.

Existen diferentes procedimientos para estimar los parámetros de un modelo de probabilidad. De entre esos procedimientos probablemente el más versátil, ya que se puede aplicar en gran cantidad de situaciones, y por ello uno de los más utilizados se conoce con el nombre de “método de máxima verosimilitud”, que selecciona como estimador del valor verdadero a aquél valor del parámetro que maximiza el valor de la probabilidad de la muestra aleatoria observada. Luego, una manera de calcular estos estimadores es derivar la función de verosimilitud con respecto a cada uno de los parámetros a estimar, igualar a cero y despejar el respectivo valor del parámetros. Podemos considerar que el método de máxima verosimilitud tal y como hoy lo conocemos e interpretamos fue propuesto por Fisher en 1912; sin embargo la resolución de los problemas numéricos planteados por este método en la mayor parte de los casos son de tal magnitud que no había sido posible su amplia utilización hasta la llegada de los ordenadores modernos.

### 2.1.1 Función de verosimilitud relativa

Es importante recordar la interpretación de la función de verosimilitud;  $L(\theta; \underline{x})$  es proporcional a la probabilidad de la muestra  $\underline{x}$ , dado que  $\theta$  es el valor verdadero, y no la probabilidad de  $\theta$  dado que hemos observado  $\underline{x}$ . También puede ocurrir que el valor de  $L(\theta; \underline{x})$  sea muy pequeño para cada valor de  $\theta$ , sin embargo, no es el valor concreto de la verosimilitud el que nos informa de la confianza para cada valor de  $\theta$ , sino el valor relativo de la verosimilitud de los diferentes valores posibles del parámetro. Lo interesante entonces, es trabajar con los cocientes de verosimilitudes, esto tiene sentido y se interpreta como una medida de plausibilidad entre dos valores del parámetro basada en la muestra observada  $\underline{x}$ . El cociente  $L(\theta_1; \underline{x})/L(\theta_2; \underline{x}) = k$  significa que el valor  $\theta_1$  es  $k$  veces más plausible que el valor de  $\theta_2$ , en el sentido de que  $\theta_1$  hace  $k$  veces más probable a la muestra observada de lo que la hace  $\theta_2$ . Por esta razón se sugiere usar la función de verosimilitud relativa que definimos a continuación para hacer inferencia sobre los parámetros del modelo probabilístico  $P(X = \underline{x}; \theta)$ .

Para que la función de verosimilitud no involucre una constante arbitraria de plausibilidad, es necesario estandarizarla con respecto a su máximo. A esta función estandarizada se le conoce con el nombre de función de verosimilitud relativa.

**Definición 2.2** La función de verosimilitud relativa  $R(\theta; \underline{x})$  de  $\theta$  está definida como la razón de la función de verosimilitud  $L(\theta; \underline{x})$  y su máximo  $L(\hat{\theta}; \underline{x})$ :

$$R(\theta; \underline{x}) = \frac{L(\theta; \underline{x})}{\max_{\theta \in \Theta} L(\theta; \underline{x})} = \frac{L(\theta; \underline{x})}{L(\hat{\theta}; \underline{x})}. \quad (2.3)$$

Donde  $\hat{\theta} \in \Theta$  es el valor del parámetro en el cual  $L(\theta; \underline{x})$  alcanza el valor máximo en  $\Theta$ , y se le conoce como *estimador de máxima verosimilitud (EMV)* de  $\theta$ . El EMV es el valor más plausible de  $\theta$  en el sentido de que le da a la muestra observada la más alta probabilidad de ocurrencia. Comúnmente no es posible calcular el estimador de máxima verosimilitud de manera analítica. Por ello son necesarios procedimientos numéricos como el método de Newton-Raphson, por ejemplo, para calcular las raíces de una función, referencia Burden[3].

Ya que  $L(\theta; \underline{x}) \propto P(X = \underline{x}; \theta)$ , y  $P(X = \underline{x}; \theta)$ ; es una función de probabilidad necesariamente acotada, el denominador de (2.3) existe y es finito. Se sigue que  $R(\theta; \underline{x})$  es una medida de plausibilidad de un valor específico de  $\theta$  relativo a  $\hat{\theta}$ , basada en la muestra observada  $\underline{x}$ . Note que  $L(\theta; \underline{x}) \leq L(\hat{\theta}; \underline{x})$  para todos los valores posibles de  $\theta$  en el espacio paramétrico, se sigue por tanto que  $0 \leq R(\theta; \underline{x}) \leq 1$ . Valores de  $\theta$  con  $R(\theta; \underline{x})$  cercanos a 1 son muy creíbles o plausibles, mientras que valores cercanos a cero son poco creíbles en base a la muestra observada.

En el siguiente ejemplo ilustraremos el algoritmo a seguir para el cálculo del estimador basados en la teoría de verosimilitud hasta ahora comprendida.

**Ejemplo 2.1** Consideremos una muestra de variables aleatorias  $\underline{x} = (x_1, x_2, \dots, x_n)$  independientes e idénticamente distribuidas con función de densidad  $f(x; \theta) = \frac{1}{\theta} \exp\{-\frac{1}{\theta}x\}$ ;  $\theta > 0$ .

De (2.1), la función de verosimilitud es:

$$L(\theta; \underline{x}) = c(\underline{x}) \prod_{i=1}^n \frac{1}{\theta} \exp\left\{-\frac{1}{\theta}x_i\right\} = c(\underline{x}) \cdot \frac{1}{\theta^n} \exp\left\{-\sum_{i=1}^n \frac{x_i}{\theta}\right\} = c(\underline{x}) \cdot \frac{1}{\theta^n} \exp\left\{-\frac{t}{\theta}\right\};$$

donde  $c(\underline{x})$  es una función constante arbitraria positiva de  $\underline{x}$  que no depende de  $\theta$  y  $t = \sum_{i=1}^n x_i$ .

Puesto que, en general, la función de verosimilitud es compleja y difícil de evaluar, y dado que existe una relación biunívoca entre una función y su logaritmo, preferimos encontrar la derivada del logaritmo de la función de verosimilitud, denotada por:

$$\ell(\theta; \underline{x}) = \log L(\theta; \underline{x}).$$

Aplicando logaritmo a la expresión  $L(\theta; \underline{x})$  obtenemos,

$$\ell(\theta; \underline{x}) = \log L(\theta; \underline{x}) = c'(\underline{x}) - n \log(\theta) - \frac{t}{\theta}.$$

Por facilidad matemática, siempre podemos trabajar con el logaritmo de  $L(\theta; \underline{x})$  ya que la función de verosimilitud es no negativa, con la convención de que  $\ell(\theta; \underline{x}) = -\infty$ , si  $L(\theta; \underline{x}) = 0$ . El objetivo que se pretende con el método de estimación es encontrar aquellos valores de los parámetros que maximicen la probabilidad de ocurrencia de la muestra. Por lo tanto, para encontrar estos estimadores se debe derivar la función de verosimilitud con respecto a cada uno de los parámetros a estimar, igualar a cero y despejar el respectivo valor, esto es, se debe determinar la “función Score”, que no es otra cosa más que la primer derivada de la función log-verosimilitud con respecto al parámetro desconocido. Resolviendo la ecuación de verosimilitud, obtenemos la función Score que denotaremos  $sc(\theta; \underline{x})$ ,

$$sc(\theta; \underline{x}) = \frac{\partial \ell(\theta; \underline{x})}{\partial \theta} = \frac{-n}{\theta} + \frac{t}{\theta^2}$$

$$sc(\theta; \underline{x}) = 0 \implies \frac{-n\theta + t}{\theta^2} = 0.$$

Por lo tanto, un punto crítico está dado por

$$\hat{\theta} = \frac{t}{n} = \bar{x}.$$

Para determinar el *EMV*, determinamos la segunda derivada con respecto a  $\theta$  que es conocida como “información de Fisher”, que denotaremos por  $I(\theta; \underline{x})$ , y verificamos que sea negativa con la finalidad de determinar si el punto crítico es un máximo. La información de Fisher mide la curvatura de la función de verosimilitud relativa o precisión local en una vecindad de su máximo. De aquí la justificación del término “información”. Esta es de particular importancia cuando la forma de la función de verosimilitud es simétrica, o tiene la forma de la normal. Cuando el valor de  $I(\theta; \underline{x})$  es pequeño la curva de la función de verosimilitud es extendida; mientras que cuando el valor de  $I(\theta; \underline{x})$  es grande la curva es angosta. Si en la información de Fisher sustituimos el valor de  $\hat{\theta}$  obtenemos lo que conocemos como “información observada”  $I(\hat{\theta}; \underline{x})$ , de aquí entonces que  $I(\hat{\theta}; \underline{x})$  es la principal característica de la forma de la función de verosimilitud

$$I(\hat{\theta}; \underline{x}) = \left[ -\frac{\partial^2}{\partial \theta^2} \log L(\theta; \underline{x}) \right]_{\theta=\hat{\theta}}.$$

Determinando ahora la información de Fisher tenemos

$$I(\theta; \underline{x}) = -\frac{\partial}{\partial \theta} sc(\theta; \underline{x}) = -\frac{n}{\theta^2} + \frac{n}{\theta^3} \sum_{i=1}^n x_i = \frac{-n\theta + 2\sum_{i=1}^n x_i}{\theta^3}.$$

Su valor en  $\theta = \hat{\theta}$ , nos proporciona la información observada,

$$I(\hat{\theta}; \underline{x}) = -\frac{n}{\hat{\theta}^2} + \frac{2}{\hat{\theta}^3} n\hat{\theta} = -\frac{n}{\hat{\theta}^2} + \frac{2n}{\hat{\theta}^2} = \frac{n}{\hat{\theta}^2}.$$

La función de verosimilitud relativa (2.3) está dada por

$$R(\theta; \underline{x}) = \left( \frac{t}{n\theta} \right)^n e^{-\left(\frac{t}{\theta} + \frac{t}{ni}\right)} = \left( \frac{t}{n\theta} \right)^n e^{-\left(\frac{t}{\theta} + n\right)};$$

pero  $\hat{\theta} = t/n$ , donde  $t = \sum_{i=1}^n x_i$ , luego

$$R(\theta; \underline{x}) = \left(\frac{\hat{\theta}}{\theta}\right)^n e^{-n\frac{\hat{\theta}}{\theta}+n}.$$

### 2.1.2 Intervalos de verosimilitud

Hasta ahora se ha visto cómo determinar el estimador de máxima verosimilitud a partir de la función de verosimilitud, sin embargo una estimación puntual no proporciona ninguna información sobre la precisión involucrada en la estimación. Una manera usual de hacer inferencia sobre un parámetro es determinar un intervalo de valores plausibles del parámetro en base a la muestra observada, Sprott[25]. A ese intervalo se le conoce como intervalo de verosimilitud.

**Definición 2.3** *El conjunto de valores de  $\theta$  para los cuales  $R(\theta; \underline{x}) \geq c$ , donde  $0 \leq c \leq 1$ , es conocido como **intervalo o región de verosimilitud** de nivel  $c$  para  $\theta$  que denotaremos por  $IV(c)$ .*

Cada valor específico de  $\theta$  dentro de la región tiene una verosimilitud relativa  $R(\theta; \underline{x}) \geq c$ , y cada valor específico de  $\theta$  fuera de la región tiene una verosimilitud relativa  $R(\theta; \underline{x}) < c$ . Por tanto, el  $IV(c)$  separa valores plausibles  $\theta$  de los no plausibles a un nivel  $c$ .

Los intervalos o regiones de verosimilitud son de gran utilidad para comparar la plausibilidad de los valores del parámetro  $\theta$  con respecto a  $\hat{\theta}$  especificando rangos de los valores del parámetro más plausibles. Cuando  $\theta$  es unidimensional, gráficamente los intervalos de verosimilitud son resultado de trazar una línea horizontal a través de la gráfica de  $R(\theta; \underline{x})$  a una distancia  $c$  del eje de  $\theta$ . Variando  $c$  de 0 a 1 se produce un conjunto jerarquizado y anidado de intervalos de verosimilitud que converge al estimador de máxima verosimilitud  $\hat{\theta}$  cuando  $c$  converge a 1. De esta manera  $\hat{\theta}$  es un valor común en todos los intervalos, y eso sirve para especificar su localización, es decir, proporciona alguna idea de la simetría de la función de verosimilitud con respecto a  $\hat{\theta}$  y de cómo cambia la plausibilidad dentro del intervalo. Este conjunto de intervalos reproduce la forma de la función de verosimilitud relativa  $R(\theta; \underline{x})$ . La desviación de  $\hat{\theta}$  del centro geométrico del intervalo nos da una idea del sesgo de la función de verosimilitud y por ende de la vecindad de plausibilidad dentro del intervalo. Usualmente consideraremos  $IV(c)$  con valores de  $c = 0.258, 0.147$  y  $0.036$ .

**Observación 2.2** *Es importante enfatizar que la verosimilitud es una función puntual, y por lo tanto el nivel de plausibilidad  $c$  de un intervalo de verosimilitud, no es una declaración de la incertidumbre del intervalo. Es una declaración acerca de la plausibilidad relativa de cualquier punto individual dentro del intervalo, cuya verosimilitud es igual o mayor a  $c$ .*

### 2.1.3 Probabilidad de cobertura de Intervalos de verosimilitud

Los resultados que veremos en esta subsección son de gran relevancia al momento de asociar un nivel de confianza a los intervalos de verosimilitud. Lo haremos a

través de la convergencia en distribución del estadístico de prueba basado en el método de máxima verosimilitud, como lo es el estadístico de razón de verosimilitud (Neyman y Pearson, 1928)  $\Lambda = -2 \log R(\theta_0; x)$ , a una variable aleatoria Ji-cuadrada con un grado de libertad. Para la demostración de este resultado citaremos algunos antecedentes relevantes de los cuales no se incluye la demostración, para una mayor referencia de su desarrollo cito a Serfling[24].

Primeramente debemos considerar condiciones de regularidad para la familia de funciones de probabilidad que a continuación describiremos: Sea  $\Theta$  un intervalo abierto, no necesariamente finito se cumplen las siguientes condiciones de regularidad:

**R1:** Para  $\theta \in \Theta$ , las derivadas  $\frac{\partial}{\partial \theta} \log f(x; \theta)$ ,  $\frac{\partial^2}{\partial \theta^2} \log f(x; \theta)$  y  $\frac{\partial^3}{\partial \theta^3} \log f(x; \theta)$ , existen para toda  $x$ .

**R2:** Para  $\theta_0 \in \Theta$  existen funciones  $g(x)$ ,  $h(x)$  y  $H(x)$ , tal que para  $\theta$  en una vecindad de  $\theta_0$ :

$$\left| \frac{\partial \log f(x; \theta)}{\partial \theta} \right| \leq g(x), \quad \left| \frac{\partial^2 \log f(x; \theta)}{\partial \theta^2} \right| \leq h(x), \quad \left| \frac{\partial^3 \log f(x; \theta)}{\partial \theta^3} \right| \leq H(x);$$

se cumplen para toda  $\underline{x}$ . Además,  $\int g(x)dx < \infty$ ,  $\int h(x)dx < \infty$  y  $E_\theta[H(x)] < \infty$ .

**R3:** Para cada  $\theta \in \Theta$ , se satisface que  $0 < E_\theta[\partial \log f(x; \theta)/\partial \theta]^2 < \infty$ .

La condición (R1) nos asegura que la función  $\partial \log f(x; \theta)/\partial \theta$  tiene, para cada  $x$ , una expansión de Taylor como una función de  $\theta$ . La condición (R2) asegura que las funciones  $\int f(x; \theta)dx$  y  $\int [\partial \log f(x; \theta)/\partial \theta] dx$  pueden ser diferenciadas con respecto a  $\theta$  bajo la integral. La condición (R3) establece que las variables aleatorias  $\partial \log f(x; \theta)/\partial \theta$  tienen varianza finita positiva.

Otro antecedente que nos será de utilidad en la demostración es la siguiente proposición.

**Proposición 2.1** *Bajo las condiciones de regularidad R1-R3, consideremos observaciones independientes e idénticamente distribuidas en la familia de funciones de probabilidad, para  $\theta \in \Theta$ . Luego, con probabilidad 1 las ecuaciones de verosimilitud admiten una sucesión  $\{\hat{\theta}_n\}$  de estimadores que satisfacen:*

- a) *Consistencia, en el caso de poblaciones infinitas se dice que un estimador  $\hat{\theta}_n$  es consistente si  $(\hat{\theta}_n - \theta) \xrightarrow{P} 0$ , cuando  $n \rightarrow \infty$ .*
- b) *Asintocidad Normal,  $\sqrt{n}(\hat{\theta}_n - \theta)$  converge en distribución a una Normal con media cero y varianza  $I^{-1}(\theta; \underline{x})$ , es decir:  $\sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{d} N(0, I^{-1}(\theta; \underline{x}))$ .*

Por último, utilizaremos en la demostración el Lema de Slutsky relacionado con convergencia en distribución de sucesiones de variables aleatorias.

**Lema 2.3** Sean  $\{X_n\}$  y  $\{Y_n\}$  sucesiones de variables aleatorias tales que  $X_n \xrightarrow{d} X$  y  $Y_n \xrightarrow{P} c$  donde  $c$  es una constante. Entonces

$$\begin{aligned} X_n + Y_n &\xrightarrow{d} X + c. \\ X_n Y_n &\xrightarrow{d} cX. \\ \frac{X_n}{Y_n} &\xrightarrow{d} \frac{X}{c}; \quad \text{si } c \neq 0. \end{aligned}$$

Para demostrar la convergencia del estadístico de razón de verosimilitud lo haremos mediante una hipótesis nula simple del tipo  $H_0 : \theta = \theta_0$ . Con este resultado determinaremos para qué niveles de  $c$  los intervalos de verosimilitud se relacionan, con ciertos niveles de significancia, como por ejemplo  $\alpha = 0.1, 0.05$  y  $0.01$ , a través de los cuantiles de un variable aleatoria Ji-cuadrada con un grado de libertad. Proporcionando de esta manera, una probabilidad de cobertura para los intervalos de verosimilitud para los mencionados niveles de  $\alpha$ .

**Teorema 2.4** Bajo las condiciones de regularidad R1-R3 y bajo la hipótesis  $H_0 : \theta = \theta_0$ , el estadístico de razón de verosimilitud  $\Lambda = -2 \log R(\theta_0; x)$  converge en distribución a una Ji-cuadrada con un grado de libertad,  $\chi_{(1)}^2$ .

**Demostración.** Sea

$$\ell_n(\theta) = \log L(\theta);$$

construyamos la expansión de Taylor de  $\ell_n(\theta)$  alrededor de  $\theta = \hat{\theta}_n$ :

$$\ell_n(\theta_0) = \ell(\hat{\theta}_n) + (\hat{\theta}_n - \theta_0)\ell'_n(\hat{\theta}_n) + \frac{(\hat{\theta}_n - \theta_0)^2}{2!}\ell''_n(\hat{\theta}_n) + R_1(\theta_0, \hat{\theta}_n)$$

$$\ell_n(\theta_0) - \ell(\hat{\theta}_n) = n(\hat{\theta}_n - \theta_0)^2 I(\theta_0) \left[ \frac{\ell''_n(\hat{\theta}_n)}{2nI(\theta_0)} + R_2(\theta_0, \hat{\theta}_n) \right];$$

de donde obtenemos:

$$\Lambda = -2 \left[ \ell_n(\theta_0) - \ell(\hat{\theta}_n) \right] - 2 \log R(\theta_0; x) = n(\hat{\theta}_n - \theta_0)^2 I(\theta_0) \left[ \frac{\ell''_n(\hat{\theta}_n)}{nI(\theta_0)} + R_3(\theta_0, \hat{\theta}_n) \right];$$

considerando que bajo  $H_0 : \theta = \theta_0$  se cumple:

$$\sqrt{n}(\hat{\theta}_n - \theta_0)\sqrt{I(\theta_0)} \xrightarrow{d} N(0, 1) \implies n(\hat{\theta}_n - \theta_0)^2 I(\theta_0) \xrightarrow{d} \chi_{(1)}^2.$$

Por la Ley Fuerte de los Grandes Números tenemos:

$$\frac{\ell''_n(\theta_0)}{n} \xrightarrow{c.s} I(\theta_0) \implies \frac{\ell''_n(\hat{\theta}_n)}{nI(\theta_0)} \xrightarrow{P} 1.$$

Por otra parte por la consistencia  $(\hat{\theta}_n - \theta) \xrightarrow{P} 0$ , cuando  $n \rightarrow \infty$  tenemos

$$R_3(\theta_0, \hat{\theta}_n) \xrightarrow{P} 0.$$

Luego, por el lema de Slutsky

$$\Lambda \xrightarrow{d} \chi_{(1)}^2.$$

Por lo tanto, el estadístico de razón de verosimilitud converge en distribución a una variable aleatoria Ji-cuadrada con un grado de libertad. ■

Si se tiene una muestra observada  $x$  que proviene de la distribución de  $X$  con parámetro  $\theta_0$ , entonces se puede calcular un intervalo  $[A, B]$  para el valor verdadero  $\theta_0$ . La *probabilidad de cobertura* del intervalo aleatorio  $[A, B]$  es la probabilidad de que el intervalo  $[A, B]$  incluya o cubra el valor verdadero del parámetro  $\theta_0$ :

$$CP(\theta_0) = P(A \leq \theta_0 \leq B; \theta = \theta_0).$$

Ésta se interpreta como la fracción de las veces que el intervalo  $[A, B]$  podría incluir el valor verdadero  $\theta_0$  en un número grande de repeticiones del experimento con parámetro  $\theta_0$ . Los extremos del intervalo  $A$  y  $B$  son variables aleatorias ya que pueden variar cuando cambia la muestra. Nótese que la distribución de probabilidad de  $A$  y  $B$  se puede calcular a partir de la distribución de probabilidad de la muestra de variables observadas y que generalmente depende de  $\theta_0$ , pudiendo incluir algunas veces el valor verdadero  $\theta_0$  o no.

El intervalo aleatorio  $[A, B]$  es llamado *intervalo de confianza (IC)* para  $\theta$  si su probabilidad de cobertura  $CP(\theta_0) = P(A \leq \theta_0 \leq B | \theta = \theta_0)$  es la misma para todos los valores  $\theta_0$  del parámetro, esto es, cuando dicha probabilidad de cobertura no depende del valor de  $\theta_0$ . En primera instancia,  $[A, B]$  es un intervalo del 95% para  $\theta$  si  $P(A \leq \theta_0 \leq B | \theta = \theta_0) = 0.95$  para todos los posibles valores de  $\theta_0$ . Un *IC* del 95% podría incluir el valor verdadero  $\theta_0$  del parámetro en un 95% de las repeticiones del experimento.

La probabilidad de cobertura de un *IV(c)*, se puede calcular a través de la distribución de probabilidad del estadístico de razón de verosimilitud  $\Lambda$  para  $\theta = \theta_0$ ,  $\Lambda \equiv -2 \log R(\theta_0; \underline{x})$ . Un valor particular del parámetro,  $\theta_0$  pertenece a un *IV(c)* para  $\theta$  si y sólo si  $R(\theta_0; \underline{x}) \geq c$ , o equivalentemente  $-2 \log R(\theta_0; \underline{x}) \leq -2 \log c$ . De esta manera, la probabilidad de cobertura del *IV(c)* es:

$$\begin{aligned} CP(\theta_0) &= P[\theta_0 \in IV(c); \theta = \theta_0] \\ &= P(\Lambda \leq -2 \log c; \theta = \theta_0). \end{aligned} \tag{2.4}$$

En general es difícil encontrar la distribución de probabilidad exacta de  $\Lambda$ ; sin embargo por el Teorema 2.4 se cumple que el estadístico de la razón de verosimilitud  $\Lambda \equiv -2 \log R(\theta_0; \underline{x})$  converge a una Ji-cuadrada con un grado de libertad,  $\chi_{(1)}^2$ , para todo  $\theta_0 \in \Theta \subset \mathbb{R}$ . Luego de (2.4) se tiene que  $P(\Lambda \leq d; \theta = \theta_0)$  es la probabilidad de cobertura del *IV(c)*.

Por (2.4),  $P(\Lambda \leq d)$  es la probabilidad de cobertura del *IV(c)*, donde  $d = -2 \log c$ ; ya que  $P(\Lambda \leq -2 \log c)$ , entonces  $c = e^{-\frac{d}{2}}$ . Si se selecciona un valor de

$(1 - \alpha)$	$c$	$d = q(\alpha, 1)$
0.90	0.258	2.706
0.95	0.147	3.841
0.99	0.036	6.635

Tabla 2.1: Confianza aproximada de  $IV(c)$  cuando  $\theta$  es unidimensional

tablas  $d = q(\alpha, 1) = 3.841$ , donde  $q(\alpha, 1)$  es el cuantil  $(1 - \alpha)$  de una Ji-cuadrada con un grado de libertad, tenemos que  $c = e^{-\frac{3.841}{2}} = 0.147$ , por tanto  $P(\Lambda \leq 3.841)$  es la probabilidad de cobertura del  $IV(c)$  con  $c = 0.147$ . Obsérvese en la Tabla 2.1 que 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, respectivamente.

Por lo que los  $IV(c)$  con  $c = 0.258$ , 0.147 y 0.036 tienen una probabilidad de cobertura aproximada del 90%, 95% y 99% respectivamente. De esta manera, el  $IV(c)$  también es un intervalo o región de confianza para el parámetro  $\theta$  y toma en consecuencia el nombre de *intervalos de verosimilitud-confianza* para  $\theta$ .

La mejor manera de construir un intervalo de confianza es a partir de la función de verosimilitud, con la finalidad de no caer en el caso en que se estén rechazando valores plausibles por no ser cubiertos por el intervalo. Cuando aplicamos un estadístico de razón de verosimilitud  $\Lambda = -2 \log R(\theta_0; \underline{x}) \approx \chi_{(1)}^2$ , el coeficiente de confianza aproximado, es decir, la probabilidad de cobertura del  $IV(c)$  está dada por:

$$CP \approx P(\chi_{(1)}^2 \leq -2 \log c).$$

Porque de la aproximación Ji-cuadrada, los intervalos de verosimilitud son exactos o se aproximan a los intervalos de confianza en la mayoría de las aplicaciones. Como el cuadrado de una variable aleatoria Ji-cuadrada con un grado de libertad es una normal estándar, por el inciso (b) de la Proposición 2.2,  $\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{d} N(0, I^{-1}(\theta_0; \underline{x}))$ , tenemos:

$$\frac{\sqrt{n}(\hat{\theta}_n - \theta_0)}{\sqrt{I^{-1}(\theta_0; \underline{x})}} \xrightarrow{d} N(0, 1);$$

por el Lema 2.3 de Slutsky,

$$\frac{\sqrt{n}(\hat{\theta}_n - \theta_0)}{\sqrt{I^{-1}(\hat{\theta}_n; \underline{x})}} \xrightarrow{d} N(0, 1);$$

donde  $\sqrt{I^{-1}(\hat{\theta}_n; \underline{x})}$  es un pivotal, ya que es una variable aleatoria que no depende de  $\theta$ . Esto es, es constante para toda  $\theta$ .

Una de las aplicaciones de la Proposición 2.2, de asintoticidad normal de  $\hat{\theta}_n$  es la siguiente definición.

**Definición 2.4** El intervalo  $\hat{\theta}_n \pm Z_{1-\alpha/2} \sqrt{I^{-1}(\hat{\theta}_n; \underline{x})/n}$  es llamado “*intervalo de Wald*” y es un intervalo asintótico de confianza  $1 - \alpha$  cuando  $n \rightarrow \infty$ .

La anterior definición es equivalente a que la probabilidad de que  $\theta_0$  sea cubierto por esos dos extremos converge a  $1 - \alpha$ , cuando  $n \rightarrow \infty$ ,

$$CP(\theta_0) = P\left(\hat{\theta}_n - Z_{\alpha/2} \sqrt{I^{-1}(\hat{\theta}_n; x)/n} \leq \theta_0 \leq \hat{\theta}_n + Z_{\alpha/2} \sqrt{I^{-1}(\hat{\theta}_n; x)/n}\right) \xrightarrow{n \rightarrow \infty} 1 - \alpha, \quad (2.2)$$

$$(2.5)$$

ya que

$$\begin{aligned} CP(\theta_0) &= P\left(-Z_{\frac{\alpha}{2}} \leq \frac{(\hat{\theta}_n - \theta_0)}{\sqrt{I^{-1}(\hat{\theta}_n; x)/n}} \leq Z_{\frac{\alpha}{2}}\right) \\ &= P\left(-Z_{\frac{\alpha}{2}} \sqrt{I^{-1}(\hat{\theta}_n; x)/n} \leq (\hat{\theta}_n - \theta_0) \leq Z_{\frac{\alpha}{2}} \sqrt{I^{-1}(\hat{\theta}_n; x)/n}\right) \\ &= P\left(\hat{\theta}_n - Z_{\frac{\alpha}{2}} \sqrt{I^{-1}(\hat{\theta}_n; x)/n} \leq \theta_0 \leq \hat{\theta}_n + Z_{\frac{\alpha}{2}} \sqrt{I^{-1}(\hat{\theta}_n; x)/n}\right) = 1 - \alpha, \end{aligned}$$

donde  $Z \sim N(0, 1)$ .

#### 2.1.4 Función de verosimilitud perfil o maximizada

Generalmente se presentan modelos con varios parámetros donde el interés es estimar un parámetro cuando se desconce todo sobre los demás. A estos últimos los llamaremos parámetros de estorbo. El problema de estimar por separado parámetros de interés en presencia de parámetros de estorbo es relevante en estadística por el impacto que éstos pueden provocar en la inferencia acerca de los parámetros de interés. Cabe mencionar que un parámetro puede ser considerado en algunos casos como de estorbo y en otros no.

Supongamos que el modelo de probabilidad para un experimento involucra dos parámetros desconocidos,  $\alpha$  y  $\beta$ , y que sólo nos interesa hacer inferencia sobre  $\beta$ , con  $\alpha$  considerado como un parámetro de estorbo; la verosimilitud perfil es una alternativa que se ha desarrollado para eliminar tal parámetro de estorbo, de tal forma que la verosimilitud se puede escribir solamente como una función del parámetro de interés.

**Definición 2.5** Definimos la *función de verosimilitud perfil* de  $\beta$ ,  $L_p(\beta; \underline{x})$ , como el máximo de  $L(\alpha, \beta; \underline{x})$  sobre  $\alpha$  con  $\beta$  fijo, en otras palabras

$$L_p(\beta; \underline{x}) = \max_{\alpha} L(\alpha, \beta; \underline{x}) = L(\hat{\alpha}(\beta), \beta; \underline{x}),$$

donde  $\hat{\alpha}(\beta)$  es el EMV restringido de  $\alpha$  por un valor específico de  $\beta$ .

Si  $P(\underline{x}; \alpha, \beta)$  es el modelo de probabilidad, tenemos que

$$L(\hat{\alpha}(\beta), \beta; \underline{x}) \propto P(\underline{x}; \hat{\alpha}(\beta), \beta).$$

La función de verosimilitud perfil relativa de  $\beta$ , es el valor máximo de  $R(\alpha, \beta; \underline{x})$  sobre  $\alpha$ , con  $\beta$  fijo:

$$R_p(\hat{\alpha}(\beta), \beta; \underline{x}) = \max_{\alpha} R(\alpha, \beta; \underline{x}).$$

Si  $\hat{\alpha}(\hat{\beta}) = \hat{\alpha}$ , entonces

$$R_p(\beta; \underline{x}) = \frac{L(\hat{\alpha}(\beta), \beta; \underline{x})}{L(\hat{\alpha}(\hat{\beta}), \hat{\beta}; \underline{x})} = R(\hat{\alpha}(\beta), \beta; \underline{x}).$$

Las propiedades de  $R_p(\beta; \underline{x})$  son similares a la función de verosimilitud relativa vista en la Sección 1, verosimilitud relativa con un parámetro.

Note que  $0 \leq R(\alpha, \beta; \underline{x}) \leq 1$  y  $R(\hat{\alpha}, \hat{\beta}; \underline{x}) = 1$ . Podemos pensar la gráfica de  $R(\alpha, \beta; \underline{x})$  como una "montaña" de conjuntos de verosimilitud sobre el plano  $\alpha, \beta$ . Así, cuando uno se posiciona sobre el eje de estorbo  $\alpha$ , obtenemos una silueta o perfil de  $L_p(\beta; \underline{x})$ , que es justamente la función de verosimilitud maximizada. La curva  $R(\alpha, \beta; \underline{x}) = c$ , la cual forma la frontera de esa región de verosimilitud, es llamada *contorno de verosimilitud* y lo denotaremos  $CV(c)$ . Equivalentemente el  $CV(c)$  para  $\alpha$  y  $\beta$  es el conjunto de valores  $(\alpha, \beta)$  tales que  $R(\alpha, \beta; \underline{x}) \geq \log c$ . Una manera conveniente de dibujar  $R(\alpha, \beta; \underline{x})$  en dos dimensiones es trazar contornos de verosimilitudes relativas constantes en el plano  $\alpha, \beta$ . Usualmente los contornos formarán un conjunto anidado de curvas de forma toscamente elípticas alrededor del punto  $(\hat{\alpha}, \hat{\beta})$ , entre más elípticos sean los contornos es señal de que existe alguna relación de dependencia entre los parámetros involucrados.

El procedimiento para el cálculo del  $EMV$   $(\hat{\alpha}, \hat{\beta})$ , es igual que el caso para un parámetro. Para el caso de dos parámetros la *información de Fisher* es una matriz simétrica 2 por 2,

$$I(\alpha, \beta; \underline{x}) = \begin{bmatrix} I_{11}(\alpha, \beta; \underline{x}) & I_{12}(\alpha, \beta; \underline{x}) \\ I_{21}(\alpha, \beta; \underline{x}) & I_{22}(\alpha, \beta; \underline{x}) \end{bmatrix} = \begin{bmatrix} -\frac{\partial^2 \ell(\alpha, \beta; \underline{x})}{\partial \alpha^2} & -\frac{\partial^2 \ell(\alpha, \beta; \underline{x})}{\partial \alpha \partial \beta} \\ -\frac{\partial^2 \ell(\alpha, \beta; \underline{x})}{\partial \alpha \partial \beta} & -\frac{\partial^2 \ell(\alpha, \beta; \underline{x})}{\partial \beta^2} \end{bmatrix} \quad (2.3)$$

Para que la matriz anterior tenga un máximo relativo, la matriz de la información observada  $I(\hat{\alpha}, \hat{\beta}; \underline{x})$  debe ser positiva definida, esto es, que el determinante del sistema sea mayor que cero. Es decir,

$$\hat{I}_{11} > 0; \hat{I}_{22} > 0; \hat{I}_{11}\hat{I}_{22} - \hat{I}_{12}^2 > 0; \quad (2.6a)$$

donde  $\hat{I}_{ij} = I_{ij}(\hat{\alpha}, \hat{\beta}; \underline{x})$ .

Esta ecuación puede ser resuelta para  $\beta$  como en el caso de un parámetro. El método de Newton-Raphson nos será de mayor utilidad, ya que nos proporciona los valores estimados de los parámetros, cuando no es posible resolver analíticamente el vector de ecuaciones formado por las funciones Score. Sea  $I^{-1}(\alpha, \beta; \underline{x})$  la matriz inversa de la matriz  $I(\alpha, \beta; \underline{x})$  y definimos la  $i$ -ésima entrada de la diagonal de  $I^{-1}(\alpha, \beta; \underline{x})$  como  $\{I^{-1}(\alpha, \beta; \underline{x})\}^{ii}$ . Así, la primera entrada de la matriz de información perfil para  $\beta$ , por ejemplo, es  $I(\beta; \underline{x}) = 1/I^{-1}(\alpha; \underline{x})^{11}$ .

Para el caso de dos parámetros tenemos

$$\begin{bmatrix} \alpha_1 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} \alpha_0 \\ \beta_0 \end{bmatrix} + \begin{bmatrix} I_{11} & I_{12} \\ I_{12} & I_{22} \end{bmatrix}^{-1} \begin{bmatrix} S_1 \\ S_2 \end{bmatrix}, \quad (2.6b)$$

donde las componentes del vector Score y la inversa de la matriz de información son todos evaluados en  $\alpha = \alpha_0$  y  $\beta = \beta_0$ . Empezando con  $\alpha_0$  y  $\beta_0$  como valores iniciales, aplicamos la expresión (2.6a) repetidamente hasta que obtenemos la convergencia, y finalmente verificamos que la condición (2.6b) se satisface para un máximo relativo.

## 2.2 Inferencia en procesos de Poisson homogéneos

El objetivo de esta sección apoyada en Basawa[1], es establecer la relación entre la teoría estadística, bajo el enfoque de verosimilitud desarrollada en la Sección 1 de este capítulo y la teoría de probabilidad, al realizar inferencia en los parámetros de un proceso de Poisson. Analizaremos algunas formas de llevar a cabo la inspección de un proceso de Poisson homogéneo con la finalidad de aplicar el método de máxima verosimilitud para estimar el parámetro del proceso y construir intervalos de confianza.

Sea  $\{N(t)\}_{t \geq 0}$  un proceso de Poisson homogéneo con parámetro  $\lambda$  de acuerdo a la Ddefinición 1.7, esto es, un proceso que cumple con:  $P[N(0) = 0] = 1$ , para  $0 \leq S_0 < S_1 < S_2 < \dots < S_n$  tenemos que  $N((S_i, S_{i+1}])$ ,  $i = 1, 2, \dots, n$  son variables aleatorias independientes y que para cada  $s, t \geq 0$ ,

$$P[N((s, t]) = n] = \frac{e^{-\lambda(t-s)} \lambda(t-s)^n}{n!}; \quad n = 0, 1, 2, \dots;$$

donde  $N(t)$  es el número de eventos ocurridos hasta un tiempo  $t$ . El objetivo es estimar el parámetro del proceso  $\lambda$  usando algunas de las realizaciones de un proceso de Poisson, que aquí se mencionan, .

### 2.2.1 Observación continua en un intervalo de tiempo fijo

Supongamos que se observa un proceso de Poisson homogéneo en un intervalo de tiempo fijo  $[0, T_0]$ . Sea  $N'$  el número total de eventos que ocurrieron durante ese intervalo. Sea el vector  $\underline{S} = (S_1, S_2, \dots, S_{N'})$  los tiempos de ocurrencia de cada uno de los eventos en el tiempo, sea  $T_k = S_k - S_{k-1}$ , para  $k = 1, 2, \dots$  con  $S_0 = 0$ , los tiempos entre eventos consecutivos.

Asumiendo que  $N(0) = 0$ , la función de verosimilitud para  $\lambda$  con base en la realización está dada por:

$$L(\lambda; \underline{S}) \propto P \left[ \begin{array}{l} N(S_1) - N(S_0) = 1, N(S_2) - N(S_1) = 1, \dots, \\ N(S_{N'}) - N(S_{N'-1}) = 1, N(T_0) - N(S_{N'}) = 0 \end{array} \right].$$

Por ser  $N$  un proceso de Poisson, se cumple que los incrementos son variables aleatorias independientes y que éstos se distribuyen Poisson; luego, se tiene que

$$\begin{aligned} L(\lambda; \underline{S}) &\propto \left( \prod_{k=1}^{N'} P[N(S_k) - N(S_{k-1}) = 1] \right) P[N(T_0) - N(S_{N'}) = 0] \\ &= \left( c(\underline{S}) \prod_{k=1}^{N'} e^{\{-\lambda(S_k - S_{k-1})\}} [\lambda(S_k - S_{k-1})] \right) e^{\{-\lambda(T_0 - S_{N'})\}} \\ &= \left( c(\underline{S}) \prod_{k=1}^{N'} e^{-\lambda T_k} \lambda T_k \right) e^{-\lambda(T_0 - S_{N'})} \\ &= c(\underline{S}) \cdot e^{-\lambda \sum_{k=1}^{N'} T_k} \lambda^{N'} \left( \prod_{k=1}^{N'} T_k \right) e^{-\lambda(T_0 - S_{N'})}, \end{aligned}$$

donde  $c(\underline{S})$  es una función constante de  $\underline{S}$ , que no depende de  $\lambda$ . Ya que  $\prod_{k=1}^{N'} T_k$  es una constante por no depender del parámetro  $\lambda$  tenemos,

$$L(\lambda; \underline{S}) = c(\underline{S}) \cdot \lambda^{N'} e^{-\lambda \sum_{k=1}^{N'} T_k} e^{-\lambda(T_0 - S_{N'})},$$

pero como  $\sum_{k=1}^{N'} T_k = S'_{N'}$ , se sigue

$$L(\lambda; \underline{S}) = c(\underline{S}) \cdot \lambda^{N'} e^{-\lambda T_0}. \quad (2.7)$$

Aplicando logaritmo a la función de verosimilitud obtenemos la función log-verosimilitud dada por

$$\ell(\lambda; \underline{S}) = c'(\underline{S}) + \log L(\lambda; \underline{S}) = c'(\underline{S}) + N' \log \lambda - \lambda T_0. \quad (2.8)$$

Considerando ahora la función Score para encontrar el máximo verosímil tenemos:

$$sc(\lambda; \underline{S}) = \frac{d\ell(\lambda; \underline{S})}{d\lambda} = \frac{N'}{\lambda} - T_0; \quad (2.9)$$

igualando a cero y resolviendo la ecuación, obtenemos la expresión del estimador de máxima verosimilitud para  $\lambda$ :

$$\hat{\lambda} = \frac{N'}{T_0}; \quad (2.10)$$

donde  $N' \sim \text{Poisson}(\lambda)$ .

Para finalizar, determinaremos un intervalo de confianza para este método de observación continua de un proceso de Poisson homogéneo en un período de tiempo fijo. Ya que  $N'$  sigue una distribución de Poisson con media  $\lambda T_0$  en  $[0, T_0]$ , tenemos que el valor esperado del estimador de máxima verosimilitud es:

$$E[\hat{\lambda}] = \frac{E[N']}{T_0} = \frac{\lambda T_0}{T_0} = \lambda.$$

Por lo tanto,  $\hat{\lambda}$  es un estimador insesgado y su varianza es:

$$V(\hat{\lambda}) = \frac{V(N')^2}{(T_0)^2} = \frac{\lambda T_0}{(T_0)^2} = \frac{\lambda}{T_0}.$$

Ya que hemos considerado la esperanza y varianza del estimador de máxima verosimilitud  $\hat{\lambda}$ , para fines comparativos con los intervalos de verosimilitud que más adelante mostraremos, un intervalo de confianza para este tipo de inspección de un proceso de Poisson homogéneo, es el intervalo de Wald con el 95% de probabilidad de cobertura, de (2.5) de la Definición 2.4 tenemos:

$$\hat{\lambda} \pm Z_{\alpha/2} \sqrt{I^{-1}(\hat{\lambda}; x)/N'} = \hat{\lambda} \mp 1.96 \sqrt{I^{-1}(\hat{\lambda}; x)} = \hat{\lambda} \mp 1.96 \sqrt{V(\hat{\lambda})}. \quad (2.11)$$

Como podemos observar el intervalo de confianza de Wald es simétrico y no depende de la función de verosimilitud. Así por ejemplo si observamos una colección de datos que presentan variación asimétrica, con el intervalo de Wald corremos el riesgo de rechazar valores plausibles por no ser cubiertos por el intervalo, o peor aún, considerar en su interior valores implausibles, como ilustraremos en los ejemplos simulados en el capítulo siguiente.

### 2.2.2 Observación continua en un período de tiempo aleatorio

Supongamos que empezamos a observar el proceso de Poisson homogéneo desde un tiempo  $T = 0$  y seguimos hasta que ocurran  $N_0$  eventos. Sean  $0 < S_1 < S_2 < \dots < S_{N_0}$  los tiempos en los que ocurren los eventos, y sean  $T_k = S_k - S_{k-1}$  los tiempos entre eventos consecutivos, variables aleatorias independientes e idénticamente distribuidas con función de densidad  $f_{T_k}(t_k) = \lambda(t_k)e^{-\lambda t_k}$ . Por el Teorema 1.8, podemos expresar los incrementos independientes que siguen una distribución de Poisson en términos de la distribución de las variables aleatorias  $T_k$  de la siguiente manera.

Tomando en cuenta las consideraciones de la realización anterior, entonces la función de verosimilitud para  $\lambda$  queda expresada por:

$$\begin{aligned} L(\lambda; \underline{S}) &= \left( c(\underline{S}) \prod_{k=1}^{N_0} P[N(S_k) - N(S_{k-1}) = 1] \right) P[N(\check{T}) - N(S_{N_0}) = 0] \\ &= \left( c(\underline{S}) \prod_{k=1}^{N_0} e^{\{-\lambda(S_k - S_{k-1})\}} [\lambda(S_k - S_{k-1})] \right) e^{\{-\lambda(\check{T} - S_{N_0})\}} \\ &= \left( c(\underline{S}) \prod_{k=1}^{N_0} e^{-\lambda T_k} \lambda T_k \right) e^{-\lambda(\check{T} - S_{N_0})} \\ &= c(\underline{S}) \cdot e^{-\lambda \sum_{k=1}^{N_0} T_k} \lambda^{N_0} \left( \prod_{k=1}^{N_0} T_k \right) e^{-\lambda(\check{T} - S_{N_0})}; \end{aligned}$$

donde  $c(\underline{S})$  es una función constante de  $\underline{S}$  que no depende de  $\lambda$  y  $\check{T}$  es el tiempo aleatorio que se tuvo que esperar para que ocurrieran  $N_0$  eventos. Debido a que

$\prod_{k=1}^{N_0} T_k$  es una constante que no depende de  $\lambda$  y  $\sum_{k=1}^{N_0} T_k = S_{N_0}$  tenemos:

$$L(\lambda; \underline{S}) = c(\underline{S}) \cdot \lambda^{N_0} e^{-\lambda \check{T}}.$$

De (1.5) la variable  $N$  es el entero  $N_0$  más grande tal que  $S_{N_0} \leq \check{T}$ , por tanto  $\check{T} = S_{N_0} = \sum_{k=1}^{N_0} T_k$ .

Siguiendo el algoritmo para determinar el estimador de máxima verosimilitud. La función log-verosimilitud correspondiente está dada por

$$\ell(\lambda; \underline{S}) = c'(\underline{S}) + N_0 \lambda - \lambda \check{T}.$$

Derivando esta función con respecto a  $\lambda$  obtenemos la función Score  $sc(\lambda; \underline{S})$ ;

$$sc(\lambda; \underline{S}) = \frac{d\ell(\lambda; \underline{S})}{d\lambda} = \frac{N_0}{\lambda} - \check{T}.$$

Por lo tanto, el estimador de máxima verosimilitud de  $\lambda$  es la solución de la ecuación  $sc(\lambda; \underline{S}) = 0$ :

$$\hat{\lambda} = \frac{N_0}{\check{T}}.$$

Procederemos ahora a determinar un intervalo de confianza para la observación continua de un proceso de Poisson homogéneo en tiempo aleatorio. Ya que  $\check{T}$  es la suma de  $N_0$  variables exponenciales independientes, está distribuida como una variable aleatoria gama con densidad

$$f_{\check{T}}(t) = \frac{\lambda(\lambda t)^{N_0-1}}{(N_0-1)!} e^{-\lambda t} \quad t > 0.$$

Como  $E\left[\frac{1}{\check{T}}\right] = \frac{\lambda}{N_0-1}$  (asumimos que  $N_0 > 1$ ) se sigue entonces que:

$$\begin{aligned} E[\hat{\lambda}] &= \int_0^\infty \frac{N_0}{t} \frac{\lambda^{N_0} t^{N_0-1}}{\Gamma(N_0)} e^{-\lambda t} dt = N_0 \int_0^\infty \frac{\lambda^{N_0} t^{(N_0-1)-1}}{\Gamma(N_0)} e^{-\lambda t} dt \\ &= \frac{\lambda N_0 \Gamma(N_0-1)}{\Gamma(N_0)} \int_0^\infty \underbrace{\frac{\lambda^{N_0-1} t^{(N_0-1)}}{\Gamma(N_0-1)}} e^{-\lambda t} dt; \end{aligned}$$

pero la expresión señalada en la integral es una función de densidad Gama con parámetros  $(N_0-1, \lambda)$  siendo 1 el valor de la integral; por tanto el valor esperado del *EMV*,

$$\begin{aligned} E[\hat{\lambda}] &= E\left[\frac{N_0}{\check{T}}\right] = \frac{\lambda N_0}{N_0-1} \\ E[\hat{\lambda}] &= \frac{\lambda N_0 \Gamma(N_0-1)}{(N_0-1)\Gamma(N_0-1)} = \frac{\lambda N_0}{(N_0-1)}, \end{aligned}$$

que como se puede ver es sesgado, pero podemos construir un estimador insesgado de  $\hat{\lambda}$  considerando un *EMV* modificado:

$$\tilde{\lambda} = \frac{N_0 - 1}{N_0} \hat{\lambda} = \frac{N_0 - 1}{N_0} \frac{N_0}{\check{T}} = \frac{N_0 - 1}{\check{T}},$$

y su varianza  $V(\tilde{\lambda}) = E[\tilde{\lambda}^2] - (E[\tilde{\lambda}])^2$  es:

$$E[\tilde{\lambda}^2] = (N_0 - 1)^2 E\left[\frac{1}{\check{T}}\right] = (N_0 - 1)^2 \int_0^\infty \frac{1}{t^2} \frac{\lambda^{N_0} t^{N_0-1}}{\Gamma(N_0)} e^{-\lambda t} dt;$$

completando la integral para una variable con densidad Gama con parámetros  $(N_0 - 2, \lambda)$ , la varianza nos da

$$V(\tilde{\lambda}) = \frac{\lambda^2}{N_0 - 2};$$

asumimos que  $N_0 > 2$ .

La función de densidad de probabilidad de  $\check{T}$ , puede ser usada para construir, mediante el método pivotal, intervalos de confianza para  $\lambda$ . Enseguida veremos que  $2\lambda\check{T}$  está distribuido como una  $\chi^2(2N_0)$ . Sea  $Y = 2\lambda\check{T}$ , la función de distribución

$$\begin{aligned} F_Y(y) &= P[Y \leq y] = P[2\lambda\check{T} \leq y] = P\left[\check{T} \leq \frac{y}{2\lambda}\right] \\ &= F_{\check{T}}\left(\frac{y}{2\lambda}\right); \end{aligned}$$

esto implica que la función de densidad generada por  $Y$  es

$$\begin{aligned} f_Y(y) &= \frac{\partial F_Y(y)}{\partial y} = \frac{\partial F_{\check{T}}\left(\frac{y}{2\lambda}\right)}{\partial y} = \frac{1}{2\lambda} f_{\check{T}}\left(\frac{y}{2\lambda}\right) = \frac{1}{2\lambda} \frac{\lambda^{N_0} \frac{y}{2\lambda}^{N_0-1}}{\Gamma(N_0)} e^{-\lambda\left(\frac{y}{2\lambda}\right)} \\ &= \frac{1}{2} \frac{\left(\frac{y}{2}\right)^{N_0-1}}{\Gamma(N_0)} e^{-\frac{y}{2}} = \frac{1}{2^{N_0}} \frac{y^{N_0-1}}{\Gamma(N_0)} e^{-\frac{y}{2}} = \frac{y^{\frac{2N_0}{2}-1}}{2^{\frac{2N_0}{2}} \Gamma\left(\frac{2N_0}{2}\right)} e^{-\frac{1}{2}y}; \quad y > 0. \end{aligned}$$

Por tanto  $2\lambda\check{T}$  es una variable aleatoria  $\chi^2(2N_0)$ , Ji-cuadrada con  $2N_0$  grados de libertad. Entonces el intervalo de confianza es obtenido de la siguiente manera:

$$P\left[\chi_{2N_0, \frac{\alpha}{2}}^2 \leq 2\lambda\check{T} \leq \chi_{2N_0, 1-\frac{\alpha}{2}}^2\right] = 1 - \alpha,$$

dividiendo todo entre  $2\check{T}$  tenemos:

$$P\left[\frac{\chi_{2N_0, \frac{\alpha}{2}}^2}{2\check{T}} \leq \frac{2\lambda\check{T}}{2\check{T}} \leq \frac{\chi_{2N_0, 1-\frac{\alpha}{2}}^2}{2\check{T}}\right] = 1 - \alpha.$$

Por lo tanto, el intervalo de confianza para  $\lambda$  es:

$$\left[\frac{\chi_{2N_0, \frac{\alpha}{2}}^2}{2\check{T}}, \frac{\chi_{2N_0, 1-\frac{\alpha}{2}}^2}{2\check{T}}\right].$$

Existen otras realizaciones o formas de inspección de los procesos de Poisson homogéneos que no serán tratadas en este trabajo de tesis, tales como: inspección continua con un plan de muestreo, realizaciones discretas, etc. Pero para mayor referencia al lector consultar Basawa[1].

Nuevamente, a pesar de que se han considerado intervalos de confianza relacionados con cada una de las realizaciones de inspección vistas en esta sección, éstos son simétricos con respecto al valor verdadero del parámetro, pudiendo no considerar valores plausibles del parámetro por quedar fuera del intervalo de confianza, o bien, considerar valores implausibles del parámetro. Por tal razón, el mejor intervalo de confianza a construir sea el intervalo de verosimilitud-confianza asociado con un nivel de confianza mediante los resultados obtenidos en la subsección 2.1.3 que relaciona una probabilidad de cobertura a los intervalos de verosimilitud. Ya que los intervalos de verosimilitud-confianza dependen directamente de la distribución que presentan los datos observados, esto es, si los datos muestran cierta tendencia asimétrica, los intervalos de verosimilitud lo reflejarán en su longitud asimétrica con respecto a la posición del *EMV*. Siendo éstos últimos por tanto, la mejor opción para la inferencia en verosimilitud que nos interesa aplicar a los procesos de Poisson.

### 2.3 Inferencia en procesos de Poisson no homogéneos

En la presente sección relacionaremos las teorías de probabilidad e inferencia estadística en verosimilitud para un proceso de Poisson no homogéneo con un modelo de función de intensidad que involucra dos parámetros desconocidos, y su desarrollo resulta de mayor relevancia para esta tesis ya que es la parte medular para el caso de aplicación, pues en la mayoría de los procesos de Poisson existentes en nuestro entorno resultan del tipo no homogéneo. Esta parte nuevamente está apoyada en la referencia Basawa[1].

Sea  $\{N(t)\}_{t \geq 0}$ , un proceso de Poisson no homogéneo o dependiente del tiempo con media  $\mu(t) = \int_0^t \lambda(u) du$ , donde  $\lambda(t)$  es la función de intensidad del proceso que representa la tasa de ocurrencia en términos del tiempo.

Sean  $0 < S_1 < S_2 < \dots < S_N$  los tiempos de ocurrencia de cada evento, de acuerdo a la Definición 1.8, el proceso cumple con las siguientes condiciones:

- (i)  $P[N(0) = 0] = 1$ .
- (ii) Para  $0 < S_1 < S_2 < \dots < S_N$  tenemos que  $N(S_k) - N(S_{k-1})$ ;  $k = 1, 2, \dots, N$  son variables aleatorias independientes.
- (iii) Para cada  $t \geq 0$

$$P[N((0, t]) = n] = \frac{e^{-\int_{s_{k-1}}^{s_k} \lambda(t) dt} \left( \int_{s_{k-1}}^{s_k} \lambda(t) dt \right)^n}{n!}; \quad n = 0, 1, 2, \dots$$

Donde  $N(t)$  denota el número de eventos que ocurren hasta el tiempo  $t$ .

Entonces los tiempos entre eventos consecutivos,  $T_k = S_k - S_{k-1}$ , son variables aleatorias independientes e idénticamente distribuidas exponenciales con densidades no homogéneas:

$$f_{T_k}(t_k) = \lambda(t_k) e^{-\int_{S_{k-1}}^{S_k} \lambda(s) ds}. \quad (2.12)$$

Ya que el proceso es no homogéneo entonces la función de intensidad, que depende de los parámetros que involucre el modelo seleccionado, no es estacionaria como en el caso homogéneo, es decir, depende del tiempo.

En los casos de inspección de un proceso de Poisson homogéneo vistos en la sección anterior se logró determinar explícitamente la función de verosimilitud en términos de las variables tiempos entre eventos consecutivos, esto es, en términos de su función de densidad que es una exponencial. Por lo que para determinar una expresión para la función de verosimilitud para un proceso de Poisson no homogéneo lo haremos de manera similar, es decir, la función de verosimilitud la pondremos en términos de la densidad de las variables  $T_k$ , dada por (2.12).

Para realizar inferencia estadística basada en verosimilitud a este tipo de procesos es necesario desarrollar un ejemplo de un proceso de Poisson no homogéneo con una función de intensidad específica, como veremos a continuación en el siguiente ejemplo.

**Ejemplo 2.5** *Supongamos que se va a inspeccionar un proceso de Poisson no homogéneo  $\{N(t)\}_{t \geq 0}$  con función de intensidad  $\lambda(t; \alpha, \beta) = \alpha t^{-\beta}$ ;  $\alpha > 0$ ,  $0 < \beta < 1$  en un intervalo de tiempo fijo  $(0, T_0)$ . Sean  $0 < S_1 < S_2 < \dots < S_N \leq T_0$  los tiempos de ocurrencia, donde  $N$  es el número de eventos que ocurren en el intervalo. Los tiempos entre eventos son las variables aleatorias  $T_k = S_k - S_{k-1}$ ,  $k = 1, 2, \dots$  con función de densidad*

$$f_{T_k}(t_k) = \lambda(t_k) e^{-\int_{S_{k-1}}^{S_k} \lambda(s) ds}.$$

Como mencionamos con anterioridad para  $\underline{S} = (S_1, S_2, \dots, S_N)$ , la función de verosimilitud que depende de los parámetros  $\alpha$  y  $\beta$  está expresada en términos de la densidad de los tiempos entre eventos consecutivos.

$$\begin{aligned} L(\alpha, \beta; \underline{S}) &\propto \prod_{k=1}^N \left( \lambda(S_k; \alpha, \beta) e^{-\int_{S_{k-1}}^{S_k} \lambda(s; \alpha, \beta) ds} \right) \\ &= c(\underline{S}) \prod_{k=1}^N (\lambda(S_k; \alpha, \beta)) e^{-\left(\sum_{k=1}^N \int_{S_{k-1}}^{S_k} \lambda(s; \alpha, \beta) ds\right)} \\ &= c(\underline{S}) \prod_{k=1}^N \left( \alpha S_k^{-\beta} \right) e^{-\int_0^{T_0} \alpha s^{-\beta} ds}, \end{aligned}$$

donde  $c(\underline{S})$  es una función constante arbitraria positiva de  $\underline{S}$  que no depende de  $\alpha$  y  $\beta$ . Aplicando el algoritmo que hemos venido utilizando para determinar los estimadores de máxima verosimilitud para cada uno de los parámetros, consideramos

ahora el logaritmo de la función de verosimilitud, luego tenemos que la función log-verosimilitud es:

$$\begin{aligned}\ell(\alpha, \beta; \underline{S}) &= \log c(\underline{S}) + \log L(\alpha, \beta; \underline{S}) \\ &= c'(\underline{S}) + N \log \alpha - \beta \sum_{k=1}^N \log S_k - \int_0^{T_0} \alpha s^{-\beta} ds,\end{aligned}$$

pero  $\int_0^{T_0} \alpha s^{-\beta} ds = \frac{\alpha T_0^{1-\beta}}{1-\beta}$ , luego

$$\ell(\alpha, \beta; \underline{S}) = c'(\underline{S}) + N \log \alpha - \beta \sum_{k=1}^N \log S_k - \frac{\alpha T_0^{1-\beta}}{1-\beta}.$$

Ya que el modelo involucra dos parámetros, obtendremos por tanto un vector de dos ecuaciones a resolver. Por tanto el vector de la función Score nos queda:

$$sc_1(\alpha, \beta; \underline{S}) = \frac{\partial \ell(\alpha, \beta; \underline{S})}{\partial \alpha} = \frac{N}{\alpha} + \frac{T_0}{1-\beta};$$

$$sc_2(\alpha, \beta; \underline{S}) = \frac{\partial \ell(\alpha, \beta; \underline{S})}{\partial \beta} = - \sum_{k=1}^N \log S_k + \frac{\alpha T_0^{1-\beta} \log T_0}{1-\beta} - \frac{\alpha T_0^{1-\beta}}{(1-\beta)^2}.$$

Resolvemos el sistema de ecuaciones simultáneas  $sc_1(\alpha, \beta; \underline{S}) = 0$  y  $sc_2(\alpha, \beta; \underline{S}) = 0$ . Primeramente resolvemos la ecuación  $sc_1(\alpha, \beta; \underline{S}) = 0$  para el parámetro  $\alpha$  y obtenemos por tanto el estimador de máxima verosimilitud para  $\alpha$  en términos de  $\beta$  como

$$\hat{\alpha}(\beta) = \frac{(1-\beta)N}{T_0^{1-\beta}}.$$

Por otro lado, sustituyendo  $\alpha = \hat{\alpha}(\beta)$  en  $sc_2(\alpha, \beta; \underline{S}) = 0$ , tenemos

$$- \sum_{k=1}^N \log S_k + \left( \frac{(1-\beta)N}{T_0^{1-\beta}} \right) \frac{T_0^{1-\beta}}{1-\beta} \log T_0 - \left( \frac{(1-\beta)N}{T_0^{1-\beta}} \right) \frac{T_0^{1-\beta}}{(1-\beta)^2} = 0.$$

Simplificando y resolviendo esta ecuación para  $\beta$ , obtenemos el estimador de máxima verosimilitud para  $\beta$  como sigue:

$$\hat{\beta} = 1 - \left[ \frac{N}{\sum_{k=1}^N \log \left( \frac{T_0}{S_k} \right)} \right].$$

Ahora determinaremos la matriz de información de Fisher  $I(\alpha, \beta; \underline{S})$  definida en (2.6) que resulta una matriz dos por dos, entonces cada una de las entradas de la matriz son respectivamente,

$$\begin{aligned}
I_{11}(\alpha, \beta; \underline{S}) &= -\frac{\partial^2 \ell(\alpha, \beta; \underline{S})}{\partial \alpha^2} = \frac{N}{\alpha^2}; \\
I_{12}(\alpha, \beta; \underline{S}) &= -\frac{\partial^2 \ell(\alpha, \beta; \underline{S})}{\partial \alpha \partial \beta} = \frac{-T_0}{(1-\beta)^2}; \\
I_{21}(\alpha, \beta; \underline{S}) &= -\frac{\partial^2 \ell(\alpha, \beta; \underline{S})}{\partial \alpha \partial \beta} = T_0[1 - \log T_0],
\end{aligned}$$

y

$$I_{22}(\alpha, \beta; \underline{S}) = -\frac{\partial^2 \ell(\alpha, \beta; \underline{S})}{\partial \beta^2} = \frac{\alpha T_0^{1-\beta} (\log T_0)^2}{1-\beta} - \frac{\alpha T_0^{1-\beta} (1 + \log T_0)}{(1-\beta)^2} + \frac{2\alpha T_0^{1-\beta}}{(1-\beta)^3},$$

sustituyendo  $\hat{\alpha}$  y  $\hat{\beta}$  en  $I(\alpha, \beta; \underline{S})$  obtenemos la matriz de información observada, luego cada una de las entradas de  $I(\hat{\alpha}, \hat{\beta}; \underline{S})$  son,

$$\begin{aligned}
I_{11}(\hat{\alpha}, \hat{\beta}; \underline{S}) &= \frac{T_0^{2(1-\beta)}}{N(1-\beta)^2}; \\
I_{12}(\hat{\alpha}, \hat{\beta}; \underline{S}) &= \frac{-T_0 \sum_{k=1}^N \log \left( \frac{T_0}{S_k} \right)}{N}; \\
I_{21}(\hat{\alpha}, \hat{\beta}; \underline{S}) &= T_0[1 - \log T_0];
\end{aligned}$$

y

$$\begin{aligned}
I_{22}(\hat{\alpha}, \hat{\beta}; \underline{S}) &= \left[ \sqrt{N}(\log T_0) - \frac{2}{N} \sum_{k=1}^N \log \left( \frac{T_0}{S_k} \right) \right]^2 \\
&\quad - \sum_{k=1}^N \log \left( \frac{T_0}{S_k} \right) \left[ 1 + (2\sqrt{2} - 1) \log T_0 \right].
\end{aligned}$$

Para verificar que  $L(\alpha, \beta; \underline{S})$  tiene un máximo relativo en  $(\hat{\alpha}, \hat{\beta})$  es suficiente que la matriz  $I(\alpha, \beta; \underline{S})$  sea positiva definida, esto es, que su determinante sea mayor que cero.

$$\hat{I}_{11} = \frac{T_0^{2(1-\beta)}}{N(1-\beta)^2} > 0,$$

ya que  $\beta \in (0, 1)$ ;

$$\begin{aligned}
\hat{I}_{22} &= \left[ \sqrt{N}(\log T_0) - \frac{2}{N} \sum_{k=1}^N \log \left( \frac{T_0}{S_k} \right) \right]^2 \\
&\quad - \sum_{k=1}^N \log \left( \frac{T_0}{S_k} \right) \left[ 1 + (2\sqrt{2} - 1) \log T_0 \right] > 0;
\end{aligned}$$

pues el primer término de la expresión es mayor que el segundo; y por último veremos que

$$\hat{I} = \hat{I}_{11}\hat{I}_{22} - \hat{I}_{12}^2 > 0;$$

$$\begin{aligned} \hat{I} = & \frac{T_0^{2(1-\beta)}}{N(1-\beta)^2} \left( \left[ \sqrt{N}(\log T_0) - \frac{2}{N} \sum_{k=1}^N \log \left( \frac{T_0}{S_k} \right) \right]^2 \right. \\ & \left. - \sum_{k=1}^N \log \left( \frac{T_0}{S_k} \right) \left[ 1 + (2\sqrt{2} - 1) \log T_0 \right] \right) \\ & - \left( \frac{-T_0 \sum_{k=1}^N \log \left( \frac{T_0}{S_k} \right)}{N} \right)^2 > 0. \end{aligned}$$

Por lo tanto, la matriz  $I(\hat{\alpha}, \hat{\beta}; \underline{S})$  es positiva definida; luego  $L(\alpha, \beta; \underline{S})$  tiene un máximo relativo en  $(\hat{\alpha}, \hat{\beta})$ .

Ahora determinemos la función de verosimilitud relativa conjunta,

$$\begin{aligned} R(\alpha, \beta; \underline{S}) &= \frac{L(\alpha, \beta; \underline{S})}{L(\hat{\alpha}, \hat{\beta}; \underline{S})} = \frac{\prod_{k=1}^N (\alpha S_k^{-\beta}) \exp \left\{ - \int_0^{T_0} \alpha s^{-\beta} ds \right\}}{\prod_{k=1}^N (\hat{\alpha} S_k^{-\hat{\beta}}) \exp \left\{ - \int_0^{T_0} \hat{\alpha} s^{-\hat{\beta}} ds \right\}} \\ &= \prod_{k=1}^N \left[ \left( \frac{\alpha}{\hat{\alpha}} \right) S_k^{\hat{\beta}-\beta} \right] \exp \left\{ -\alpha \frac{T_0^{1-\beta}}{1-\beta} + \hat{\alpha} \frac{T_0^{1-\hat{\beta}}}{1-\hat{\beta}} \right\} \\ &= \left( \frac{\alpha}{\hat{\alpha}} \right)^N \left[ \prod_{k=1}^N S_k^{\hat{\beta}-\beta} \right] \exp \left\{ -\alpha \frac{T_0^{1-\beta}}{1-\beta} + \hat{\alpha} \frac{T_0^{1-\hat{\beta}}}{1-\hat{\beta}} \right\}. \end{aligned}$$

Consideremos que nuestro parámetro de interés es  $\beta$ , por tanto  $\alpha$  es el parámetro de estorbo, entonces la función de verosimilitud perfil para  $\beta$  nos queda:

$$L_p(\beta; \underline{S}) = L(\hat{\alpha}(\beta), \beta; \underline{S}) = \prod_{k=1}^N (\hat{\alpha}(\beta) S_k^{-\beta}) \exp \left\{ - \int_0^{T_0} \hat{\alpha}(\beta) s^{-\beta} ds \right\}.$$

Por otro lado la función de verosimilitud relativa maximizada o relativa perfil es:

$$\begin{aligned} R_p(\beta; \underline{S}) &= \left( \frac{\hat{\alpha}(\beta)}{\hat{\alpha}} \right)^N \left[ \prod_{k=1}^N S_k \right]^{\hat{\beta}-\beta} \exp \left\{ -\hat{\alpha}(\beta) \frac{T_0^{1-\beta}}{1-\beta} + \hat{\alpha} \frac{T_0^{1-\hat{\beta}}}{1-\hat{\beta}} \right\} \\ &= \left( \frac{(1-\beta)N}{\hat{\alpha} T_0^{1-\beta}} \right)^N \left[ \prod_{k=1}^N S_k \right]^{\hat{\beta}-\beta} \exp \left\{ - \left( \frac{(1-\beta)N}{T_0^{1-\beta}} \right) \frac{T_0^{1-\beta}}{1-\beta} + \hat{\alpha} \frac{T_0^{1-\hat{\beta}}}{1-\hat{\beta}} \right\} \\ &= \left( \frac{(1-\beta)N}{\hat{\alpha} T_0^{1-\beta}} \right)^N \left[ \prod_{k=1}^N S_k \right]^{\hat{\beta}-\beta} \exp \left\{ -N + \hat{\alpha} \frac{T_0^{1-\hat{\beta}}}{1-\hat{\beta}} \right\}. \end{aligned}$$

La función de verosimilitud perfil relativa es de suma importancia para llevar a cabo la estimación del parámetro de interés  $\beta$  en primera instancia, así como la estimación para el parámetro  $\alpha$  en su momento. Para luego realizar inferencia sobre los mismos a través del cálculo, graficación e interpretación de los intervalos de verosimilitud y de sus probabilidades de cobertura que se aproximan a los intervalos de confianza, de cada uno de dichos parámetros.

Como se puede ver, los cálculos realizados para la estimación de los parámetros vía verosimilitud para este ejemplo son algo engorrosos, por lo que para determinar específicamente el valor de los estimadores de máxima verosimilitud así como sus intervalos, cuando analíticamente no sea conveniente o sencillo realizarlo, es posible recurrir a dispositivos numéricos de aproximación por simulación. La pregunta que ahora toma relevancia es cómo convertir en material operativo el conjunto de métodos descritos, es precisamente la dirección que toma el siguiente capítulo.



## Capítulo 3

### Simulación de procesos de Poisson y validación de probabilidades de cobertura

Este capítulo contiene el material operativo de simulación de datos, basado en las teorías de probabilidad y estadística desarrolladas en los capítulos precedentes. Mediante métodos numéricos, realizaremos la inferencia estadística bajo la teoría de verosimilitud a procesos de Poisson homogéneos y no homogéneos, tanto de un parámetro como de dos.

Los objetivos de este capítulo son: primero, describir los algoritmos que nos ayudarán a simular procesos de Poisson tanto homogéneos como no homogéneos. Segundo, a la luz de los datos simulados implementar la inferencia estadística basada en verosimilitud desarrollada en la Sección 2.1 para cada caso a tratar. Finalmente, validar las propiedades frecuentistas del método de verosimilitud corroborando las probabilidades de cobertura real de los intervalos de verosimilitud determinadas en la Subsección 2.1.3.

El Capítulo 3 se compone de dos secciones; la primera inicia con la simulación de un proceso de Poisson homogéneo inspeccionado en un intervalo de tiempo fijo, para luego hacer la simulación de un proceso de Poisson inspeccionado en un intervalo de tiempo aleatorio, y con la ayuda de los datos simulados, procederemos a inferir al parámetro mediante verosimilitud. La segunda sección desarrolla el caso de un proceso de Poisson no homogéneo con dos parámetros. Obtendremos y graficaremos la función e intervalos de verosimilitud, y finalizaremos con la construcción de los intervalos de verosimilitud-confianza a partir de la función de verosimilitud relativa verificando su probabilidad de cobertura por medio de simulaciones repetidas del proceso.

#### 3.1 Generación de un proceso de Poisson homogéneo

Para simular un proceso de Poisson ya sea homogéneo o no homogéneo usaremos el algoritmo de la transformada inversa, el cual se basa en la siguiente proposición, Ross[20].

**Proposición 3.1** *Sea  $U$  una variable aleatoria uniforme en  $(0, 1)$ . Para cualquier función de distribución continua  $F$ , invertible, la variable aleatoria  $X$  definida como  $X = F^{-1}(U)$  tiene distribución  $F$ , donde  $F^{-1}$  denota la función inversa de  $F$ .*

**Demostración.** Sea  $F_X$  la función de distribución de  $X = F^{-1}(U)$ . Por lo que

$$F_X(x) = P[X \leq x] = P[F^{-1}(U) \leq x]. \quad (3.1)$$

Ahora como  $F$  es una función de distribución, se tiene que  $F(x)$  es una función monótona creciente de  $x$  y por lo tanto la desigualdad " $a \leq b$ " es equivalente a la desigualdad " $F(a) \leq F(b)$ ". Así, de (3.1) vemos que

$$F_X(x) = P[F(F^{-1}(U) \leq F(x))] = P[U \leq F(x)] = F(x).$$

Por lo que la variable aleatoria  $X = F^{-1}(U)$  tiene distribución  $F$ . ■

**Ejemplo 3.1** Sea  $X$  una variable aleatoria exponencial con parámetro  $\lambda$ , entonces su función de distribución está dada por  $F(x) = 1 - e^{-\lambda x}$ .

Si hacemos  $x = F^{-1}(u)$ , entonces  $u = F(x) = 1 - e^{-\lambda x}$ , o bien  $1 - u = e^{-\lambda x}$ . Al obtener los logaritmos,  $x = -1/\lambda \log(1 - u)$ . Por lo tanto, para generar una exponencial con parámetro  $\lambda$  (media  $1/\lambda$ ) generamos un número aleatorio  $U$  y hacemos  $X = F^{-1}(U) = -1/\lambda \log(1 - U)$ . Ya que  $1 - U$  es también una uniforme en  $(0, 1)$ , entonces  $-1/\lambda \log(1 - U)$  tiene la misma distribución que  $-1/\lambda \log(U)$ , de esta manera podemos hacer  $X = F^{-1}(U) = -1/\lambda \log(U)$ .

**Observación 3.2** Lo anterior también nos proporciona el algoritmo para generar una variable aleatoria Poisson. Recordemos que un proceso de Poisson con parámetro  $\lambda$ , surge cuando los tiempos entre eventos consecutivos  $T_k$  son exponenciales independientes con razón  $\lambda$ . Para tal proceso  $N(t)$ , el número de eventos hasta el instante  $t$ , se distribuye Poisson con media  $\lambda$ . Entonces el  $n$ -ésimo evento ocurrirá en el instante  $\sum_{k=1}^n T_k = S_n$ , de modo que el número de eventos hasta el instante  $t$  de acuerdo a (1.5) se expresa como,

$$N(t) = \max\{n : S_n = \sum_{k=1}^n T_k \leq t\}.$$

Es decir, el número de eventos hasta el instante  $t$  es igual a la máxima  $n$  para la cual el  $n$ -ésimo evento ha ocurrido hasta el instante  $t$ .

Con los resultados del Ejemplo 3.1 es posible generar  $N = N(t)$ , una variable aleatoria Poisson con media  $\lambda$ , produciendo números aleatorios  $U_1, U_2, \dots, U_n$  y haciendo:

$$\begin{aligned} N &= \max\{n : \sum_{k=1}^n -\frac{1}{\lambda} \log(U_k) \leq t\} = \max\{n : \sum_{k=1}^n \log(U_k) \geq -\lambda t\} \\ &= \max\{n : \log(U_1 U_2 \dots U_n) \geq -\lambda t\} = \max\{n : U_1 U_2 \dots U_n \geq e^{-\lambda t}\}. \end{aligned} \quad (3.2)$$

Por lo tanto, una variable aleatoria Poisson  $N$  con media  $\lambda$  se puede obtener al generar de manera sucesiva números aleatorios hasta que su producto sea menor que  $e^{-\lambda t}$ , o equivalentemente,  $N = \min\{n : U_1 U_2 \dots U_n < e^{-\lambda t}\} - 1$ , donde  $t$  es el

tiempo,  $N$  es el número de eventos que han ocurrido hasta el tiempo  $t$ . Además el algoritmo proporciona la sucesión de tiempos,  $0 < S_1 < S_2 < \dots < S_N$ , en los que ocurrieron cada uno de los  $N$  eventos.

En base al algoritmo descrito simularemos un proceso de Poisson homogéneo, y retomando la teoría desarrollada en la Sección 2.2 de inferencia en procesos de Poisson homogéneos, implementaremos los resultados obtenidos para los casos que ahí se trataron.

### 3.1.1 Primer caso: Observación continua en un intervalo de tiempo fijo.

#### Simulación del proceso de Poisson homogéneo

**Ejemplo 3.3** *En base al algoritmo descrito anteriormente, simulemos un proceso de Poisson homogéneo  $N$  con un valor verdadero del parámetro  $\lambda_0 = 5$ , en un intervalo de tiempo fijo  $[0, 50]$ .*

Para simular el proceso de Poisson homogéneo en un tiempo fijo, iniciamos proporcionando  $T_0 = 50$ , la longitud del intervalo de tiempo fijo, luego proporcionamos el valor verdadero con el que se simula el proceso de Poisson,  $\lambda_0 = 5$ . Se aplican los resultados del Ejemplo 3.1 para generar de esta manera los tiempos de ocurrencia de cada uno de los eventos, el proceso se detiene hasta que (3.2) no se cumpla. A la luz de los datos observados en la simulación contamos con una realización de  $N = 252$  eventos ocurridos en los tiempos  $0 < S_1 < S_2 < \dots < S_{252} \leq 50$ , que siguen una distribución de Poisson. Procederemos ahora a implementar la teoría en verosimilitud desarrollada en la Subsección 2.2.1 con la finalidad de estimar el parámetro  $\lambda$  involucrado en el proceso de Poisson.

#### Estimación puntual por verosimilitud

De las funciones de verosimilitud (2.7), log-verosimilitud (2.8), Score (2.9) que obtuvimos en la Subsección 2.2.1, y de acuerdo a (2.10), el estimador de máxima verosimilitud para un proceso Poisson homogéneo en tiempo fijo es,

$$\hat{\lambda} = \frac{N'}{T_0};$$

donde  $N'$  es el número de eventos aleatorios ocurridos hasta el tiempo fijo  $T_0 = 50$ , por tanto el valor del estimador de máxima verosimilitud de  $\lambda$  es:  $\hat{\lambda} = 252/50 = 5.04$ , que es muy cercano al valor verdadero  $\lambda_0 = 5$  utilizado para simular el proceso.

Hasta ahora se ha determinado el estimador de máxima verosimilitud a partir de la función de verosimilitud, sin embargo una estimación puntual no proporciona ninguna información sobre la precisión involucrada en la estimación. Por ello determinaremos un intervalo de valores plausibles del parámetro en base a la muestra observada en la simulación.

#### Intervalos de verosimilitud y de confianza

Con la función de verosimilitud dada por (2.7) y el valor del  $EMV$ ,  $\hat{\lambda} = 5.04$  proporcionado por los datos simulados, es posible determinar la función de verosimilitud relativa y en consecuencia los intervalos de verosimilitud para ciertos niveles de  $c$ .

El siguiente gráfico ilustra la función de verosimilitud relativa junto con los  $IV(c)$  con  $c = 0.20, 0.50$  y  $0.80$  del parámetro  $\lambda$ .

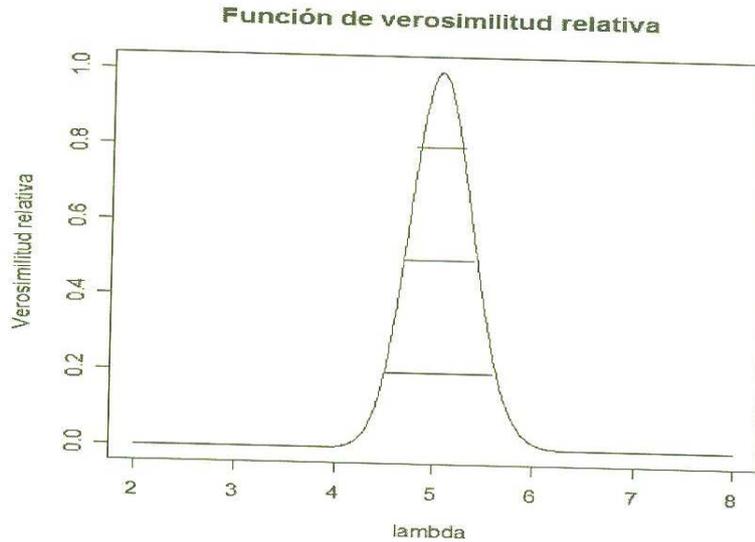


Fig. 3.1 Intervalos de verosimilitud para un proceso de Poisson en tiempo fijo

La forma de la función de verosimilitud es casi simétrica con respecto al estimador  $\hat{\lambda}$ ; observamos que para valores del parámetro menores de 4 y mayores de 6 resultan ser muy implausibles. Los límites de los intervalos de verosimilitud indicados en la Figura 3.1 para  $c = 0.20, 0.50$  y  $0.80$ , son  $(4.49, 5.59)$ ,  $(4.69, 5.39)$  y  $(4.79, 5.29)$  respectivamente.

Para construir ahora un intervalo de confianza, recordemos de la Subsección 2.2.1 que para este tipo de inspección del proceso hemos construido intervalos de Wald  $\hat{\lambda} \pm 1.96\sqrt{V(\hat{\lambda})}$ , requiriendo por tanto lo siguiente.

La varianza del estimador que está dada por.

$$V(\hat{\lambda}) = \frac{\lambda}{T_0},$$

es decir,  $V(\hat{\lambda}) = 5.04/50 = 0.101$ , por tanto el intervalo de Wald de acuerdo con (2.11) es:

$$5.04 \pm 1.96\sqrt{(0.101)} = (4.417, 5.663);$$

que es una aproximación al intervalo de confianza del 95% para  $\lambda$ . Este intervalo de Wald resulta ser muy similar al  $IV(c)$  con  $c = 0.147 \approx 0.15$ ,  $(4.443, 5.687)$ , que también es una aproximación al intervalo de confianza del 95% para  $\lambda$ . El intervalo de Wald, si observamos la forma de la función de verosimilitud relativa (Fig. 3.1)

en ese nivel  $c$ , se encuentra ligeramente trasladado hacia la izquierda con respecto al límite inferior del intervalo de verosimilitud. Esto significa que el intervalo de Wald por el extremo izquierdo está considerando valores implausibles de  $\lambda$ , por el extremo derecho está rechazando valores plausibles del parámetro.

Finalmente, para cubrir los objetivos descritos en la introducción de este capítulo, validaremos las propiedades frecuentistas del método de verosimilitud corroborando las probabilidades de cobertura real de los intervalos de verosimilitud determinadas en la Subsección 2.1.3. Para este caso lo haremos únicamente para el valor de  $c = 0.147$  relacionado con el intervalo de confianza con probabilidad de cobertura del 95%.

### Validación de la probabilidad de cobertura

Para validar la probabilidad de cobertura de los intervalos de verosimilitud, realizamos una serie de 5000 simulaciones de un proceso de Poisson homogéneo con valor verdadero  $\lambda_0 = 5$  en el intervalo de tiempo fijo  $[0, 50]$ , para cada simulación calculamos el  $IV(c)$  con  $c = 0.147$  y validamos que éste cubra al valor verdadero  $\lambda_0 = 5$ . De los 5000 intervalos obtenidos, 4730 cubren el valor verdadero, siendo la estimación de la probabilidad de cobertura de 0.946, que es muy próximo al 95%.

Como podemos observar los límites de los intervalos ya sea de Wald (4.417, 5.663) o el  $IV(c)$  con  $c = 0.147$  (4.443, 5.687) son similares. Sin embargo resulta conveniente utilizar los intervalos de verosimilitud-confianza, para evitar la situación como ya se dijo anteriormente, de rechazar valores plausibles por no ser cubiertos por el intervalo o de aceptar valores implausibles. Confirmando también que para el nivel de  $c = 0.147$  los límites del intervalo de verosimilitud siempre están cerca de la probabilidad de cobertura de 0.95.

### 3.1.2 Segundo caso: Observación continua en un intervalo de tiempo aleatorio.

Como segundo caso, realicemos ahora la inspección de observaciones continuas de un proceso de Poisson homogéneo en un intervalo de tiempo aleatorio. La diferencia con el primer caso radica en que ahora el número de eventos a ocurrir en un tiempo  $t$ ,  $N_0 = N(t)$  es fijo en un intervalo de tiempo aleatorio, esto es, simularemos un proceso hasta que ocurran  $N_0$  eventos Poisson, quedando de esta manera aleatorio el tiempo.

### Simulación del proceso de Poisson homogéneo

**Ejemplo 3.4** *En base al algoritmo descrito anteriormente, simularemos un proceso de Poisson homogéneo con razón  $\lambda_0 = 15$  en un intervalo de tiempo aleatorio, iniciando con  $t = 0$  hasta que ocurran  $N_0 = 30$  eventos.*

Para simular el proceso de Poisson homogéneo en un tiempo aleatorio, iniciamos proporcionando  $t = 0$ , luego proporcionamos el valor verdadero con el que se simula el proceso de Poisson,  $\lambda_0 = 15$  y el número de eventos que se van a generar  $N_0 = 30$ .

Nuevamente se aplican los resultados del Ejemplo 3.1 para generar de esta manera los tiempos de ocurrencia de cada uno de los eventos, el proceso se detiene hasta que se obtiene el evento número 30.

A la luz de los datos observados en la simulación contamos con una realización de  $N_0 = 30$  eventos, ocurridos en los tiempos  $0 < S_1 < S_2 < \dots < S_{30}$  que siguen una distribución de Poisson homogéneo en un intervalo de tiempo aleatorio. Procederemos ahora a implementar la teoría en verosimilitud desarrollada en la Subsección 2.2.2 con la finalidad de estimar el parámetro  $\lambda$  involucrado en el proceso de Poisson.

### Estimación puntual por verosimilitud

De las funciones de verosimilitud, log-verosimilitud y Score que obtuvimos, el estimador de máxima verosimilitud para un proceso Poisson homogéneo en tiempo aleatorio es,

$$\hat{\lambda} = \frac{N_0}{\bar{T}},$$

donde  $N_0 = 30$  es el tamaño de la muestra fija y  $\bar{T} = S_{30}$ . Por lo tanto el estimador de máxima verosimilitud obtenidos es  $\hat{\lambda} = 15.201$ , que es muy cercano al valor verdadero  $\lambda_0 = 15$ .

### Intervalos de verosimilitud y de confianza

Con la función de verosimilitud y el valor del *EMV*,  $\hat{\lambda} = 15.201$  proporcionado por los datos simulados, es posible determinar la función de verosimilitud relativa y en consecuencia los intervalos de verosimilitud para ciertos niveles de  $c$ . De los datos simulados, el siguiente gráfico ilustra la función de verosimilitud relativa junto con los intervalos de verosimilitud para  $c = 0.20, 0.50$  y  $0.80$  del parámetro  $\lambda$ .

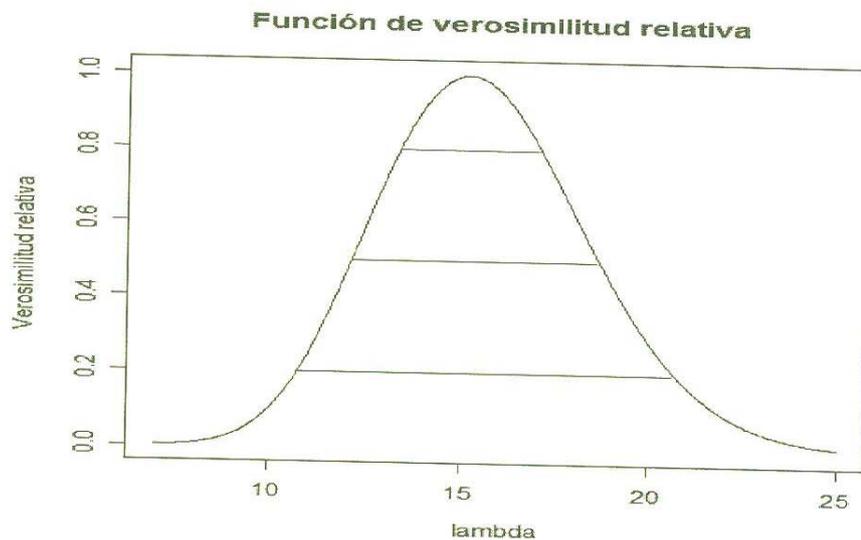


Fig. 3.2 Intervalos de verosimilitud para un proceso de Poisson homogéneo en tiempo aleatorio

Los límites de los intervalos en la Figura 3.2 son (10.750, 20.601), (12.201, 18.601) y (13.450, 17.101) respectivamente a los valores de  $c$  indicados, los cuales son de longitudes amplias debido a la curvatura amplia, no simétrica y sesgada ligeramente a la derecha que muestra la función de verosimilitud relativa con respecto al valor de  $\hat{\lambda}$ .

Procedemos ahora a construir un intervalo de confianza. Como ya se vio en la Subsección 2.2.2, el valor esperado para el estimador  $\hat{\lambda}$  del parámetro de un proceso de Poisson homogéneo inspeccionado de manera continua en un intervalo de tiempo aleatorio es sesgado, modificando el estimador de tal forma que contemple la característica de ser insesgado, el estimador fue

$$\tilde{\lambda} = (N_0 - 1)/\tilde{T},$$

y su varianza está expresada por

$$V(\tilde{\lambda}) = \lambda^2 / (N_0 - 2),$$

es decir,

$$V(\tilde{\lambda}) = (15.201)^2 / (30 - 2) = 8.252.$$

Siguiendo la misma analogía que en el ejemplo anterior, de acuerdo a la expresión (2.11), el intervalo de Wald correspondiente es

$$15.201 \pm 1.96\sqrt{(8.252)} = (9.571, 20.831);$$

que es una aproximación del intervalo de confianza del 95% para  $\lambda$ . El  $IV(c)$  con  $c = 0.147$  es (10.310, 21.131), que comparado con el de Wald vemos que se presenta la misma situación que en el caso anterior, el intervalo de Wald rechaza valores plausibles del parámetros y a la vez acepta valores implausibles. Lo cual hace decidirnos que los intervalos de verosimilitud-confianza son la mejor alternativa para realizar la inferencia estadística a los parámetros involucrados.

### Validación de la probabilidad de cobertura

Siguiendo el mismo procedimiento que el primer caso, para validar la probabilidad de cobertura de los intervalos de verosimilitud realizamos una serie de 5000 simulaciones de un proceso de Poisson homogéneo en tiempo aleatorio, para cada simulación calculamos el  $IV(c)$  con  $c = 0.147$  y validamos la probabilidad de cobertura del mismo. De los 5000 intervalos obtenidos, 4641 cubren el valor verdadero  $\lambda_0 = 15$ ; siendo la estimación de la probabilidad de cobertura de 0.928, que se aproxima al 95%.

## 3.2 Generación de un proceso de Poisson no homogéneo

Un proceso de conteo extremadamente importante para fines de modelación es el proceso de Poisson no homogéneo (PPNH), el cual relaja la hipótesis de incrementos estacionarios en el proceso de Poisson. Con ello permite que la tasa de ocurrencia de

los eventos no sea constante, sino que varíe con el tiempo, por ende la probabilidad de ocurrencia de los eventos no es la misma en intervalos de tiempo que no se traslapan. Si los eventos ocurren de manera aleatoria en el intervalo del tiempo y  $N(t)$  es el número de eventos que ocurren hasta el tiempo  $t$ ; una forma útil de interpretar un proceso de Poisson no homogéneo es la siguiente.

Supongamos que ciertos eventos  $N(t)$  ocurren de acuerdo con un proceso de Poisson con razón  $\lambda$ , supongamos también que independientemente de lo que haya pasado antes, un evento que ocurre en el instante  $t$  se cuenta con probabilidad  $p(t)$ . El proceso  $N'(t)$  de eventos contados constituye un proceso de Poisson no homogéneo con función de intensidad  $\lambda(t) = \lambda p(t)$ .

Simulamos entonces un proceso de Poisson homogéneo  $N(t)$ , con parámetro  $\lambda$ . Tenemos hasta un instante  $t$  una serie de eventos ocurridos en el tiempo, seleccionamos un intervalo  $(t, t+h]$  de tal manera que contemos los eventos ocurridos en el intervalo:

$$P[\text{un evento sea contado entre } t \text{ y } t+h] = P[\text{Ocurre un evento y sea contado}] \\ + P[\text{Ocurren dos o más eventos y exactamente uno sea contado}].$$

En términos de incrementos independientes la probabilidad anterior se expresa:

$$P[N'(t+h) - N'(t) = 1] = P[N'(t+h) - N'(t) = 1 \cap N(t+h) - N(t) = 1] \\ + P[N'(t+h) - N'(t) = 1 \cap N(t+h) - N(t) \geq 2];$$

las intersecciones las podemos expresar en términos de las probabilidades condicionales

$$P[N'(t+h) - N'(t) = 1] = \\ = P[N'(t+h) - N'(t) = 1 | N(t+h) - N(t) = 1] \cdot P[N(t+h) - N(t) = 1] \\ + P[N'(t+h) - N'(t) = 1 | N(t+h) - N(t) \geq 2] \cdot P[N(t+h) - N(t) \geq 2];$$

luego de (1.14) y (1.15) de la Condición 4<sup>o</sup>, tenemos:

$$P[N'(t+h) - N'(t) = 1] = p(t)\lambda h + p(t)o(h),$$

dividiendo por  $h$ ,

$$\frac{P[N'(t+h) - N'(t) = 1]}{h} = \lambda p(t) + p(t) \frac{o(h)}{h}.$$

En el límite, la parte izquierda de la igualdad representa la probabilidad de que un evento que ocurre en el instante  $t$  sea contado, es decir,

$$\lim_{h \rightarrow 0} \frac{P[N'(t+h) - N'(t) = 1]}{h} = \lambda(t);$$

de esta manera determinamos

$$\lambda(t) = \lambda p(t).$$

Por lo tanto la probabilidad con la que se cuenta un evento que ha ocurrido en el instante de tiempo  $t$ , está dada por  $p(t) = \lambda(t)/\lambda$ .

Con la ayuda de simulación generaremos las primeras  $T$  unidades de tiempo de un proceso de Poisson no homogéneo con función de intensidad  $\lambda(t)$ . El método que utilizaremos para generar el proceso comienza eligiendo un valor  $\lambda$  tal que

$$\lambda(t) \leq \lambda = \max_{s \leq t} \lambda(s) \quad \text{para toda } s \leq T.$$

Ahora, el proceso no homogéneo puede generarse mediante una selección aleatoria de los tiempos de evento de un proceso de Poisson con razón  $\lambda$ . Es decir, si se cuenta un evento de un proceso de Poisson con razón  $\lambda$  que ocurre en el instante  $t$  (de forma independiente a lo que ha ocurrido antes) con probabilidad  $\lambda(t)/\lambda$ , entonces el proceso de eventos contados es un proceso de Poisson no homogéneo con función de intensidad  $\lambda(t)$ ,  $0 \leq t \leq T$ . Por lo tanto, al simular un proceso de Poisson y luego contar de manera aleatoria sus eventos, generamos el proceso de Poisson no homogéneo.

El procedimiento anterior, conocido como algoritmo de adelgazamiento (pues reduce los puntos Poisson homogéneos) es más eficiente, en el sentido de tener el menor número de tiempos de eventos rechazados, cuando  $\lambda(t)$  está cerca de  $\lambda$  en todo el intervalo.

### 3.2.1 Primer caso: PPNH con función de intensidad exponencial

En base al procedimiento descrito para la obtención de un proceso de Poisson no homogéneo (PPNH) a partir de un proceso de Poisson homogéneo, procederemos a simular un PPNH con una función de intensidad exponencial con la finalidad de obtener una muestra observada. Para luego retomar los resultados desarrollados en el capítulo 2 referentes a verosimilitud con dos parámetros, estimar los parámetros, determinar intervalos de verosimilitud ilustrando gráficamente su función de verosimilitud relativa así como sus mapas de contorno. Finalmente validaremos las probabilidades de cobertura de los intervalos de verosimilitud de acuerdo a la teoría desarrollada en 2.1.3.

#### Simulación del proceso de Poisson no homogéneo.

**Ejemplo 3.5** Consideremos el intervalo de tiempo fijo  $(0, 1]$  donde se va a inspeccionar el proceso de Poisson no homogéneo  $\{N(t)\}_{t \geq 0}$  con función de intensidad  $\lambda(s; \theta, \mu) = e^{\theta + \mu s}$ ;  $\theta > 0$ ,  $\mu > 0$ . Sean  $0 < S_1 < S_2 < \dots < S_N \leq T_0$  los tiempos de ocurrencia, donde  $N$  es el número de eventos que ocurren en esos tiempos. Los tiempos entre eventos consecutivos son las variables aleatorias  $T_k = S_k - S_{k-1}$ ,  $k = 1, 2, \dots$  con función de densidad

$$f_{T_k}(t_k) = \lambda(t_k) \exp \left\{ - \int_{S_{k-1}}^{S_k} \lambda(s) ds \right\}.$$

Para simular el PPNH, iniciamos estableciendo  $T_0 = 1$ , la longitud del intervalo de tiempo fijo. Luego se proporcionan los valores verdaderos para cada uno de los

parámetros involucrados en la función de intensidad, esto es,  $\theta_0 = 3$  y  $\mu_0 = 2$ , enseguida se especifica la forma de la función de intensidad, que en nuestro caso es  $\lambda(s; \theta, \mu) = e^{\theta + \mu s}$ , y se establece el valor máximo alcanzado en la función, es decir,  $\lambda = \max \lambda(s)$  para toda  $s \leq T_0 = 1$ , por lo que  $\lambda = e^{3+2(1)} = 149$ . Se simula un proceso de Poisson homogéneo como lo hemos venido realizando, y cada uno de esos tiempos generados son seleccionados de manera aleatoria y contados con cierta probabilidad  $p(t)$ . La probabilidad como mencionamos se calcula  $\lambda(s; \theta, \mu)/\lambda$ , esto es, evaluamos los tiempos  $t$  obtenidos en el proceso de Poisson homogéneo en la función de intensidad y dividimos entre  $\lambda$ . De tal modo que el algoritmo arroja el número de eventos contados  $N$ , además de los tiempos  $0 < S_1 < S_2 < \dots < S_N$ , en los que fueron contados dichos eventos, que siguen una distribución de Poisson con función de intensidad  $\lambda(s; \theta, \mu)$ .

De los datos simulados del proceso obtuvimos los siguientes resultados, el número de eventos ocurridos en el intervalo de tiempo  $(0, 1]$  fueron  $N = 66$  que ocurrieron en los tiempos  $0 < S_1 < S_2 < \dots < S_{66} \leq 1$ . Procederemos ahora a implementar la teoría en verosimilitud desarrollada en la Sección 2.3 con la finalidad de estimar los parámetros  $\theta$  y  $\mu$  involucrados en la función de intensidad.

### Estimación puntual por verosimilitud

Primeramente haremos el desarrollo analítico para este caso, basándonos en la teoría vista en la Subsección 2.1.4. referente a inferencia en verosimilitud para el caso de dos parámetros. Para la muestra observada en la simulación del PPNH  $\underline{S} = (S_1, S_2, \dots, S_{66})$ , la función de verosimilitud de  $\theta$  y  $\mu$  está expresada por:

$$\begin{aligned} L(\theta, \mu; \underline{S}) &\propto \prod_{k=1}^N \left( \lambda(S_k; \theta, \mu) \cdot \exp \left\{ - \int_{S_{k-1}}^{S_k} \lambda(s; \theta, \mu) ds \right\} \right) \\ &= c(\underline{S}) \prod_{k=1}^N (\exp\{\theta + \mu S_k\}) \cdot \exp \left\{ - \int_0^1 e^{\theta + \mu t} dt \right\}; \end{aligned}$$

donde  $c(\underline{S})$  es una función constante de  $\underline{S}$  que no depende de los parámetros  $\theta$  y  $\mu$ . Considerando ahora el logaritmo, tenemos que la función log-verosimilitud es

$$\begin{aligned} \ell(\theta, \mu; \underline{S}) &= \log L(\theta, \mu; \underline{S}) \\ &= c'(\underline{S}) + N\theta + \mu \sum_{k=1}^N S_k - \int_0^1 e^{\theta + \mu t} dt, \end{aligned}$$

pero  $\int_0^1 e^{\theta + \mu t} dt = \frac{1}{\mu} e^{\theta} [e^{\mu} - 1]$ , luego

$$\ell(\theta, \mu; \underline{S}) = c'(\underline{S}) + N\theta + \mu \sum_{k=1}^N S_k - \frac{e^{\theta}}{\mu} [e^{\mu} - 1].$$

Determinaremos ahora el vector de la función Score,

$$sc_1(\theta, \mu; \underline{S}) = \frac{\partial \ell(\theta, \mu; \underline{S})}{\partial \theta} = N - \frac{e^{\theta}}{\mu} [e^{\mu} - 1];$$

$$sc_2(\theta, \mu; \underline{S}) = \frac{\partial \ell(\theta, \mu; \underline{S})}{\partial \mu} = \sum_{k=1}^N S_k - \frac{e^\theta}{\mu} [2e^\mu - 1].$$

Resolvemos el sistema de ecuaciones simultáneas  $sc_1(\theta, \mu; \underline{S}) = 0$  y  $sc_2(\theta, \mu; \underline{S}) = 0$ . Primeramente resolvemos la ecuación  $sc_1(\theta, \mu; \underline{S}) = 0$  para el parámetro  $\theta$  y obtenemos por tanto el estimador de máxima verosimilitud para  $\theta$  en términos de  $\mu$  como

$$\hat{\theta}(\mu) = \log \left( \frac{\mu N}{e^\mu - 1} \right). \quad (3.3)$$

Por otro lado, sustituyendo  $\theta = \hat{\theta}(\mu)$  en  $sc_2(\theta, \mu; \underline{S}) = 0$ , tenemos

$$\sum_{k=1}^N S_k - \frac{e^{\log(\frac{\mu N}{e^\mu - 1})}}{\mu} [2e^\mu - 1] = 0.$$

Simplificando y resolviendo esta ecuación para  $\mu$ , obtenemos el estimador de máxima verosimilitud para  $\mu$  como

$$\hat{\mu} = \log \left( \frac{N - \sum_{k=1}^N S_k}{2N - \sum_{k=1}^N S_k} \right). \quad (3.4)$$

Hasta aquí hemos determinado explícitamente una expresión de los estimadores para cada uno de los parámetros, donde  $N = 66$ , es el número de eventos contados hasta un tiempo  $t$  y  $S_k$  son los tiempos de ocurrencia de dichos eventos. Ahora con los datos simulados y de manera numérica alimentamos las expresiones (3.3) y (3.4) con la finalidad de determinar el valor específico de cada estimador.

El programa nos calculó los estimadores de máxima verosimilitud  $\hat{\theta} = 3.025$  y  $\hat{\mu} = 2.005$  que son muy próximos a los valores verdaderos con los que fue simulado el PPNH,  $\theta_0 = 3$  y  $\mu_0 = 2$ .

### Función de verosimilitud con datos simulados

Con la expresión de la función de verosimilitud y los valores del EMV,  $\hat{\theta}$  y  $\hat{\mu}$  proporcionados por los datos simulados, es posible determinar la forma que toma la función de verosimilitud relativa  $R(\theta, \mu; \underline{S})$  que en esta ocasión será una superficie de verosimilitud. El valor máximo 1 de  $R(\theta, \mu; \underline{S})$  ocurre en  $(3.025, 2.005)$ , es decir, en el plano  $(\theta, \mu)$  el punto  $(3.025, 2.005)$  hace que  $R(\theta, \mu; \underline{S}) = 1$ .

La Figura 3.3, nos ilustra para los tiempos de ocurrencia  $\underline{S} = (S_1, S_2, \dots, S_{66})$  la superficie de verosimilitud  $R(\theta, \mu; \underline{S})$  del proceso de Poisson no homogéneo con función de intensidad  $\lambda(\underline{S}; \theta, \mu) = e^{\theta + \mu s}$  simulado en el intervalo de tiempo fijo  $(0, 1]$  con valores verdaderos  $\theta = 3$ ,  $\mu = 2$ .

## SUPERFICIE DE VEROSIMILITUD

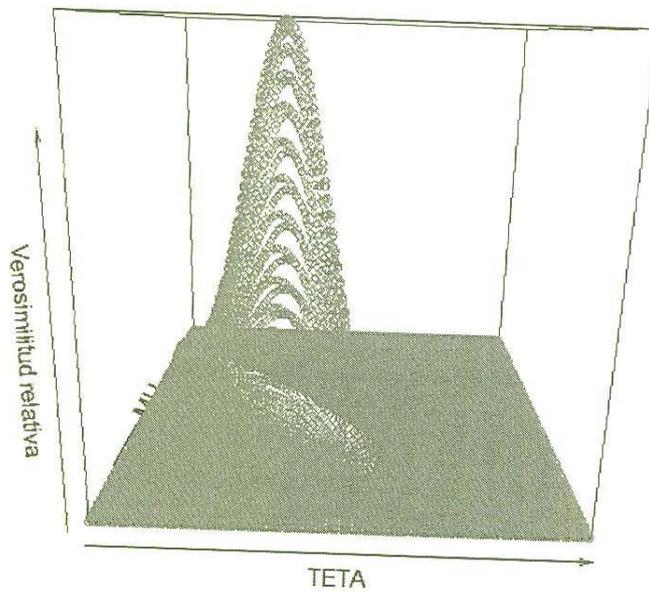


Fig. 3.3 Superficie de verosimilitud relativa de un proceso con dos parámetros.

En un modelo que involucra varios parámetros también se puede hablar de mapas de contorno. Recordemos que la región de verosimilitud  $CV(c)$ , es el conjunto de valores paramétricos  $(\theta, \mu)$  tales que  $R(\theta, \mu; \underline{S}) \geq c$  y la curva  $R(\theta, \mu; \underline{S}) = c$ , la cual forma la frontera de esa región de verosimilitud es llamada *contorno de verosimilitud*.

Usualmente los contornos formarán un conjunto anidado de curvas de forma elípticas alrededor del punto  $(\hat{\theta}, \hat{\mu})$ , para ver en que nivel del plano  $(\theta, \mu)$  se encuentra la pareja de valores estimados  $(\hat{\theta}, \hat{\mu}) = (3.025, 2.005)$  el siguiente gráfico nos será de mayor utilidad.

La Figura 3.4 nos muestra para los diferentes niveles de  $c = 0.20, 0.50$  y  $0.80$ , los contornos de verosimilitud.

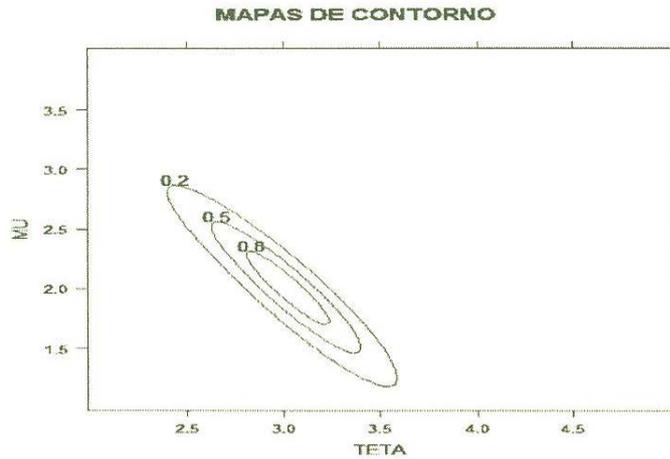
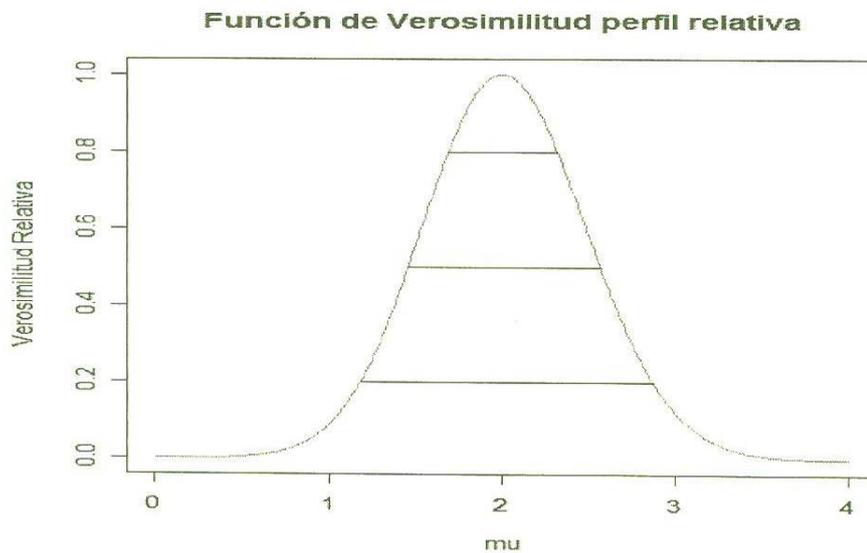


Fig. 3.4 Contornos de verosimilitud.

La pareja estimada  $(\hat{\theta}, \hat{\mu}) = (3.025, 2.005)$  se localiza en el centro de de la curva  $R(\theta, \mu; \underline{S}) = 0.8$ , es decir, dentro del  $CV(c)$  con  $c = 0.8$ .

#### Función de verosimilitud perfil con datos simulados

Considerando a  $\mu$  como parámetro de interés, de los datos simulados, la Figura 3.5 nos muestra la función de verosimilitud perfil de  $\mu$ , de un proceso Poisson no homogéneo simulado en el intervalo de tiempo fijo  $(0, 1]$ , con un valor verdadero  $\mu = 2$ . Cabe mencionar que la corrida del proceso de Poisson no homogéneo simulado se fijó para realizar la gráfica de la función de verosimilitud perfil, y en consecuencia los intervalos de verosimilitud para ciertos niveles de  $c$ .

Fig. 3.5 Función de verosimilitud perfil relativa para  $\mu$ .

De acuerdo al gráfico se aprecia la función de verosimilitud perfil de  $\mu$  centrada en el valor del estimador de máxima verosimilitud  $\hat{\mu} = 2.005$ , y los intervalos de verosimilitud para los niveles de  $c = 0.20, 0.50$  y  $0.80$ , son  $(1.183, 2.869)$ ,  $(1.462, 2.567)$  y  $(1.695, 2.321)$  respectivamente.

### Validación de la probabilidad de cobertura

Para validar la probabilidad de cobertura de los intervalos de verosimilitud, realizaremos una serie de 5000 simulaciones de un proceso de un PPNH en el intervalo de tiempo fijo  $[0, 1]$ , para cada simulación calcularemos el  $IV(c)$  con  $c = 0.147$  para el parámetro de interés  $\mu$  y validamos que éste cubra al valor verdadero  $\mu_0 = 5$ .

Para realizar los cálculos de las probabilidades de cobertura de los intervalos de verosimilitud en el caso de dos parámetros, nos centraremos en el resultado obtenido en la Subsección 2.1.3, donde el estadístico de razón de verosimilitud  $\Lambda = -2 \log R(\mu; \underline{S})$  converge en distribución a una variable aleatoria Ji-cuadrada con un grado de libertad  $\chi_{(1)}^2$ .

El experimento consiste en repetir un número grande de veces, los siguientes pasos:

- (i) Simulación de un proceso de Poisson no homogéneo con parámetros  $(\theta, \mu)$ .
- (ii) Considerando  $\theta$  como parámetro de estorbo, proceder a construir los intervalos de verosimilitud. para  $\mu$ ,
- (iii) Examinar cada intervalo de verosimilitud calculado para determinar por comparación con el valor verdadero de  $\mu$  si se obtuvo cobertura o no.

El objetivo es comparar el valor nominal de cobertura de los intervalos de verosimilitud, con el valor de cobertura real, que como ya discutimos en el capítulo previo se aproxima a los intervalos de confianza para ciertos valores de  $c$ .

El resultado de iterar 5000 veces el experimento anterior para un nivel de  $c = 0.147$  es el siguiente: de las 5000 repeticiones que se simularon 4724 cubren el valor verdadero  $\mu_0 = 2$ , con una probabilidad de 0.944 que se aproxima al 95%. Efectivamente la aproximación es muy buena.

### 3.2.2 Segundo caso: PPNH con función de intensidad gama

El siguiente ejemplo es de suma importancia para el desarrollo del siguiente capítulo, ya que con el pretendemos contrastar las diferentes estimaciones de los intervalos de verosimilitud y de confianza seleccionados en 2.1.3 y que utilizaremos para el caso de estudio. Para tal efecto utilizaremos en este caso  $IV(c)$  con niveles de  $c = 0.258, 0.147$  y  $0.036$  por tener una probabilidad de cobertura aproximada del 90%, 95% y 99%, respectivamente. Se resalta aquí que el cálculo computacional de la verosimilitud perfil de cada uno de los parámetros involucrados, las estimaciones a los parámetros y sus respectivos intervalos de verosimilitud es tarea fácil y rápida, no así, el cálculo de las probabilidades de cobertura que requirieron de mayor tiempo.

### Simulación del proceso de Poisson no homogéneo.

**Ejemplo 3.6** Consideremos el intervalo de tiempo fijo  $(0, 365]$  donde se va a inspeccionar el proceso de Poisson no homogéneo  $\{N(t)\}_{t \geq 0}$  con función de intensidad  $\lambda(s) = 1/\Gamma(\alpha)\beta^\alpha s^{\alpha-1} \exp(-s/\beta)$ ,  $\alpha, \beta > 0$ . Sean  $0 < S_1 < S_2 < \dots < S_N \leq T_0$  los tiempos de ocurrencia, donde  $N$  es el número de eventos que ocurren en esos tiempos. Los tiempos entre eventos consecutivos son las variables aleatorias  $T_k = S_k - S_{k-1}$ ,  $k = 1, 2, \dots, N$  con función de densidad

$$f_{T_k}(t_k) = \lambda(t_k) \exp \left\{ - \int_{S_{k-1}}^{S_k} \lambda(s) ds \right\}.$$

Llevamos a cabo la simulación del proceso de Poisson no homogéneo con valores verdaderos  $\alpha_0 = 4$  y  $\beta_0 = 12$ , de la misma manera que el primer caso de esta sección. De los datos simulados del proceso obtuvimos los siguientes resultados, el número de eventos ocurridos en el intervalo de tiempo  $(0, 365]$  fueron  $N = 15$  que ocurrieron en los tiempos  $0 < S_1 < S_2 < \dots < S_{15} \leq 365$ . Procederemos ahora a implementar la teoría en verosimilitud desarrollada en la Sección 2.3 con la finalidad de estimar los parámetros  $\alpha$  y  $\beta$  involucrados en la función de intensidad  $\lambda(s)$ .

#### Estimación puntual por verosimilitud

Para  $\underline{S} = (S_1, S_2, \dots, S_{15})$ , la función de verosimilitud de  $\alpha$  y  $\beta$  está expresada por:

$$\begin{aligned} L(\alpha, \beta; \underline{S}) &\propto \prod_{k=1}^N \left( \lambda(S_k; \alpha, \beta) \cdot \exp \left\{ - \int_{s_{k-1}}^{s_k} \lambda(s; \alpha, \beta) ds \right\} \right) \\ &= c(\underline{S}) \prod_{k=1}^N \left( \frac{1}{\Gamma(\alpha)\beta^\alpha} S_k^{\alpha-1} \exp(-S_k/\beta) \right) \cdot \exp \left\{ - \frac{1}{\Gamma(\alpha)\beta^\alpha} \int_0^{T_0} s^{\alpha-1} e^{-s/\beta} ds \right\}, \end{aligned}$$

donde  $c(\underline{S})$  es una función constante de  $\underline{S}$  que no depende de los parámetros  $\alpha$  y  $\beta$ . Considerando ahora el logaritmo, tenemos que la función log-verosimilitud es

$$\begin{aligned} \ell(\alpha, \beta; \underline{S}) &= \log L(\alpha, \beta; \underline{S}); \\ &= c'(\underline{S}) + \sum_{k=1}^N \log \left( \frac{1}{\Gamma(\alpha)\beta^\alpha} S_k^{\alpha-1} \exp(-S_k/\beta) \right) \\ &\quad - \frac{1}{\Gamma(\alpha)\beta^\alpha} \int_0^{T_0} s^{\alpha-1} \exp(-s/\beta) ds, \end{aligned}$$

ya que analíticamente la expresión de la función de verosimilitud resulta ser complicada, procederemos a obtener la estimación de verosimilitud para los parámetros  $\alpha$  y  $\beta$  de manera numérica. De los datos simulados del proceso de Poisson no

homogéneo con función de intensidad  $\lambda(s; \alpha, \beta) = 1/\Gamma(\alpha)\beta^\alpha s^{\alpha-1} \exp(-s/\beta)$ , las estimaciones obtenidas son  $\hat{\alpha} = 5.08$  y  $\hat{\beta} = 11.94$  que son muy cercanos a los valores verdaderos.

### Función de verosimilitud perfil con datos simulados

Considerando ahora a  $\alpha$  como parámetro de interés, la Figura 3.6 nos muestra la función de verosimilitud perfil relativa a  $\alpha$ , los respectivos intervalos de verosimilitud para los distintos niveles de  $c$  y la ubicación del *EMV*  $\hat{\alpha} = 5.08$ , de un proceso de Poisson no homogéneo con función de intensidad gama simulado en un intervalo de tiempo fijo  $(0, 365]$ , con un valor verdadero  $\alpha_0 = 4$ . Es importante comentar aquí, que los datos obtenidos de la corrida del proceso de Poisson no homogéneo simulado se fijaron para graficar las funciones de verosimilitud perfiles para cada uno de los parámetros.

Se observa una función de verosimilitud perfil ligeramente asimétrica con cola hacia la derecha respecto al valor del estimador  $\hat{\alpha} = 5.08$ . Además valores del parámetro  $\alpha$  menores de 1 y mayores de 12 son muy implausibles.

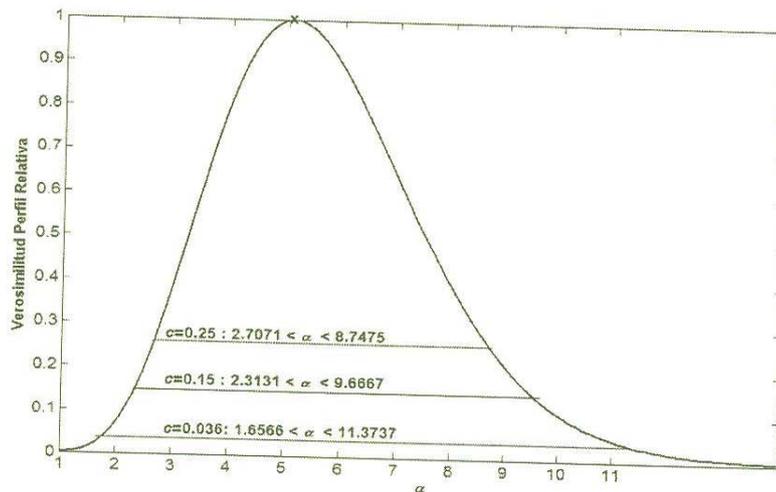
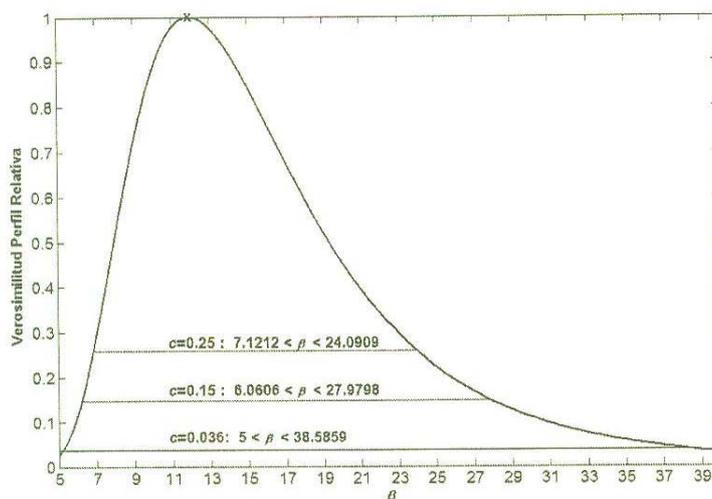


Fig. 3.6 Función de verosimilitud perfil de  $\alpha$ .

Considerando ahora a  $\beta$  como parámetro de interés y a  $\alpha$  como parámetro de estorbo, la Figura 3.7 nos muestran la función de verosimilitud perfil relativa a  $\beta$  y los intervalos de verosimilitud para los distintos niveles de  $c$ . Se observa una clara asimetría con cola pesada hacia la derecha con respecto a  $\hat{\beta} = 11.94$ .

Fig. 3.7 Función de verosimilitud perfil de  $\beta$ .

### Validación de la probabilidad de cobertura

Ahora ejecutando el programa que nos determina la probabilidad de cobertura de los  $IV(p)$  con  $c = 0.147, 0.258$  y  $0.036$ , obtuvimos los siguientes resultados, se simularon 5,000 procesos de Poisson no homogéneo con valores verdaderos  $\alpha_0 = 4$  y  $\beta_0 = 12$ , considerando  $\alpha$  como parámetro de interés, se calcularon 5,000 intervalos de verosimilitud para  $\alpha$  de tal manera que  $N$  contabiliza el número de intervalos que cubren el valor verdadero  $\alpha_0 = 4$ , con una probabilidad de cobertura  $PC$ . Los resultados se presentan en la siguiente tabla:

Verosim. relativa	Nivel de confianza	N	PC( $\alpha$ )
$c=0.25$	0.90	4431	0.886
$c=0.15$	0.95	4713	0.942
$c=0.036$	0.99	4916	0.983

De igual manera, se realizó el mismo procedimiento para  $\beta$  ahora considerado como parámetro de interés, se calcularon 5,000 intervalos de verosimilitud para  $\beta$  y se contabilizaron cuántos de ellos cubrieron el valor verdadero  $\beta_0 = 12$ .

Verosim. relativa	Nivel de confianza	N	PC( $\beta$ )
$c=0.25$	0.90	4366	0.873
$c=0.15$	0.95	4664	0.932
$c=0.036$	0.99	4797	0.979

Si observamos los resultados obtenidos para las probabilidades de cobertura de los intervalos de verosimilitud, notamos que efectivamente éstas se aproximan a los niveles de confianza específicos. Validando de esta manera que para los niveles de  $c$  seleccionados los intervalos de verosimilitud son ahora intervalos de verosimilitud-confianza

En la Subsección 2.1.3 mencionamos que los intervalos de confianza no dependen del valor verdadero  $\theta_0$ , por lo que la probabilidad de cobertura es la misma para todos los valores que  $\theta_0$  pueda tomar. Para validar esta conjetura, realizamos cambios a los valores verdaderos de  $\alpha_0$  y  $\beta_0$  con la finalidad de obtener las probabilidades de cobertura del  $IV(c)$  con  $c = 0.147$  para cada uno de los parámetros. Dentro del mismo algoritmo que hemos venido utilizando para validar las probabilidades de cobertura en este segundo caso, hemos considerado ahora los valores verdaderos  $\alpha_0 = 12$  y  $\beta_0 = 4$ , resultando:

	Verosim. relativa	Nivel de confianza	N	PC
$\alpha_0 = 12$	$c=0.15$	0.95	4698	0.939
$\beta_0 = 4$	$c=0.15$	0.95	4733	0.946

Lo anterior ratifica la hipótesis antes mencionada, la probabilidad de cobertura de un intervalo de confianza no depende del valor verdadero que tome el parámetro.

## Capítulo 4

### Caso de estudio

En este capítulo aplicaremos al caso de estudio las teorías de Probabilidad y Estadística y los resultados de los capítulos anteriores, en problemas de inferencia estadística. Usualmente empezamos con un conjunto de datos, y con alguna información acerca de la manera en que éstos datos fueron recolectados. Luego caracterizamos un modelo probabilístico para el experimento el cual es generado por los datos; comúnmente el modelo de probabilidad involucrará uno o más parámetros desconocidos el o los cuales deberán ser estimados a partir de los datos.

Para tal efecto, consideremos aquí las temperaturas registradas en la ciudad de Guanajuato. Se distinguen tres tipos de clima: semiseco, templado y semicálido. El clima semiseco, se le denomina también seco estepario y se caracteriza porque en él la evaporación excede a la precipitación. El clima templado es en general intermedio en cuanto a temperatura (mesotérmico); el clima semicálido es el más cálido de los tres tipos básicos presentes en el estado, sus rangos de temperatura oscilan entre los 18 y los 22 grados Celsius, Figura 4.1.

#### Fuente de Información

Los eventos observados para este caso de estudio, están integrados en una base de datos en Excel; éstos refieren las mediciones de las temperaturas durante los meses de diciembre del año 2005 a diciembre del año 2006, registradas por la estación meteorológica ubicada en el Centro de Investigación de Matemáticas (CIMAT) en el municipio de Guanajuato, Gto.

Los registros muestran las mediciones de cada media hora aproximadamente de: temperaturas tanto interior como exterior en grados Celsius, dirección, velocidad y rachas del viento en millas por hora, porcentaje de humedad interior como exterior, presión barométrica y lluvia en pulgadas, así como presión barométrica al nivel del mar, y otros. Presentaremos la información por estadios: primavera, verano, otoño e invierno. Para ello, con el apoyo del programa “R”, hemos obtenido las temperaturas máximas diarias durante los diferentes estadios del año, cada uno de ellos con una duración de 92 días.

#### Análisis de datos

Iniciaremos el planteamiento de este caso definiendo la variable de estudio que nos interesa, diremos que ocurre una “*racha de calor*” si se presentan temperaturas máximas diarias que exceden los  $31^{\circ}C$ , tal como lo muestra la Figura 4.2,

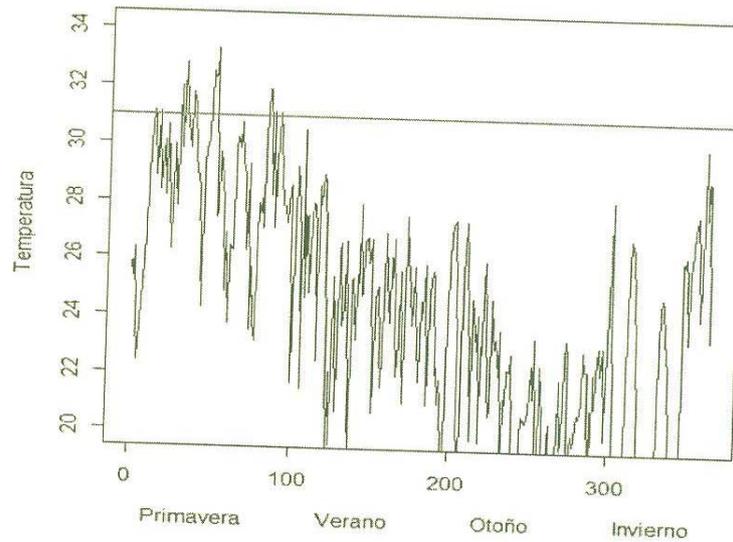


Fig. 4.2 Presencia de rachas de calor.

Los objetivos en este caso de estudio son dos: modelar la distribución que siguen las rachas de calor como un proceso de Poisson no homogéneo, y determinar la función de intensidad del modelo de acuerdo a las tendencias que las rachas de calor presentan durante el año. Para ello es necesario el siguiente gráfico que nos muestra la información por estadios.

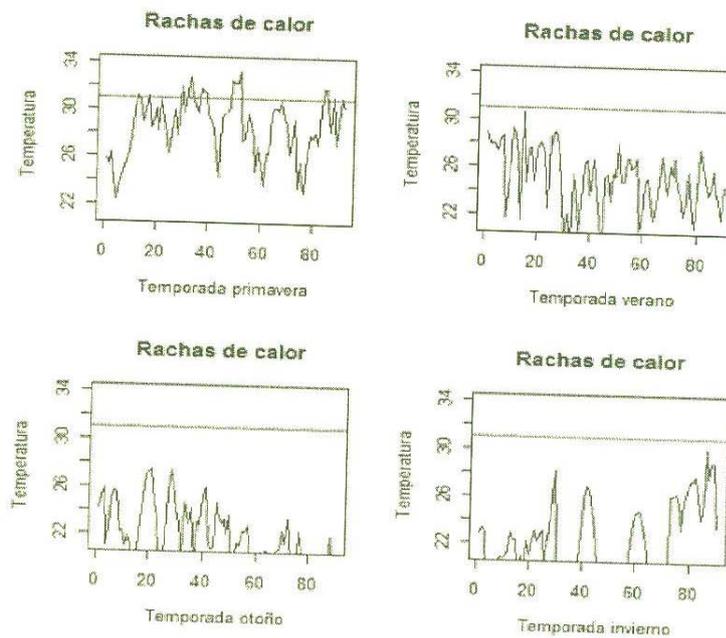


Fig. 4.3 Rachas de calor en los diferentes estadios del año.

Como podemos apreciar en la Figura 4.3 las rachas de calor se hacen presentes en el estadio de primavera, mientras que en el resto del año es nula la presencia de las mismas. Los datos observados fueron registrados de manera continua en un intervalo de tiempo fijo  $[0, 365]$ ; en el conjunto de observaciones, denotaremos por  $N(t)$  al proceso que cuenta el número de rachas de calor presentadas durante el tiempo  $t$ . El siguiente vector reporta los días en que se presentaron temperaturas máximas mayores de  $31^\circ \text{C}$  en los 365 días considerados,

$$(13, 17, 30, 32, 33, 34, 37, 38, 39, 49, 50, 51, 52, 84, 85, 88, 91).$$

Notemos que hay rachas de calor que duran sólo un día como lo reflejan las tres primeras entradas del vector anterior, que representan los tiempos de ocurrencia de las tres primeras rachas de calor, la cuarta racha inicia en el día 32 teniendo ésta una duración de tres días, luego la quinta racha inicia el día 37 durando también tres días; la racha de calor más larga es precisamente la sexta racha con una duración de cuatro días y así sucesivamente.

Denotemos por  $\underline{S}$  al vector que reporta los tiempos donde inicia una racha de calor

$$\underline{S} = (13, 17, 30, 32, 37, 49, 84, 88, 91);$$

además  $N(365) = 9$ , es decir, se presentaron 9 rachas de calor hasta el tiempo 365. Así mismo los tiempos entre rachas consecutivas  $T_k = S_k - S_{k-1}$ ,  $k = 1, 2, \dots, n$ , están expresados por el vector:  $T_k = (13, 4, 13, 2, 5, 12, 35, 4, 3)$ .

Por el análisis exploratorio visual de los datos de la Figura 4.3, vemos que la presencia de las rachas de calor en su totalidad se presentan en el estadio de primavera, esto es, en el primer cuarto del año. Si analizamos con detalle el estadio de primavera, la presencia de rachas de calor se hace notar con mayor frecuencia durante la primera mitad de la temporada, mientras que disminuye de manera paulatina durante la segunda mitad, siendo nula la presencia de las mismas conforme termina el año, razón por la que el proceso de conteo es no homogéneo, es decir, que el número promedio de rachas de calor esperadas está en función del tiempo. Ésto lo cuantifica la función de intensidad del proceso.

Una distribución aceptable al comportamiento de la intensidad de las rachas de calor sería una curva creciente hasta alcanzar un punto máximo alrededor de la primera mitad del estadio, decreciente hasta alcanzar el cero dentro de la segunda mitad del estadio y estacionarse en 0 en el resto del año; por tanto propondremos como función de intensidad  $\lambda(t)$  del modelo estadístico:

$$\lambda(t) = \frac{\delta}{\Gamma(\alpha)\beta^\alpha} t^{\alpha-1} \exp(-t/\beta), \quad \alpha, \beta, \delta > 0; \quad (4.1)$$

donde  $\alpha$ ,  $\beta$  y  $\delta$  son parámetros desconocidos. Bien pudimos haber considerado otras caracterizaciones para la función de intensidad, que también cumplen con las características descritas por el comportamiento de los datos, pero elegimos la gama por las bondades que presenta esta distribución bajo el contexto que nos interesa.

### Reescalamiento de los tiempos

Tenemos hasta ahora un proceso que cuenta las rachas de calor que se distribuyen de manera no homogénea. Enseguida consideraremos una transformación a los tiempos de ocurrencia  $\underline{S}$  del proceso de conteo no homogéneo con la finalidad de reescalar los tiempos, es decir, aplicaremos al vector  $\underline{S}$  la transformación  $m(t) = \int_0^t \lambda(s) ds$ , donde  $\lambda(s)$  es la función de intensidad gama del proceso dada por (4.1). De esta manera obtenemos un nuevo vector de tiempos de ocurrencia  $\underline{S}'$  dado por:

$$\underline{S}' = (0.5649, 0.9813, 2.7235, 3.0095, 3.7139, 5.2439, 7.8975, 8.0579, 8.1634).$$

De la misma manera obtenemos ahora el nuevo vector de tiempos entre rachas consecutivas

$$T'_k = (0.5649, 0.4164, 1.7422, 0.2860, 0.7044, 1.5300, 2.6536, 0.1604, 0.1055).$$

Analizando el comportamiento que presentan los tiempos entre rachas consecutivas  $T'_k$  la Figura 4.4 nos proporciona el histograma de dichos tiempos,

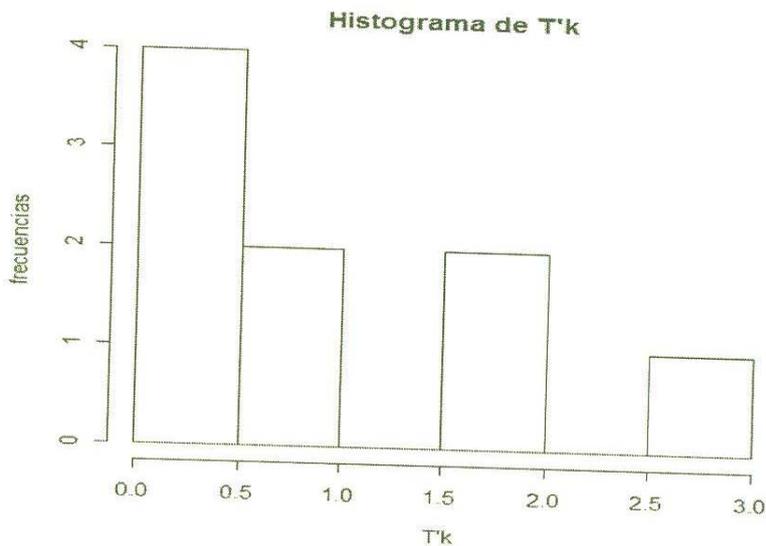


Fig. 4.4 Histograma de los tiempos  $T'_k$ .

por la manera en que decrecen las frecuencias, pareciera que los tiempos  $T'_k$  muestran una tendencia a distribuirse de manera exponencial. Con la finalidad de verificar que éstos tiempos entre eventos consecutivos se distribuyen de manera exponencial utilizaremos un gráfico de probabilidad qq plot Sidney[?], un método que nos permite comparar la función de distribución empírica de los datos con una función de distribución teórica específica. El gráfico qq plot consiste en lo siguiente: en el punto  $x$ , la función de distribución empírica toma por valor la proporción de los datos que son inferiores a  $x$ . Si los datos se simulan a partir de una distribución teórica específica, esta proporción debería estar cercana al valor correspondiente de

la función de distribución teórica. Así que una vez observados los tiempos entre rachas de calor consecutivas,  $T'_1, T'_2, \dots, T'_9$ , ordenamos las observaciones de menor a mayor. Luego determinamos los valores  $p_i = i - 0.5/n$ ;  $i = 1, 2, \dots, 9$ , por  $Q_T(p)$  denotamos al cuantil de orden  $p$ , ( $0 < p < 1$ ) de las observaciones, tenemos entonces  $T'_i = Q_T(p_i)$ ;  $i = 1, 2, \dots, 9$ . Determinamos ahora bajo el algoritmo de la función inversa visto en el Ejemplo 3.1 los cuantiles de orden  $p_i$ ,  $i = 1, 2, \dots, 9$ . De la distribución teórica representada por la función de distribución  $F$ , que para nuestro caso es una exponencial con parámetro  $\lambda = 1$ , es decir:  $Q_t(p_i) = F^{-1}(p_i)$ ;  $i = 1, 2, \dots, 9$ . Representamos el conjunto  $(Q_t(p_i), Q_{T'}(p_i))$ ,  $i = 1, 2, \dots, 9$ , o lo que es lo mismo los puntos  $(F^{-1}(p_i), T'_i)$ . Si la distribución teórica constituye una buena aproximación de la distribución empírica, cabría esperar que los cuantiles de los datos están muy próximos a la distribución teórica. Gráficamente esto se puede ver con un distribución de los puntos muy cercanos a la recta  $y = x$ .

Analicemos entonces por medio de un gráfico qq plot (Fig. 4.5), la distribución que presentan los tiempos entre rachas consecutivas.

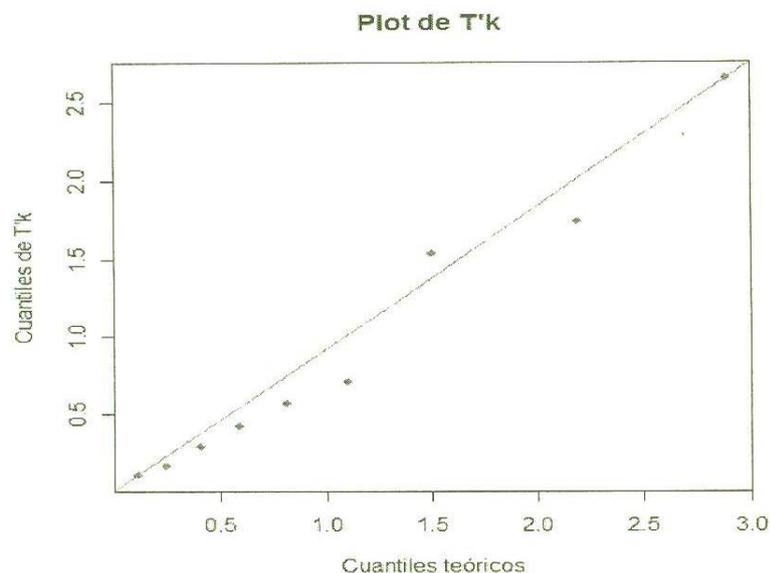


Fig. 4.5 Gráficos QQ-plot de  $T'_n$ .

Obsérvese en la Fig. 4.5 que los tiempos entre rachas consecutivas que muestra el gráfico cuantil-cuantil de los datos no distan mucho de la recta con pendiente 1. Luego entonces, éstas sustentan que una distribución exponencial es razonable para los tiempos entre rachas consecutivas, de esta manera podemos suponer que el proceso de conteo  $N$ , que cuenta las rachas de calor es de Poisson homogéneo con  $\lambda = 1$ , debido a la transformación  $m(t)$  que hemos llevado a cabo al proceso puntual  $\underline{S}$ .

Resumiendo entonces, en el intervalo de tiempo fijo  $(0, 365]$  se ha inspeccionado el fenómeno aleatorio rachas de calor que pueden ser modeladas como un proceso de Poisson no homogéneo  $\{N(t)\}_{t \geq 0}$ , con función de intensidad  $\lambda(t; \alpha, \beta)$  dada por

(4.1). Que el vector  $\underline{S} = (13, 17, 30, 32, 37, 49, 84, 88, 91)$ ; son los tiempos donde inicia cada racha de calor, y que los tiempos entre rachas consecutivas son las variables aleatorias  $T_k = S_k - S_{k-1}$ ,  $k = 1, 2, \dots$  con función de densidad no homogénea

$$f_{T_k}(t_k) = \lambda(t_k) \cdot \exp\left\{-\int_{st_{k-1}}^{s_k} \lambda(s) ds\right\}.$$

### Inferencia estadística basada en verosimilitud

Apoyados en las teorías de probabilidad y estadística desarrolladas en los capítulos precedentes y que hemos validado con simulación, mediante métodos numéricos, realizaremos la inferencia estadística bajo el enfoque de verosimilitud sobre los parámetros  $\alpha, \beta$  y  $\delta$  involucrados en la función de intensidad del proceso de Poisson no homogéneo.

La función de verosimilitud para  $\alpha, \beta, \delta$  está dada entonces por la expresión :

$$L(\alpha, \beta, \delta; \underline{S}) \propto \prod_{k=1}^N \left( \frac{\delta}{\Gamma(\alpha)\beta^\alpha} S_k^{\alpha-1} \exp(-S_k/\beta) \right) \cdot \exp\left\{-\frac{\delta}{\Gamma(\alpha)\beta^\alpha} \int_0^{T_0} s^{\alpha-1} \exp(-s/\beta) ds\right\},$$

mientras que la función log-verosimilitud correspondiente queda expresada por,

$$\ell(\alpha, \beta, \delta; \underline{S}) = c(\underline{S}) + \sum_{k=1}^N \log\left(\frac{\delta}{\Gamma(\alpha)\beta^\alpha} S_k^{\alpha-1} \exp(-S_k/\beta)\right) - \frac{\delta}{\Gamma(\alpha)\beta^\alpha} \int_0^{365} s^{\alpha-1} \exp(-s/\beta) ds;$$

donde  $c(\underline{S})$  es una constante que no depende de los parámetros  $\alpha, \beta$  y  $\delta$ .

Por la expresión de la función log-verosimilitud resulta muy complicado proporcionar de manera explícita una expresión de los estimadores de máxima verosimilitud, utilizando métodos numéricos, procedemos a calcular los valores de los parámetros que maximizan la probabilidad de ocurrencia de la muestra observada. Los resultados obtenidos son:

$$\begin{aligned} \hat{\alpha} &= 2.6; \\ \hat{\beta} &= 18.91; \\ \hat{\delta} &= 9.03. \end{aligned}$$

Por lo que la función de intensidad específica (4.1) de las rachas de calor está expresada por

$$\lambda(t; \alpha, \beta) = 0.0062 t^{1.6} \exp(-t/18.91),$$

y la gráfica de la función está dada por la Figura 4.6,

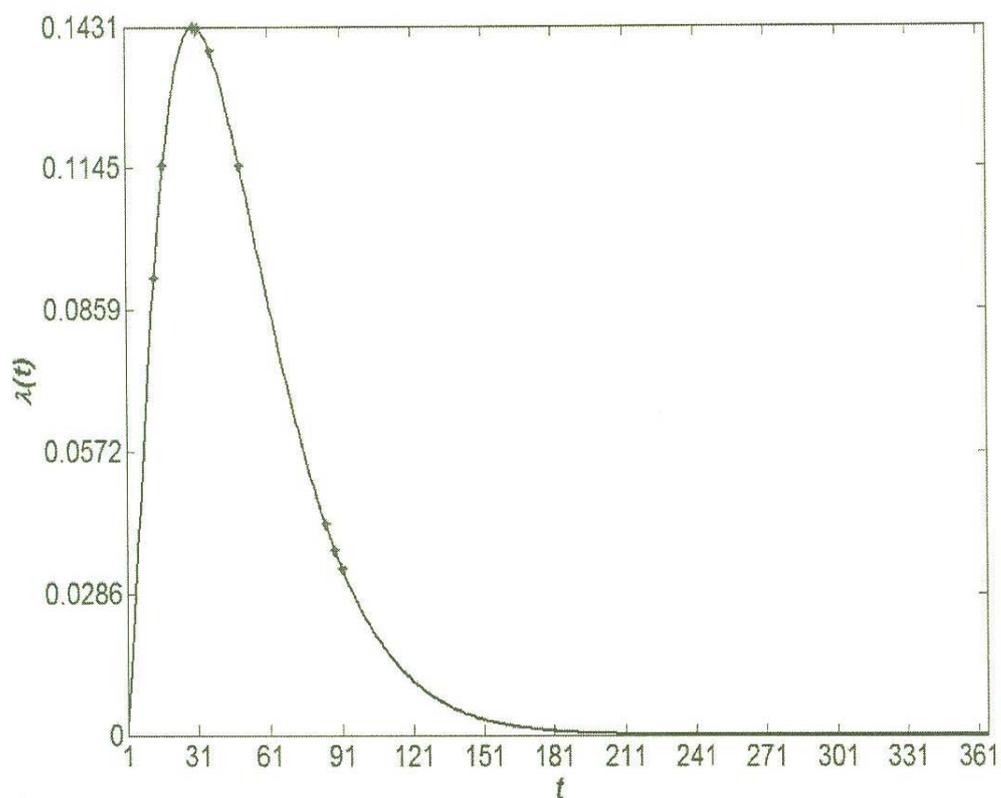


Fig. 4.6 Función de Intensidad de rachas de calor.

Por la forma que presenta la función podemos afirmar entonces, que la intensidad de las rachas de calor en el intervalo de observación  $(0, 365]$  muestra un incremento conforme va transcurriendo el tiempo en el primer cuarto de observación, llega a un máximo rápidamente en 0.143, y muestra efectivamente una tendencia decreciente, tal como lo supusimos, cabe mencionar que el cero de la gráfica de la función de intensidad es a partir de la presencia de las rachas de calor, esto es, a partir del inicio de la primavera.

Para cuantificar con formalidad lo anterior, es preciso realizar inferencia estadística apropiada, siendo este el punto de la metodología propuesta y estudiada. Es a partir de la función de verosimilitud como cuantificaremos los intervalos de verosimilitud. Una manera conveniente de dibujar la función de verosimilitud relativa en dos dimensiones, es trazar contornos de verosimilitud relativa constantes en el plano  $(\alpha, \beta)$ . Los contornos de verosimilitud correspondientes a los niveles de  $c = 0.90, 0.95$  y  $0.99$  con  $\hat{\delta} = 9.03$  fijo, se muestran en el siguiente gráfico.

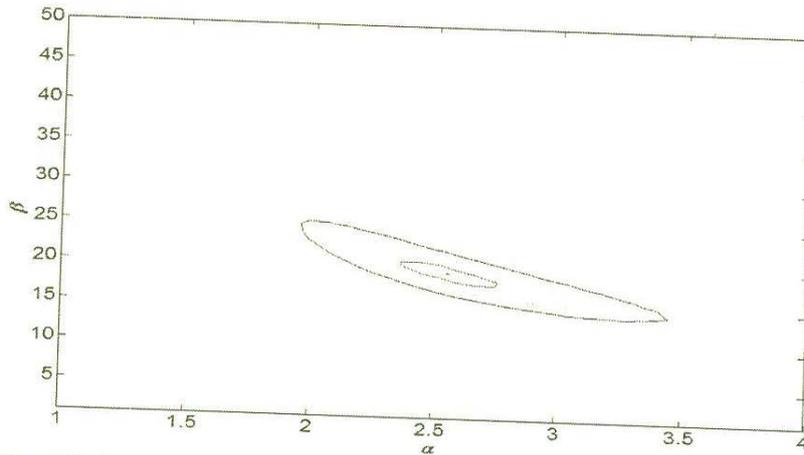


Fig. 4.7 Mapas de contorno para la verosimilitud relativa en el plano  $(\alpha, \beta)$ .

El análisis de estos contornos deberá dar inferencia acerca de la posición  $(\hat{\alpha}, \hat{\beta})$  y de los límites de los intervalos de verosimilitud para cada uno de los parámetros involucrados en cada uno de los niveles de  $c$  tratados. Debido a la variabilidad presente en los datos reflejada en la función de intensidad, es intuitivo esperar que la incertidumbre en la inferencia se reflejará en intervalos de verosimilitud con cierta tendencia de cola hacia la derecha como se puede observar también en los contornos.

Como nuestro modelo involucra tres parámetros es preciso determinar la función de verosimilitud perfil, de tal forma que la función de verosimilitud se pueda escribir solamente como una función del parámetro de interés, con la finalidad de aproximar los intervalos de verosimilitud para cada uno de los parámetros. Considerando primeramente a  $\alpha$  como parámetro de interés, la Figura 4.8 nos muestra la función de verosimilitud perfil de  $\alpha$ , con sus respectivos intervalos de verosimilitud para los valores de  $c = 0.25, 0.15$  y  $0.036$  que como vimos en el capítulo suelen estar asociados a niveles del 90, 95 y 99% de confianza:

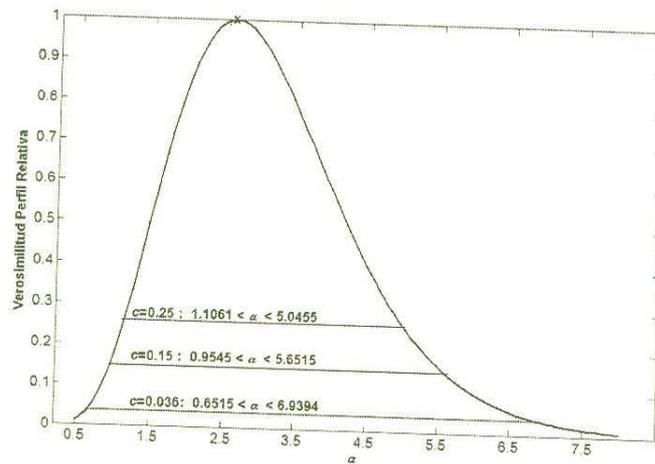


Fig. 4.8 Función de verosimilitud perfil relativa de  $\alpha$ .

Como lo muestra la Figura 4.8, a la luz de los datos observados la función de verosimilitud perfil para  $\alpha$  se presentan de manera asimétrica con respecto al valor de  $\hat{\alpha} = 2.6$ , para valores de  $\alpha$  menores que 0.6 y mayores que 7 son muy implausibles. Ahora considerando a  $\beta$  como parámetro de interés, obtenemos su función de verosimilitud perfil y sus intervalos de verosimilitud aproximados correspondientes a los niveles de  $c = 0.25, 0.15$  y  $0.036$ ,

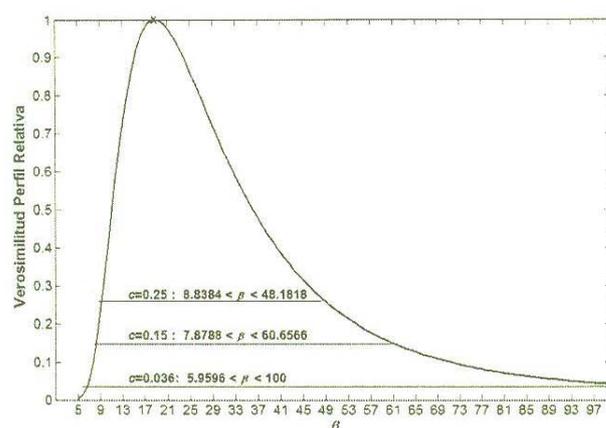


Fig. 4.9 Función de verosimilitud perfil relativa de  $\beta$ .

que como podemos observar en la Figura 4.9, la función de verosimilitud perfil para  $\beta$  es un tanto asimétrica con respecto al valor de  $\hat{\beta} = 18.91$  con cola claramente marcada hacia la derecha. Los límites inferior y superior de los intervalos de verosimilitud para los diferentes niveles de  $c$  para cada uno de los estimadores están dados en el gráfico.

Por último considerando a  $\delta$  como parámetro de interés, la función de verosimilitud perfil para este parámetro se presenta en la Figura 4.10:

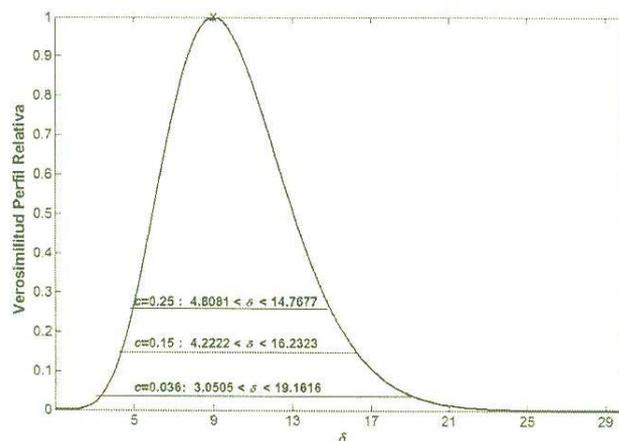


Fig. 4.10 Función de verosimilitud perfil relativa de  $\delta$ .

A la luz de los datos observados, los valores estimados de los parámetros  $\alpha$ ,  $\beta$  y  $\delta$  no tienen interpretación física dentro de la situación problemática. Lo que verdaderamente tiene sentido, es hablar acerca de un rango de la forma que la función de intensidad toma, con la finalidad de predecir el comportamiento de futuras observaciones. Es decir, para las siguientes mediciones se espera que la intensidad de futuras rachas de calor tengan un comportamiento dentro de dicha franja establecida, que es precisamente lo que la Figura 4.11 nos muestra.

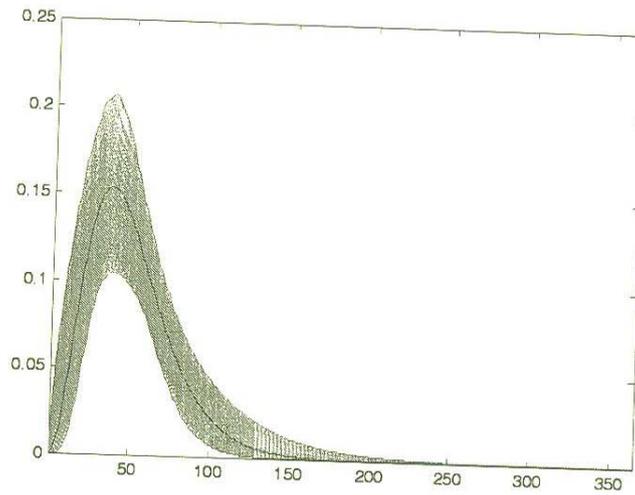


Fig. 4.11. Rangos de la función de intensidad gama del PPNH.

Como podemos observar marca una función de intensidad mínima y una máxima a partir de la función de intensidad media obtenida en este caso de estudio.

## Conclusiones

Una de las motivaciones de este trabajo de tesis fue relacionar las teorías de probabilidad y estadística general, para realizar inferencia estadística en la estimación de los parámetros de un proceso de Poisson por el método de verosimilitud. En ese sentido la aportación principal de esta tesis es proporcionar una metodología específica para generar inferencia en los parámetros de la función de intensidad o tasa de incidencia de un proceso de Poisson, el cual no tendría mayor relevancia si no encontráramos en nuestro entorno la aplicación de esta formulación a un problema real. La solución involucró la concepción detallada del modelo probabilístico adecuado a los datos y la prescripción e implementación de una de las técnicas formales de estimación.

El trabajo realizado requirió considerablemente el apoyo de aspectos computacionales; la aplicación de cómputo que resultó crucial para el desarrollo de esta tesis fueron los programas "R" y MATLAB. Con este último, se trabajó en la implementación de una herramienta para la simulación de procesos de Poisson y en la programación de un método para realizar la elaboración numérica y gráfica de los resultados correspondientes a inferencia. Incluyó una fase de experimentación con el modelo para diversos casos de inspección de este tipo de procesos. El estudio permitió entender de mejor manera el modelo formulado y apreciar el efecto que tienen los parámetros postulados para éste. Por otro lado se utilizó también esta implementación para la exploración y desarrollo de la metodología de inferencia para evaluar distintos escenarios simulados y casos reales.

El análisis de los datos capturados en campo real ha dado lugar a la inferencia estadística acerca de la función de intensidad de un proceso de Poisson no homogéneo; por el contexto y el análisis exploratorio de los datos, se puede decir que presentaron una disposición acorde al método estadístico antes discutido. Por tanto fueron un buen punto de partida para contrastar la teoría y los hechos a los que hace alusión esta tesis.

Es posible utilizar la metodología de inferencia basada en verosimilitud desarrollada en este trabajo de tesis en otros contextos, sólo basta determinar que los datos modelen un proceso de Poisson inspeccionado en un intervalo de tiempo fijo o aleatorio, por nombrar algunos, para inferir propiedades acerca de los parámetros involucrados en el modelo con base en las observaciones realizadas del fenómeno aleatorio.

Las aportaciones principales de esta tesis fueron:

- La caracterización de un modelo para la función de intensidad de un proceso de Poisson no homogéneo con tres parámetros desconocidos.
- El desarrollo de la teoría de inferencia estadística basada en verosimilitud

específicamente para este tipo de procesos de Poisson. Tanto para el caso homogéneo como no homogéneo, exhibiendo los casos para uno y dos parámetros.

- Ejemplificación de la realización del cálculo de los estimadores máximos verosímiles e intervalos de verosimilitud.
- La implementación computacional de programas de simulación para este tipo de procesos, donde simulamos procesos de Poisson con la finalidad de validar y contrastar propiedades con datos simulados versus con datos reales.
- Aplicación de las teorías de probabilidad y estadística desarrolladas en un caso real.



Fig. 4.1 Zonas definidas por el clima en el estado de Guanajuato.

## Bibliografía

- [1] Basawa, I. & Prakasa, R. (1980). *Statistical Inference for Stochastic Process*, Academic Press Inc.
- [2] Billingsley, P. (1979). *Probability and Measure*, John Wiley & Sons, Inc.
- [3] Burden, R. & Douglas, F. (1985). *Análisis Numérico*, Ed. Iberoamérica.
- [4] Canavos George (1999). *Probabilidad y Estadística. Aplicaciones y métodos*, Ed. McGraw Hill.
- [5] Castillo, S. y Lozano A. D. (2007). Q-Q Plot Normal. Los puntos de posición gráfica, *Statistical Papers*.
- [6] Cinlar, E. (1975). *Introduction to Stochastic Processes*, Prentice-Hall, Inc.
- [7] Chem Mok, S. (2004). Definiendo la función de distribución de un parámetro para el análisis de tráfico. *Statistical Paper de la Revista InrerSedes*.
- [8] Davenport, W. Jr (1970). *Probability and random processes*, McGraw-Hill Inc.
- [9] Evans Michael J. & Rosenthal Jeffrery S. (2005). *Probability and Statistics. The Science of Uncertainty*, Ed. Reverté.
- [10] Gómez V. M. A. (2005). *Inferencia Estadística*, Ed. Díaz de Santos.
- [11] Kalbfleisch, J. (1985). *Probability and Statistical Inference*, Springer-Verlag.
- [12] Karlin, S. (1975). *A First Course in Sthocastic Processes*, Academic Press Inc.
- [13] Karlin, S. & Howard, T. (1998). *An Introduccion to Stochastic Modeling*, Academic Press Inc.
- [14] Kulkarni, V. (1995). *Modeling and analisis of stochastic systems*, Chapman & Hall.
- [15] Meeker, William Q. & Escobar, Luis A. (1995). Teaching about Approximate Confidence Regions Based on Maximum Likelihood Estimation. *Statistical Papers The American Statistician*
- [16] Montoya, J. A. (2008). *La verosimilitud perfil en inferencia estadística*, Tesis Doctoral en ciencias con orientación en Probabilidad y Estadística.
- [17] Montoya, J.A., Díaz-Francis, E. y Sprott, D. (2007). On a Criticism of the Profile likelihood Function. *Statistical Paper*.

- 
- [18] Parzen, E. (1962). *Stochastic processes*, Holden-Day Inc.
- [19] Pranab K. & Singer, J. (1993). *Large sample methods in statistics and introduction with applications*, Chapman & Hall.
- [20] Ross, S. (1999). *Simulación*, Prentice Hall Hispanoamericana.
- [21] Resnick, S. (1992). *Adventures in Stochastic Processes*, Ed. Birkhäuser Boston.
- [22] Resnick, S.(2007). *Heavy-Tail Phenomena Probabilistic and Statistical Modeling*. Springer.
- [23] Rusell, A. Boyles. (2008). The Role of Likelihood in Interval Estimation. *Statistical Papers The American Statistician*.
- [24] Serfling, R.J. (1980). *Aproximations Theorems of Mathematical Statistics*. John Wiley and Sons.
- [25] Sprott. D. (2000). *Statistical Inference in Science*, Springer.
-



