

CIMAT

Centro de Investigación en Matemáticas A.C

**PROCESOS PUNTUALES:
EL MODELO DE HAWKES Y SUS APLICACIONES**

T E S I S

PARA OBTENER EL GRADO DE:
**MAESTRO EN CIENCIAS CON ESPECIALIDAD EN
PROBABILIDAD Y ESTADÍSTICA**

PRESENTA:
DARÍO MARTÍN ROMERO CRUZ

DIRECTOR DE TESIS:
DR. DANIEL HERNÁNDEZ HERNÁNDEZ

Guanajuato, Gto., a 8 de febrero de 2018.



Centro de Investigación en Matemáticas, A.C.

Acta de Examen de Grado

Acta No.: 131

Libro No.: 002

Foja No.: 131

En la Ciudad de Guanajuato, Gto., siendo las 14:00 horas del día 8 de febrero del año 2018, se reunieron los miembros del jurado integrado por los señores:

DRA. ELOÍSA DÍAZ-FRANCÉS MURGUÍA (CIMAT)
DR. JOSÉ LUIS ÁNGEL PÉREZ GARMENDIA (CIMAT)
DR. DANIEL HERNÁNDEZ HERNÁNDEZ (CIMAT)

bajo la presidencia del primero y con carácter de secretario el segundo, para proceder a efectuar el examen que para obtener el grado de

MAESTRO EN CIENCIAS CON ESPECIALIDAD EN PROBABILIDAD Y ESTADÍSTICA

Sustenta

DARÍO MARTÍN ROMERO CRUZ

en cumplimiento con lo establecido en los reglamentos y lineamientos de estudios de posgrado del Centro de Investigación en Matemáticas, A.C., mediante la presentación de la tesis

" PROCESOS PUNTUALES: EL MODELO DE HAWKES Y SUS APLICACIONES "

Los miembros del jurado examinaron alternadamente al (la) sustentante y después de deliberar entre sí resolvieron declararlo (a):

Aprobado

DRA. ELOÍSA DÍAZ-FRANCÉS MURGUÍA
Presidente

DR. JOSÉ LUIS PÉREZ GARMENDIA
Secretario

DR. VÍCTOR MANUEL RIVERO MERCADO
Director General

DR. DANIEL HERNÁNDEZ HERNÁNDEZ
Vocal

Agradecimientos

A mis padres, Adolfo Romero y Alma Delia Cruz, que durante toda mi vida han estado cerca para apoyarme y me han brindaron toda su confianza. Sin ustedes no hubiera podido llegar hasta donde estoy ahora, un logro más que compartimos juntos.

A mi director de tesis, el Dr. Daniel Hernández, quien dedicó su tiempo y conocimientos a la realización de este trabajo. También le agradezco el apoyo como profesor y tutor, sepa que se lleva mi completa admiración. Asimismo, un sincero agradecimiento a mis sinodales, la Dra. Eloisa Díaz-Francés y el Dr. José Luis Pérez, por todos los comentarios brindados.

A mis amigos, nuevos y viejos, que contribuyeron a hacer de estos dos años un tiempo muy agradable en mi vida. Y aunque todos son importantes para mi, quiero agradecer de manera especial a Paulina Pérez, una amiga que estuvo presente tanto en los mejores momentos como en los más difíciles, y de quien aprendí mucho. Pau, hoy podemos decir con alegría: lo logramos.

A CIMAT y cada uno de los profesores con los que tuve el placer de convivir.

A CONACyT, que me dio la oportunidad de contar con una de sus becas durante todo el tiempo que conllevó la realización este posgrado.

Resumen

Los procesos puntuales son modelos probabilísticos que frecuentemente se emplean para describir la forma en que ocurren distintos tipos de eventos a lo largo de un periodo de tiempo determinado. En el presente trabajo, plantearemos las bases para dar una definición formal de estos procesos. Además, se discutirán temas relacionados con estos, tales como el proceso de conteo, la función de intensidad y el compensador de un proceso, entre otros. Se describirán los resultados estadísticos, como estimación paramétrica, pruebas de bondad de ajuste, simulación y predicción. Por último, introduciremos un tipo particular de proceso puntal, denominado el proceso de Hawkes, mostrando aplicaciones y ejemplos.

Contenido

Introducción	1
1. Procesos Puntuales	3
1.1. Procesos puntuales y el proceso de conteo	3
1.2. Medidas aleatorias	6
1.2.1. Medidas de conteo	6
1.2.2. Medida media	7
1.3. El proceso de Poisson	8
1.4. Martingalas y el compensador del proceso	9
1.4.1. Definiciones básicas	9
1.4.2. Aplicación a procesos puntuales	10
1.5. Función de intensidad	11
1.6. Procesos puntuales regulares	12
1.7. Función de supervivencia	13
1.8. Función de densidad conjunta	15
1.9. Procesos puntuales marcados	16
1.9.1. Definiciones básicas	17
1.9.2. Función de intensidad condicional para procesos marcados	18
1.9.3. Clasificación de procesos puntuales marcados	19
1.9.4. El proceso de Poisson compuesto	19
1.10. Procesos puntuales multivariados	20
2. Estadística aplicada a los procesos puntuales	21
2.1. Estimadores de máxima verosimilitud	21
2.1.1. Estimación para procesos univariados	21
2.1.2. Estimación para el proceso de Poisson homogéneo	22
2.1.3. Estimación para procesos marcados y multivariados	24
2.2. Transformación de tiempos aleatorios	24

2.3.	Bondad de ajuste	26
2.3.1.	Pruebas de uniformidad	27
2.3.2.	Pruebas de exponencialidad	28
2.3.3.	Prueba ji-cuadrada	28
2.3.4.	Aproximación al movimiento browniano	32
2.4.	Simulación de procesos puntuales	33
2.4.1.	Simulación del proceso de Poisson homogéneo	33
2.4.2.	Método de la inversa del compensador	33
2.4.3.	Método de adelgazamiento	33
2.4.4.	Simulación de un proceso puntual con función de intensidad acotada a trozos	34
2.4.5.	Simulación de un proceso puntual multivariado	36
2.5.	Predicción	37
2.6.	Ejemplo	38
2.6.1.	Simulación	38
2.6.2.	Bondad de ajuste	39
3.	El Proceso de Hawkes	41
3.1.	Caso unidimensional	41
3.1.1.	Verosimilitud del proceso de Hawkes	42
3.1.2.	Simulación del proceso de Hawkes	43
3.2.	Selección de la función de excitación	44
3.2.1.	Función con decrecimiento exponencial	44
3.2.2.	Ejemplos	46
3.2.3.	Otras funciones de excitación	50
3.3.	El proceso de Hawkes multidimensional	51
3.3.1.	Ejemplo	52
3.4.	Relación del proceso de Hawkes con los procesos de ramificación	54
	Bibliografía	58

Introducción

Los procesos puntuales son una herramienta muy importante dentro de la aplicación de las matemáticas. Su uso nos permite abordar una amplia gama de problemas en los cuales el objeto de estudio son puntos y nos interesa conocer la dispersión de los mismos dentro de algún espacio. Frecuentemente, la connotación que se le da a este espacio es el tiempo, mientras que cada punto representa un instante en el cual sucede algún evento determinado. De este modo, el problema se transforma en tratar de describir la forma en que ocurren diversos eventos, por ejemplo, la llegada de clientes a una cola de espera, el momento en el que nace un individuo dentro de una población o el instante en el que ocurre un accidente vial.

Este tema no es nuevo, y en la actualidad se tiene desarrollada bastante teoría para tratarlo. En la literatura, el modelo más famoso que podemos encontrar en el conjunto de procesos puntuales es sin lugar a dudas el denominado proceso de Poisson, nombrado de esta manera por la distribución que sigue el número de eventos ocurridos en un intervalo de tiempo fijo. La ventaja que ofrece la implementación de este modelo radica en la sencillez de sus cálculos, cosa que muy pocos procesos puntuales mantienen.

Sin embargo, con la creciente capacidad de procesamiento informático, resulta cada vez más fácil el utilizar modelos que tengan una mayor complejidad en su estructura. En este contexto, destaca el conocido proceso de Hawkes, introducido hacia la década de los 70's, que aborda los casos en los que los fenómenos que se estudian presentan agrupamientos durante ciertos periodos. Más aún, estos agrupamientos son provocados por la denominada propiedad de *auto-excitación* del proceso, esto es, que cada que ocurre un nuevo evento entonces aumenta la probabilidad, por un cierto lapso de tiempo, de que suceda otro posterior al primero. Lo anterior se da gracias a un incremento instantáneo en la tasa de llegada de los eventos con cada nuevo suceso.

El presente trabajo sirve como introducción a toda esta clase de modelos. En el Capítulo 1 presentaremos la teoría que existe detrás de los procesos puntuales, iniciaremos planteando su definición formal viendo al modelo como una sucesión de tiempos, una función escalonada e incluso una medida de probabilidad. Además, introduciremos algunos otros procesos relacionados con éste, tales como los procesos de conteo o el compensador, entre otros. Este capítulo tiene un mayor enfoque en teoría de la probabilidad, sentando las bases para los resultados que se den en los dos capítulos siguientes.

En el Capítulo 2 se discutirá sobre la estadística que se puede aplicar a los procesos puntuales. Temas que se discutirán aquí son estimación, pruebas de bondad de ajuste,

simulación y predicción, entre otros.

Finalmente, el Capítulo 3 está dedicado esencialmente al modelo de Hawkes, describiendo sus características y aplicando los temas vistos en los capítulos anteriores. Aquí se incluyen una serie de ejemplos que ayudan al lector a comprender la estructura del proceso. El capítulo finaliza vinculando el proceso de Hawkes con los procesos de ramificación.

Para profundizar más en el tema, se incluye una extensa bibliografía, la cual puede ser consultada conforme se sigue este escrito.

1

Procesos Puntuales

Los procesos puntuales son modelos probabilísticos que, como se dijo en la introducción, describen la dispersión de puntos aleatorios sobre algún espacio, y con frecuencia son empleados para modelar los instantes en que ocurren determinados tipos de eventos dentro de un intervalo de tiempo dado. En tal caso, el espacio donde se encuentran dispersos suele ser el conjunto de los números reales no negativos o un subconjunto de éste, cada punto representa un nuevo evento y su ubicación es el tiempo en el cual dicho evento tiene lugar.

Por simplicidad, comenzaremos enfocándonos en el caso en que los puntos son unidimensionales, por lo cual, la única distinción entre dos eventos será el instante en el que ocurren. La generalización de esto se da cuando empezamos a incluir diferentes tipos de eventos, para lo cual se incorporarán los conceptos de procesos puntuales marcados y multivariados que serán nuestro objeto de estudio en las últimas secciones del capítulo. Gran parte de este capítulo se basa particularmente en [1],[3],[8],[17] y [18]. Fuentes que pueden ser consultados por el lector para extender su conocimiento de este tema.

1.1. Procesos puntuales y el proceso de conteo

A continuación presentaremos una definición formal de proceso puntual. Más específicamente, la siguiente definición corresponde a un proceso puntual simple y sin explosión, estas consideraciones adicionales serán importantes en la teoría que pretendemos desarrollar y se describirán con mayor detalle más adelante.

Consideremos un espacio de probabilidad (Ω, \mathcal{A}, P) , con \mathcal{A} una σ -álgebra para los subconjuntos en Ω y P una medida de probabilidad en \mathcal{A} .

Definición 1.1 (Proceso puntual). Sea $T = \{T_n; n \in \mathbb{N}\}$ una secuencia de variables aleatorias definidas en (Ω, \mathcal{A}, P) y que toman valores en $\bar{\mathbb{R}}_0 = \mathbb{R}_0 \cup \{\infty\}$, con $\mathbb{R}_0 = [0, \infty)$. Si satisface que:

- i) $P(0 < T_1 \leq T_2 \leq \dots) = 1$,
- ii) $P(T_n < T_{n+1}, T_n < \infty) = P(T_n < \infty), n \geq 1$, y

iii) $P(\lim_{n \rightarrow \infty} T_n = \infty) = 1$,

entonces T se denominará un *proceso puntual simple*.

De esta manera, un proceso puntual se puede ver como una secuencia positiva, posiblemente infinita, y estrictamente creciente de variables aleatorias. La interpretación de T_n , cuando es finito, será el instante en que el n -ésimo evento tiene lugar. Por otro lado, cuando $T_n = \infty$, lo interpretaremos como que el número total de eventos ocurridos a lo largo de \mathbb{R}_0 será menor a n .

El término *simple* en la Definición 1.1 hace referencia a que nuestro proceso excluye los casos en los cuales ocurren más de un evento en un mismo instante de tiempo (casi seguramente), lo anterior es implicación de la Condición (ii). Además, la Condición (iii) es equivalente a decir que en cada intervalo de tiempo finito solo puede ocurrir un número finito de eventos, en otras palabras, el proceso no presenta *explosión*. Si esta última condición no se satisface, entonces diremos que la secuencia T será un *proceso puntual simple con explosión*.

Si denotamos por \mathbb{K} al espacio de secuencias de puntos dado por

$$\mathbb{K} = \{ \{t_n; n \geq 1\} \in \bar{\mathbb{R}}_0^{\mathbb{N}} : 0 < t_1 \leq t_2 \leq \dots \uparrow \infty, t_n < t_{n+1} \text{ si } t_n < \infty \}, \quad (1.1)$$

y sea \mathcal{K} la σ -álgebra canónica, generada por la secuencia de variables aleatorias $\{T_1, T_2, \dots\} \in \mathbb{K}$. Entonces tenemos que T será una variable aleatoria definida sobre $(\mathbb{K}, \mathcal{K})$ y cuya distribución se obtiene de P a través de la transformación

$$\Pr(T \in K) = P(\{\omega : T(\omega) \in K\}), \quad K \in \mathcal{K}.$$

Ahora introduciremos un nuevo proceso, el cual está fuertemente ligado al proceso puntual y que llamaremos el proceso de conteo. Para ello, comenzamos definiendo la función \mathcal{N} . Sea $T = \{T_n; n \in \mathbb{N}\}$ un proceso puntual, denotaremos por $\mathcal{N}(A)$ al número de eventos (o puntos) de T que ocurren dentro de algún intervalo A , esto es,

$$\mathcal{N}(A) = \sum_n \mathbb{1}_{\{T_n \in A\}}, \quad (1.2)$$

para todo $A \subseteq \mathbb{R}$.

Notemos que $\mathcal{N}(A)$ tomará valor en $\bar{\mathbb{N}}_0 = \mathbb{N}_0 \cup \{\infty\}$, con $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$. Además, dada la Condición (iii) de la Definición 1.1, sabemos que será finito para todo A que sea un conjunto acotado.

Adicionalmente, supongamos que A es formada por la unión de los conjuntos disjuntos A_1, \dots, A_k , es decir,

$$A = \bigcup_{i=1}^k A_i, \quad \text{con } A_i \cap A_j = \emptyset, i \neq j.$$

Entonces se satisface que

$$\mathcal{N}(A) = \sum_{i=1}^k \mathcal{N}(A_i).$$

Esta consideración se puede extender al caso en que $k = \infty$.

Como estaremos interesados en calcular valores tales como la probabilidad conjunta

$$P(\mathcal{N}(A_1) = n_1, \dots, \mathcal{N}(A_k) = n_k),$$

para $n_i = 0, 1, 2, \dots$, $i = 1, \dots, k$ y $k = 1, 2, \dots$, con A_i conjuntos arbitrarios en \mathbb{R} . Para que estas probabilidades estén bien definidas, en lo sucesivo impondremos la restricción de que los A_i , en los cuales evaluaremos nuestra función \mathcal{N} , pertenezcan a los conjuntos de Borel de \mathbb{R} .

Definición 1.2 (Proceso de conteo). Sea T un proceso puntual en (Ω, \mathcal{A}, P) . El proceso $N = \{N(t); t \geq 0\}$, con $N(0) \equiv 0$ y

$$N(t) = \mathcal{N}((0, t]), \quad t > 0,$$

es el *proceso de conteo* asociado a T .

Para simplificar la notación, emplearemos en algunos casos $N_t = N(t)$.

De la definición anterior se observa que, dado un tiempo fijo $t < \infty$, $N_t(\omega)$ se puede ver como una variable aleatoria que toma valores en \mathbb{N}_0 . Por otro lado, si lo que se considera fijo es una realización $\omega \in \Omega$, entonces $N_t(\omega)$ forma una función en t escalonada y continua por la derecha.

Podemos definir el espacio de procesos de conteo \mathbb{W} como

$$\mathbb{W} = \left\{ w \in \mathbb{N}_0^{\mathbb{R}_0} : w(t) \text{ es creciente y continua por la derecha con } w(0) = 0 \right. \\ \left. \text{y } \Delta w(t) \in \{0, 1\}, \forall t \geq 0 \right\},$$

donde $\Delta f(t) = f(t) - f(t-)$.

Si denotamos por \mathcal{W} a la menor σ -álgebra de subconjuntos de \mathbb{W} generada por

$$\mathcal{W} = \sigma(\{N_t; t \geq 0\}),$$

para $N = \{N_t; t \geq 0\} \in \mathbb{W}$. Entonces N se puede ver como una variable aleatoria definida sobre $(\mathbb{W}, \mathcal{W})$ y cuya distribución esta dada por la transformación

$$\Pr(N \in W) = P(\{\omega : N(\omega) \in W\}), \quad W \in \mathcal{W}.$$

Existe una relación biyectiva entre un proceso puntual y su proceso de conteo, esto debido a la equivalencia de eventos:

$$\{N_t < n\} \equiv \{T_n > t\}.$$

Por lo cual, dado un proceso de conteo N definido en el espacio de probabilidad (Ω, \mathcal{A}, P) , podremos obtener el proceso puntual $T = \{T_n\}$ al tomar

$$T_n = \inf \{t \geq 0 : N_t \geq n\}, \quad \forall n \in \mathbb{N},$$

con $\inf \emptyset \equiv \infty$.

Por esta relación biyectiva entre ambos procesos, algunos autores emplean los términos *proceso puntual* y *proceso de conteo* indistintamente tanto para referirse a N como a T . A pesar de esto, en lo subsecuente mantendremos la diferencia entre ambos términos. Por último, considerando ahora que $T = \{T_n; n \in \mathbb{Z}\}$ puede contener puntos en cualquier lugar de $\bar{\mathbb{R}} = [-\infty, \infty]$, tal que

$$T_n < T_{n+1}, \quad T_0 < 0 < T_1.$$

Una extensión conveniente de la Definición 1.2, para todo $t \in \mathbb{R}$, se da de la siguiente manera:

$$N_t = \begin{cases} \mathcal{N}((0, t]), & \text{si } t > 0, \\ 0, & \text{si } t = 0, \\ -\mathcal{N}((t, 0]), & \text{si } t < 0. \end{cases}$$

Así, N_t seguirá siendo una función continua por la derecha con valores enteros. Los resultados que se presentan en las siguientes secciones solo considerarán el escenario en que los eventos ocurren en tiempos $t > 0$, sin embargo, no se necesitan muchas modificaciones a estos resultados para abarcar el caso en el cual sucedan a lo largo de toda la recta real.

1.2. Medidas aleatorias

Hasta ahora hemos dado dos maneras de identificar la realización de un proceso puntual:

- Considerando un conjunto numerable de puntos T , y
- A través de una función escalonada y continua por la derecha N_t .

Sin embargo existe una definición más general y que es de gran utilidad cuando se pretenda extender el concepto de proceso puntual para considerar los casos en que los puntos se encuentran ubicados en un espacio diferente a $\bar{\mathbb{R}}_0$, esto es, definir al proceso como una sucesión de medidas aleatorias.

1.2.1. Medidas de conteo

Dado un espacio de medida (E, \mathcal{E}) , donde E es un conjunto no vacío y \mathcal{E} representa una σ -álgebra de los subconjuntos de E , por ejemplo, su conjunto potencia en el caso de que sea discreto o la σ -álgebra de Borel de E en el caso de que sea euclidiano. Además, sean X_1, X_2, \dots variables aleatorias que toman valores en E .

Definición 1.3 (Medida de conteo). Una *medida de conteo* (también llamada *medida puntual*), μ , sobre (E, \mathcal{E}) es una medida determinada por la secuencia finita o infinita numerable de puntos en E , $\{x_i; i \in I\}$, tal que $\mu(A)$ es el número de puntos que se encuentran en A , para todo $A \in \mathcal{E}$. Esto es,

$$\mu(A) = |\{x_i; i \in I\} \cap A| = \sum_{i \in I} \mathbb{1}_{\{x_i \in A\}} = \sum_{i \in I} \delta_{x_i}(A),$$

donde $\delta_x(\cdot)$ es la medida de Dirac con masa uno en x .

Notemos que la Definición 1.3 permite la existencia de puntos iguales en el conjunto $\{x_i; i \in I\}$. Esta multiplicidad se puede dar frecuentemente en el caso en que E sea un espacio de estados discreto. Una medida de conteo se dice *simple* si solo contiene puntos distintos entre sí (casi seguramente),

$$\mu(\{x\}) \leq 1, \quad \forall x \in E.$$

Denotamos por $\mathbb{M} = \mathbb{M}(E, \mathcal{E})$ al *espacio de medidas de conteo* en \mathcal{E} , y $\mathcal{M} = \mathcal{M}(E, \mathcal{E})$ representará la menor σ -álgebra de los subconjuntos de \mathbb{M} dada por

$$\mathcal{M} = \sigma(\{\mu \in \mathbb{M} : \mu(A) \leq k\} : A \in \mathcal{E}, 0 \leq k \leq \infty).$$

Usaremos la notación \mathbb{M}_s para representar el *espacio de medidas de conteo simples* dado por

$$\mathbb{M}_s = \{\mu \in \mathbb{M} : \mu\{x\} \leq 1, \forall x \in E\}.$$

Definición 1.4 (Proceso de conteo). Sea (Ω, \mathcal{A}, P) un espacio de probabilidad. Un mapeo

$$M : \Omega \rightarrow \mathbb{M}(E, \mathcal{E})$$

se dice un *proceso de conteo* (o *proceso puntual*) en (E, \mathcal{E}) si es medible con respecto a \mathcal{A} y $\mathcal{M}(E, \mathcal{E})$. En este caso, se dice que M forma una medida aleatoria de conteo sobre (E, \mathcal{E}) .

Un proceso de conteo M será *simple* si $P(M \in \mathbb{M}_s) = 1$. Además, diremos que el proceso es σ -finito sobre E si M corresponde a una medida σ -finita, esto es, si $M(E_i) < \infty$, $i \in I$, para todos los conjuntos E_i disjuntos a pares tales que $\cup_{i \in I} E_i = E$. Más adelante, estas definiciones nos ayudarán a presentar al proceso en espacios de dimensión mayor a uno. Sin embargo, por el momento seguiremos enfocándonos en el caso unidimensional.

Proposición 1.5. \mathcal{N} , definida en (1.2), es una medida aleatoria de conteo sobre $(\mathbb{R}_0, \mathcal{B}_0)$, donde $\mathcal{B}_0 = \mathcal{B}(\mathbb{R}_0)$ representa la σ -álgebra de Borel de \mathbb{R}_0 .

De esta proposición se sigue que el proceso de conteo N , presentado en la Definición 1.2, sea en efecto una sucesión de medidas.

1.2.2. Medida media

La medida media juega un papel importante en la caracterización de un proceso puntual, pese a que no existe una correspondencia uno a uno entre estos dos como en el caso del proceso puntual y de conteo.

Sea T un proceso puntual en (Ω, \mathcal{A}, P) . Denotamos por $\mu_T(A)$ al número esperado de eventos de T que ocurren dentro de A , esto es,

$$\mu_T(A) := E(\mathcal{N}(A)) = E\left(\sum_n \mathbb{1}_{\{T_n \in A\}}\right), \quad A \in \mathcal{B}_0. \quad (1.3)$$

Lema 1.6. μ_T es una medida.

Demostración. Observemos que $\mu_T(A) \geq 0$, dado que $\mathcal{N}(A) \geq 0$, para todo $A \in \mathcal{B}_0$. Y $\mu_T(\emptyset) = \mathbb{E}(\mathcal{N}(\emptyset)) = 0$.

Además, dados los conjuntos disjuntos $\{A_i; i \in I\}$, con $A_i \in \mathcal{B}_0$ y $A_i \cap A_j = \emptyset$ para $i \neq j$, entonces

$$\mu_T\left(\sum_i A_i\right) = \mathbb{E}\left[\mathcal{N}\left(\sum_i A_i\right)\right] = \mathbb{E}\left[\sum_i \mathcal{N}(A_i)\right] = \sum_i \mathbb{E}[\mathcal{N}(A_i)] = \sum_i \mu_T(A_i),$$

con lo cual se satisface la σ -aditividad de μ_T , concluyendo así la demostración. \square

Definición 1.7. μ_T , definida en (1.3), es conocida como la *medida media* (también llamada *medida de intensidad* o *medida de primer momento*) de T .

Además, denotaremos por

$$\mu_t = \mu_T((0, t]),$$

a la *función de valor medio* del proceso.

La densidad de Lebesgue de la función de valor medio es llamada *función de intensidad* o simplemente *intensidad*. Esta función, suponiendo que exista, también será de gran importancia en el desarrollo de este trabajo, por lo cual seguiremos haciendo mención de ella en las secciones siguientes. Si la función de intensidad es igual a una constante λ , a este valor lo llamaremos la *tasa* o *intensidad* del proceso.

1.3. El proceso de Poisson

El proceso de Poisson es uno de los procesos puntuales más sencillos para trabajar y de los que se han desarrollado mayores resultados. Una vez más, aunque también es posible definir un proceso de Poisson para puntos que se encuentren localizados dentro de algún espacio arbitrario E , nos enfocaremos en el caso en que $E = \mathbb{R}_0$. Y además, la ubicación de cada punto la utilizaremos para representar el instante de tiempo en el que ocurre cierto evento.

Definición 1.8 (Proceso de Poisson). Un proceso puntual T definido en $(\Omega, \mathcal{A}, \mathbb{P})$ se dice un *proceso de Poisson* con medida media μ si:

- i) Para todos los conjuntos disjuntos $A_1, \dots, A_k \in \mathcal{B}_0$, las variables aleatorias $\mathcal{N}(A_1), \dots, \mathcal{N}(A_k)$ son independientes.
- ii) Para todo $A \in \mathcal{B}_0$ y $k \geq 0$,

$$\mathbb{P}(\mathcal{N}(A) = k) = \frac{e^{-\mu(A)} [\mu(A)]^k}{k!}.$$

Donde \mathcal{B}_0 denota la σ -álgebra de Borel de \mathbb{R}_0 . Por convención $\mathbb{P}(\mathcal{N}(A) = \infty) = 1$, si $\mu(A) = \infty$.

Sea N el proceso de conteo asociado a T , la primera condición implica que N tiene incrementos independientes. Por otro lado, (ii) indica que el número de eventos ocurridos en A seguirá una distribución Poisson de parámetro $\mu(A)$ cuando $\mu(A) < \infty$.

Teorema 1.9. *Sea μ una medida σ -finita. Existe un único proceso de conteo que satisface las condiciones (i) y (ii) de la Definición 1.8.*

La unicidad a la que hace referencia el teorema anterior es en el sentido de que si existen dos procesos que satisfagan ambas condiciones, estos procesos tienen la misma distribución. La demostración de éste teorema se excluye del presente trabajo, pero puede observarse en [17].

Un caso particular del proceso de Poisson se da cuando el proceso presenta homogeneidad temporal, es decir, la distribución del número de eventos en cierto intervalo solo depende de la longitud de dicho intervalo pero no de su posición en el tiempo.

Definición 1.10 (Proceso de Poisson homogéneo). Sea T un proceso de Poisson, diremos que es *homogéneo* si y solo si su medida media está dada por $\mu(A) = \lambda |A|$, donde $|A|$ representa la medida de Lebesgue de A y $\lambda > 0$ será una constante denominada la tasa o intensidad de T . En caso contrario, el proceso se dirá *no homogéneo*.

Consideremos un proceso de Poisson homogéneo, $\{T_n; n \in \mathbb{N}\}$, cuya tasa es λ . Una característica a destacar de este proceso es que los tiempos entre llegadas, dados por $W_n = T_n - T_{n-1}$ para $n \in \mathbb{N}$ y con $S_0 \equiv 0$, serán variables aleatorias independientes con distribución exponencial y media $1/\lambda$.

1.4. Martingalas y el compensador del proceso

En esta sección describiremos otro proceso directamente relacionado al proceso puntual y el proceso de conteo denominado compensador. Para ello comenzamos dando un pequeño repaso de algunos conceptos básicos en la teoría de procesos y martingalas.

1.4.1. Definiciones básicas

Sea (Ω, \mathcal{A}, P) un espacio de probabilidad.

Definición 1.11 (Filtración). Una familia $\mathcal{F} = \{\mathcal{F}_t; t \geq 0\}$ de sub- σ -álgebras de \mathcal{A} es llamada *filtración* si

$$\mathcal{F}_s \subset \mathcal{F}_t, \quad \forall s < t.$$

Adicionalmente, sea

$$\mathcal{F}_{t+} \equiv \bigcap_{s>t} \mathcal{F}_s,$$

si $\mathcal{F}_t = \mathcal{F}_{t+}$ entonces diremos que la filtración \mathcal{F} es *continua por la derecha*.

Consideremos ahora un proceso $X = \{X_t; t \geq 0\}$ definido sobre (Ω, \mathcal{A}, P) y que toma valores en los reales, esto es, $X_t : (\Omega, \mathcal{A}) \rightarrow (\mathbb{R}, \mathcal{B})$ tal que $(t, \omega) \mapsto X_t(\omega)$, donde \mathcal{B} denota la σ -álgebra de Borel de \mathbb{R} .

Definición 1.12 (Proceso adaptado). Decimos que X es *adaptado* a la filtración \mathcal{F} o \mathcal{F} -*adaptado* si X_t es \mathcal{F}_t -medible para todo $t \geq 0$.

Definición 1.13 (Proceso predecible). Sea X un proceso \mathcal{F} -adaptado. Decimos que X es \mathcal{F} -*predecible* si X_t es medible respecto a la σ -álgebra predecible:

$$\sigma(\{(s, t] \times A : 0 < s < t, A \in \mathcal{F}_t\}).$$

De esta forma, si X es \mathcal{F} -predecible, entonces será \mathcal{F}_{t-} -medible. De forma práctica, diremos que X es predecible si es adaptado a \mathcal{F} y continuo por la izquierda.

Definición 1.14 (Martingala). El proceso $X = \{X_t; t \geq 0\}$ se dice *martingala* con respecto a \mathcal{F} o \mathcal{F} -*martingala* si

$$X_s = E(X_t | \mathcal{F}_s), \quad c.s.$$

para todo $s \leq t$. Más aún, X es *submartingala* con respecto a \mathcal{F} si $X_s \leq E(X_t | \mathcal{F}_s)$ para todo $s \leq t$, casi seguramente. De forma análoga, se dirá *supermartingala* si $X_s \geq E(X_t | \mathcal{F}_s)$.

1.4.2. Aplicación a procesos puntuales

Consideremos un proceso de conteo $N = \{N_t, 0 \leq t \leq \tau\}$ definido sobre el espacio de probabilidad (Ω, \mathcal{A}, P) hasta un tiempo τ finito. En este caso, sabemos que $N_t < \infty$ para $t \in [0, \tau]$ (c.s.).

Definición 1.15 (Filtración natural). Denotamos por \mathcal{H}_t la menor σ -álgebra tal que las variables aleatorias N_s con $s \in [0, t]$ sean medibles, es decir,

$$\mathcal{H}_t = \sigma(\{N_s : 0 \leq s \leq t\}), \quad \forall t \in [0, \tau]. \quad (1.4)$$

A la filtración dada por $\mathcal{H} = \{\mathcal{H}_t; t \geq 0\}$ la llamaremos *filtración natural* o *canónica* de N .

Tomemos otra filtración $\mathcal{F} = \{\mathcal{F}_t; 0 \leq t \leq \tau\}$, a la cual denominaremos la “historia del proceso”, de modo tal que $\mathcal{H}_t \subseteq \mathcal{F}_t$ para todo t . De esta manera, nuestro proceso N será \mathcal{F} -adaptado.

Sería válido emplear $\mathcal{F} = \mathcal{H}$, sin embargo, de forma más general utilizaremos esta otra filtración con el objetivo de representar la posible existencia de información adicional y externa al proceso. Por ejemplo, sea \mathcal{F}_0 la información con la que se cuenta al momento en el que nuestro proceso comienza, entonces podemos tomar $\mathcal{F}_t = \mathcal{F}_0 \cup \mathcal{H}_t$, para todo $t \geq 0$.

Notemos que $N_s \leq N_t$ para todo $s \leq t$, por lo tanto tenemos que

$$E(N_t | \mathcal{F}_s) \geq E(N_s | \mathcal{F}_s) = N_s, \quad c.s.$$

de lo cual se sigue que N es una \mathcal{F} -submartingala.

Por lo tanto, utilizando el Teorema de descomposición de Doob, sabemos que es posible sustraer cierto valor al proceso N para hacerlo una martingala. Este valor formará a su vez un nuevo proceso, al cual llamaremos compensador, y se define de la siguiente manera.

Definición 1.16 (Compensador). El proceso $\Lambda = \{\Lambda(t); 0 \leq t \leq \tau\}$ se conoce como \mathcal{F} -compensador de N si es monótono no decreciente y \mathcal{F} -predecible, tal que el proceso

$$\{N_t - \Lambda(t), 0 \leq t \leq \tau\}$$

sea una \mathcal{F} -martingala con media cero.

Definición 1.17 (Función de intensidad). Si además existe un proceso no negativo y \mathcal{F} -predecible $\lambda = \{\lambda(t); 0 \leq t \leq \tau\}$, tal que

$$\Lambda(t) = \int_0^t \lambda(s) ds, \quad t \in [0, \tau]. \quad (1.5)$$

Entonces λ es llamado el *proceso de intensidad* de N .

De igual forma que se hizo para $N(t)$, emplearemos con frecuencia la notación Λ_t y λ_t para hacer referencia a $\Lambda(t)$ y $\lambda(t)$ respectivamente.

Notemos que, dada una realización $\omega \in \Omega$, $\lambda_t(\omega)$ puede verse como una función continua por la izquierda. λ_t será conocida como la *función de intensidad* (o *función de intensidad condicional*) del proceso.

1.5. Función de intensidad

Dado un proceso de conteo $N = \{N_t; 0 \leq t \leq \tau\}$ sobre el espacio de probabilidad filtrado $(\Omega, \mathcal{A}, \mathcal{F}, \mathbb{P})$, con $\mathcal{F} = \{\mathcal{F}_t; t \geq 0\}$ una filtración de \mathcal{A} que representa la historia del proceso, esto es, $\mathcal{H}_t \subset \mathcal{F}_t$ para todo $t \in [0, \tau]$, donde $\mathcal{H}_t = \sigma\{N_s; 0 \leq s \leq t\}$ corresponde a la filtración natural del proceso.

Sean Λ y λ el compensador y el proceso de intensidad asociados a N respectivamente. Entonces, para todo $h \geq 0$, se tiene que

$$\mathbb{E}(N_{(t+h)-} - \Lambda_{(t+h)-} | \mathcal{F}_{t-}) = N_{t-} - \Lambda_{t-},$$

y agrupando términos,

$$\mathbb{E}(N_{(t+h)-} - N_{t-} | \mathcal{F}_{t-}) = \mathbb{E}(\Lambda_{(t+h)-} - \Lambda_{t-} | \mathcal{F}_{t-}).$$

Supongamos que existe la función de intensidad λ_t del proceso, dada en la Definición 1.17, entonces podemos demostrar que

$$\Lambda_{(t+h)-} - \Lambda_{t-} = \int_t^{t+h} \lambda_u du = \Lambda_{t+h} - \Lambda_t,$$

por lo tanto

$$\mathbb{E}(N_{(t+h)-} - N_{t-} | \mathcal{F}_{t-}) = \mathbb{E}(\Lambda_{t+h} - \Lambda_t | \mathcal{F}_{t-}).$$

Dividiendo ambos lados de la última igualdad entre h y tomando el límite cuando $h \downarrow 0$, tenemos que

$$\begin{aligned} \lim_{h \downarrow 0} \frac{1}{h} \mathbb{E}(N_{(t+h)-} - N_{t-} | \mathcal{F}_{t-}) &= \lim_{h \downarrow 0} \frac{1}{h} \mathbb{E}(\Lambda_{t+h} - \Lambda_t | \mathcal{F}_{t-}), \\ &= \frac{d\mathbb{E}(\Lambda_t | \mathcal{F}_{t-})}{dt}. \\ &= \frac{d\Lambda_t}{dt} = \lambda_t. \end{aligned}$$

Cabe destacar que en la ecuación anterior fue posible sacar a Λ_t de la esperanza condicional dado que, por definición, Λ es un proceso \mathcal{F} -predecible. De todo lo anterior, se obtiene el siguiente lema.

Lema 1.18. *Sea λ_t la función de intensidad de N , entonces*

$$\lambda_t = \frac{E(dN_t | \mathcal{F}_{t-})}{dt}, \quad t \in [0, \tau], \quad (1.6)$$

donde $dN_t = N_{(t+dt)-} - N_{t-}$ representa el número de eventos ocurridos en el intervalo $[t, t + dt)$.

Equivalentemente, si suponemos que N es un proceso de conteo simple, entonces

$$\begin{aligned} \lambda_t dt &= P(dN_t = 1 | \mathcal{F}_{t-}) \\ &= 1 - P(dN_t = 0 | \mathcal{F}_{t-}). \end{aligned}$$

Por último, sea N un proceso de conteo markoviano, entonces la ecuación (1.6) se reduce a

$$\lambda_t dt = E(dN_t | N_{t-}), \quad t \in [0, \tau].$$

1.6. Procesos puntuales regulares

A continuación describiremos un conjunto más específico de procesos puntuales a los cuales los llamaremos procesos puntuales regulares. Subsecuentemente, nos enfocaremos exclusivamente a procesos en éste conjunto.

Sea $T = \{T_n; n \in \mathbb{N}\}$ un proceso puntual definido sobre el espacio de probabilidad filtrado $(\Omega, \mathcal{A}, \mathcal{F}, P)$, con \mathcal{F} representando la historia del proceso, y $N = \{N_t; 0 \leq t \leq \tau\}$ su proceso de conteo asociado.

Definición 1.19 (Proceso puntual regular). Si $\lambda_t = \lambda_t(\omega)$, definido en (1.5), existe para cada realización $\omega \in \Omega$, es no negativa, continua a trozos sobre $[0, \tau]$ y continua por la izquierda en sus puntos de discontinuidad. Entonces el proceso T es llamado un *proceso puntual regular*.

Asimismo, N se dirá que es un proceso de conteo regular.

Para un proceso puntual regular, la probabilidad de que un evento ocurra o no en el intervalo $[t, t + h)$ estará dada por

$$\begin{aligned} P(N_{(t+h)-} - N_{t-} = 1 | \mathcal{F}_{t-}) &= \lambda_t h + o(h), \\ P(N_{(t+h)-} - N_{t-} = 0 | \mathcal{F}_{t-}) &= 1 - \lambda_t h + o(h), \end{aligned}$$

respectivamente. Donde

$$\lim_{h \downarrow 0} \frac{o(h)}{h} = 0.$$

Proposición 1.20. *Sea T un proceso puntual regular, entonces su función de intensidad determina su estructura de probabilidad de forma única.*

1.7. Función de supervivencia

La función de supervivencia condicional se define como la probabilidad de que a partir del instante en el cual ocurrió el último evento, al tiempo t no haya ocurrido ningún evento más. Para encontrar dicha probabilidad, empezaremos enunciando el siguiente lema.

Lema 1.21. *Sea N un proceso de conteo regular con función de intensidad λ_t . Entonces*

$$P(N_t = N_s | \mathcal{F}_s) = \exp \left[- \int_s^t \lambda_u du \right],$$

para todo $t \geq s \geq 0$.

Demostración. Dado un valor fijo de $s \geq 0$, iniciamos calculando la derivada de $P(N_u = N_s | \mathcal{F}_s)$ con respecto a $u \geq s$. Notemos que, dada la filtración \mathcal{F}_s , el valor de N_s será conocido. De este modo tenemos

$$\frac{d}{du} P(N_u = N_s | \mathcal{F}_s) = \lim_{h \downarrow 0} \frac{1}{h} [P(N_{u+h} = N_s | \mathcal{F}_s) - P(N_u = N_s | \mathcal{F}_s)], \quad (1.7)$$

con

$$\begin{aligned} P(N_{u+h} = N_s | \mathcal{F}_s) &= P(N_{u+h} = N_s | N_u = N_s, \mathcal{F}_s) P(N_u = N_s | \mathcal{F}_s) \\ &= P(N_{u+h} = N_u | \mathcal{F}_u) P(N_u = N_s | \mathcal{F}_s). \end{aligned} \quad (1.8)$$

Sustituyendo (1.8) en (1.7), nos queda

$$\begin{aligned} \frac{d}{du} P(N_u = N_s | \mathcal{F}_s) &= \lim_{h \downarrow 0} \frac{1}{h} [(P(N_{u+h} = N_u | \mathcal{F}_u) - 1) P(N_u = N_s | \mathcal{F}_s)] \\ &= \left[\lim_{h \downarrow 0} \frac{1}{h} P(N_{u+h} = N_u | \mathcal{F}_u) - 1 \right] P(N_u = N_s | \mathcal{F}_s) \\ &= -\lambda_{u+} P(N_u = N_s | \mathcal{F}_s). \end{aligned}$$

Con lo cual obtenemos

$$\frac{dP(N_u = N_s | \mathcal{F}_s)}{P(N_u = N_s | \mathcal{F}_s)} = -\lambda_{u+} du.$$

Integrando desde $u = s$ hasta $u = t$,

$$\log P(N_t = N_s | \mathcal{F}_s) - \log P(N_s = N_s | \mathcal{F}_s) = - \int_s^t \lambda_u du,$$

donde $\log P(N_s = N_s | \mathcal{F}_s) = \log 1 = 0$. Aplicando la función exponencial a ambos lados de esta expresión, concluimos que

$$P(N_t = N_s | \mathcal{F}_s) = \exp \left[- \int_s^t \lambda_u du \right],$$

que es la expresión a la que queríamos llegar. \square

Sea $T = \{T_n; n \in \mathbb{N}\}$ el proceso puntual asociado a N , y denotando por $\mathcal{F}_{(n)} = \mathcal{F}_{T_n}$ a la historia del proceso hasta la llegada del n -ésimo evento.

Definición 1.22. La *función de supervivencia* del $(n + 1)$ -ésimo evento está dada por

$$G_{n+1}(t) = P(T_{n+1} > t | \mathcal{F}_{(n)}), \quad t > T_n, \quad (1.9)$$

para $n \geq 0$ y con $T_0 \equiv 0$.

Notemos que, condicionado a $\mathcal{F}_{(n)}$, el evento $\{T_{n+1} > t\}$ es equivalente a $\{N_t = n\}$ para $t \geq T_n$. Así, utilizando el Lema 1.21, podemos demostrar el siguiente enunciado:

Lema 1.23. Sea $T_n = t_n$, entonces

$$G_{n+1}(t) = \exp \left[- \int_{t_n}^t \lambda_u du \right], \quad t > t_n. \quad (1.10)$$

Además, si denotamos por h_{n+1} a la función de densidad de T_{n+1} condicionada a $\mathcal{F}_{(n)}$, entonces

$$h_{n+1}(t | \mathcal{F}_{(n)}) = \lim_{h \downarrow 0} \frac{1}{h} P(T_{n+1} \in [t, t+h] | \mathcal{F}_{(n)}), \quad t > t_n, n \geq 0. \quad (1.11)$$

Considerando que

$$\begin{aligned} P(T_{n+1} \in [t, t+h] | \mathcal{F}_{(n)}) &= P(N_{t+h} = n+1, N_t = n | \mathcal{F}_{(n)}) \\ &= P(N_{t+h} - N_t = 1 | N_t = n, \mathcal{F}_{(n)}) P(N_t = n | \mathcal{F}_{(n)}) \\ &= P(N_{t+h} - N_t = 1 | \mathcal{F}_t) G_{n+1}(t), \end{aligned}$$

y sustituyendo en (1.11), tenemos

$$\begin{aligned} h_{n+1}(t | \mathcal{F}_{(n)}) &= \left[\lim_{h \downarrow 0} \frac{1}{h} P(N_{t+h} - N_t = 1 | \mathcal{F}_t) \right] G_{n+1}(t) \\ &= \lambda_t G_{n+1}(t), \quad t > t_n, \end{aligned} \quad (1.12)$$

para todo $n > 0$, con $t_0 \equiv 0$.

De esta manera, concluimos que la función de intensidad de un proceso puntual regular también puede ser representada a través de su función de supervivencia, de la siguiente forma:

$$\lambda_t = \frac{h_n(t | \mathcal{F}_{(n-1)})}{G_n(t)}, \quad T_{n-1} < t \leq T_n, \quad (1.13)$$

con T_n el tiempo de ocurrencia del n -ésimo evento y $T_0 \equiv 0$.

Notemos que $G_n(t) = \int_t^\infty h_n(s | \mathcal{F}_{(n-1)}) ds$, así

$$\lambda_t = h_n^*(t | \mathcal{F}_{(n-1)}), \quad T_{n-1} < t \leq T_n, \quad (1.14)$$

donde $h_n^*(t | \mathcal{F}_{(n-1)}) = \frac{h_n(t | \mathcal{F}_{(n-1)})}{\int_t^\infty h_n(s | \mathcal{F}_{(n-1)}) ds}$ es conocida como la *función de riesgo*.

1.8. Función de densidad conjunta

Sea T un proceso puntual regular asociado a N y definido sobre el espacio de probabilidad filtrado $(\Omega, \mathcal{A}, \mathcal{F}, P)$. Denotamos la función de distribución conjunta de nuestro proceso como

$$F_T(t_1, \dots, t_k) = P(T_1 \leq t_1, \dots, T_k \leq t_k),$$

para

$$0 < t_1 \leq \dots \leq t_k.$$

Lema 1.24. *Sea f_T la función de densidad asociada a F_T , entonces*

$$f_T(t_1, \dots, t_k) = \left(\prod_{i=1}^k \lambda_{t_i} \right) \exp(-\Lambda_{t_k}), \quad (1.15)$$

donde Λ es el compensador del proceso dado por $\Lambda_t = \int_0^t \lambda_s ds$.

Demostración. Denotemos por $\mathcal{F}_{(n)} = \mathcal{F}_{T_n}$. Utilizando probabilidad condicional, podemos escribir

$$f_T(t_1, \dots, t_k) = \prod_{n=1}^k h_n(t_n | \mathcal{F}_{(n-1)}), \quad (1.16)$$

con h_n representando la función de densidad de T_n dada la historia hasta T_{n-1} , expresada en (1.12).

Utilizando además la ecuación (1.10), tenemos que

$$h_{n+1}(t | \mathcal{F}_{(n)}) = \lambda_t \exp\left(-\int_{t_n}^t \lambda_u du\right),$$

para todo $n > 0$, con $T_0 \equiv 0$.

Sustituyendo en (1.16),

$$\begin{aligned} f_T(t_1, \dots, t_k) &= \prod_{n=1}^k \lambda_{t_n} \exp\left(-\int_{t_{n-1}}^{t_n} \lambda_u du\right) \\ &= \left(\prod_{n=1}^k \lambda_{t_n} \right) \exp\left(-\sum_{i=1}^k \int_{t_{i-1}}^{t_i} \lambda_u du\right) \\ &= \left(\prod_{n=1}^k \lambda_{t_n} \right) \exp\left(-\int_0^{t_k} \lambda_u du\right). \end{aligned}$$

Compleatando así la demostración. □

Por otro lado, si consideramos un intervalo fijo $[0, \tau]$, puede que sea de nuestro interés conocer la función de densidad conjunta de (N_τ, T) . Donde N_τ es el total de eventos ocurridos en dicho intervalo y $T = \{T_1, \dots, T_{N_\tau}\}$ son los tiempos en que los eventos suceden.

Teorema 1.25. *La función de densidad conjunta de un proceso puntual regular T sobre el intervalo $[0, \tau]$ está dada por*

$$f_{N_\tau, T}(k, t_1, \dots, t_k) = \left(\prod_{n=1}^k \lambda_{t_n} \right) \exp(-\Lambda_\tau), \quad (1.17)$$

para $k \in \mathbb{N}$, y $f_{N_\tau, T}(\cdot) = \exp(-\Lambda_\tau)$ si $k = 0$.

Demostración. Para $k \in \mathbb{N}$ arbitrario, tenemos que

$$\begin{aligned} f_{N_\tau, T}(k, t_1, \dots, t_k) &= P(N_\tau = k | T_1 = t_1, \dots, T_k = t_k) f_T(t_1, \dots, t_k) \\ &= G_{k+1}(\tau) f_T(t_1, \dots, t_k). \end{aligned}$$

Por (1.10) sabemos que

$$G_{k+1}(\tau) = \exp\left(-\int_{t_k}^{\tau} \lambda_u du\right),$$

y junto a (1.15), obtenemos

$$\begin{aligned} f_{N_\tau, T}(k, t_1, \dots, t_k) &= \left(\prod_{n=1}^k \lambda_{t_n} \right) \exp\left(-\Lambda_{t_k} - \int_{t_k}^{\tau} \lambda_u du\right) \\ &= \left(\prod_{n=1}^k \lambda_{t_n} \right) \exp(-\Lambda_\tau). \end{aligned}$$

Por otro lado, si $k = 0$, entonces simplemente

$$f_{N_\tau, T}(k) = P(N_\tau = N_0 | \mathcal{F}_0) = \exp\left(-\int_0^{\tau} \lambda_u du\right).$$

Completando así la demostración. □

Equivalentemente, tenemos que

$$\begin{aligned} \log [f_{N_\tau, T}(k, t_1, \dots, t_k)] &= \sum_{n=1}^k \log(\lambda_{t_n}) - \Lambda_\tau \\ &= \int_0^{\tau} \log(\lambda_t) dN_t - \Lambda_\tau. \end{aligned} \quad (1.18)$$

1.9. Procesos puntuales marcados

Ahora, supongamos que no estamos interesados en trabajar exclusivamente con el proceso puntual T , sino que éste solo representa un componente de un modelo más complejo. Consideremos que T sólo sirve para indicar el momento en que tiene lugar algún otro proceso estocástico, el cual cuenta con su propia estructura y puede ser independiente o no de nuestro proceso puntual.

Para modelar este escenario, en la presente sección introduciremos el concepto de *proceso puntual marcado*. La idea fundamental en la construcción de este modelo es asignar una variable aleatoria Z_n , a la cual llamaremos *marca*, al n -ésimo evento de un proceso puntual, para todo $n \in \mathbb{N}$.

1.9.1. Definiciones básicas

Sea (Q, \mathcal{Q}) un espacio medible. Supongamos que cada vez que ocurre un nuevo evento, la marca asociada tomará un valor en Q , por lo cual al espacio (Q, \mathcal{Q}) lo llamaremos el *espacio de marcas*. Asimismo, definiremos la marca ∇ , la cual asociaremos a los eventos que nunca ocurren, es decir, cuando $T_n = \infty$. Denotaremos por $\bar{Q} = Q \cup \{\nabla\}$ y $\bar{\mathcal{Q}} = \sigma(\mathcal{Q}, \{\nabla\})$.

Definición 1.26 (Proceso puntual marcado). Un *proceso puntual marcado* con espacio de marcas Q es una secuencia doble de variables aleatorias,

$$T = \{(T_n, Z_n); n \geq 1\},$$

definidas sobre algún espacio de probabilidad (Ω, \mathcal{A}, P) y con $T_n \in \bar{\mathbb{R}}_0$ y $Z_n \in \bar{Q}$, tal que $\bar{T} = \{T_n; n \geq 1\}$ forma un proceso puntual y además satisface que

- i) $P(Z_n \in Q, T_n < \infty) = P(T_n < \infty)$,
- ii) $P(Z_n = \nabla, T_n = \infty) = P(T_n = \infty)$,

para toda $n \in \mathbb{N}$.

Con el fin de proporcionar una mayor estructura a nuestro proceso, consideremos que Q corresponda a un espacio de Borel. Adicional a esto, es posible que la marca tenga dimensión mayor a uno, si este es el caso entonces diremos que el proceso será un proceso puntual *vector valuado*.

Al proceso marginal \bar{T} , que contiene los tiempos de ocurrencia de los eventos, lo nombraremos *proceso base*.

Definición 1.27. Un proceso puntual marcado es *simple* si y solo si su proceso base \bar{T} es simple.

Si denotamos por \mathbb{K}_Q al conjunto de secuencias definida por

$$\mathbb{K}_Q = \{ \{(t_n, z_n); n \geq 1\} \in \bar{\mathbb{R}}_0^{\mathbb{N}} \times \bar{Q}^{\mathbb{N}} : \{t_n, n \geq 1\} \in \mathbb{K}, z_n \in Q \text{ si y solo si } t_n < \infty \},$$

donde \mathbb{K} esta dado por (1.1). Y denotamos por \mathcal{K}_Q la σ -álgebra canónica de subconjuntos de \mathbb{K}_Q . Entonces, un proceso puntual marcado puede verse simplemente como una variable aleatoria definida sobre $(\mathbb{K}_Q, \mathcal{K}_Q)$.

Cabe señalar que un proceso puntual marcado puede verse también como un proceso puntual en el espacio producto $(\mathbb{R}_0 \times Q, \mathcal{B}_0 \otimes \mathcal{Q})$. De la Definición 1.3, una medida de conteo sobre este espacio se encuentra dada por

$$\mathcal{N}(C) = \sum_n \mathbb{1}_{\{(T_n, Z_n) \in C\}} = \sum_n \delta_{(S_n, T_n)}(C), \quad C \in \mathcal{B}_0 \otimes \mathcal{Q},$$

donde $\delta_{(T_n, Z_n)}(\cdot) = \delta_{(T_n(\omega), Z_n(\omega))}(\cdot; \omega)$ es la medida de Dirac con masa uno en el punto (T_n, Z_n) .

Notemos que, para toda medida de conteo μ , se satisface lo siguiente:

- i) $\mu(C) \in \bar{\mathbb{N}}_0, C \in \mathcal{B}_0 \otimes \mathcal{Q}$,
- ii) $\mu(\{0\} \times Q) = 0$,
- iii) $\mu(\{t\} \times Q) \in \{0, 1\}$,
- iv) $\mu([0, t] \times Q) < \infty$.

De esta forma, será posible caracterizar a nuestro proceso puntual marcado mediante una secuencia de medidas de conteo $N(A) = \{N_t(A); t \geq 0\}$, con

$$N_t(A) = \mathcal{N}([0, t] \times A) = \sum_n \mathbf{1}_{\{T_n \leq t, Z_n \in A\}},$$

para todo conjunto $A \in \mathcal{Q}$.

Finalmente, denotaremos por $\bar{N} = \{\bar{N}_t; t \geq 0\}$ al *proceso de conteo base*, en el cual \bar{N}_t representa el número total de eventos ocurridos en $[0, t]$ indistintamente de la marca asignada a cada evento, es decir,

$$\bar{N}_t = N_t(Q) = \sum_n \mathbf{1}_{\{T_n \leq t\}}.$$

1.9.2. Función de intensidad condicional para procesos marcados

La extensión a la definición de la función de intensidad para los procesos marcados es trivial siempre y cuando la dimensión principal siga representando el tiempo de ocurrencia de nuestros eventos.

Definición 1.28. Sea $T = \{(T_n, Z_n); n \geq 1\}$ un proceso puntual regular marcado con espacio de marcas Q e historia dada por la filtración $\{\mathcal{F}_t; t \geq 0\}$. Diremos que $\lambda = \{\lambda_t(z); z \in Q, t \geq 0\}$ es la función de intensidad condicional de T si

$$\lambda_t(z) dz dt = \mathbb{E}[\mathcal{N}(dt \times dz) | \mathcal{F}_{t-}],$$

donde $\mathcal{N}(dt \times dz)$ representará el número de eventos ocurridos en el intervalo $[t, t + dt)$ y con marca en $[z, z + dz)$.

Además, si denotamos por $h_n(t, z | \mathcal{F}_{(n-1)})$ la función de densidad conjunta de (T_n, Z_n) , condicionada a la historia del proceso hasta el $(n - 1)$ -ésimo evento. Entonces, análogo a la fórmula (1.13), la función de intensidad condicional estará dada por

$$\lambda_t(z) = \frac{h_n(t, z | \mathcal{F}_{(n-1)})}{\int_t^\infty \int_Q h_n(s, z | \mathcal{F}_{(n-1)}) \ell_Q(dz) ds}, \quad T_{n-1} < t \leq T_n,$$

con $\mathcal{F}_{(n)} = \mathcal{F}_{T_n}$ y donde ℓ_Q representa una medida de referencia de Q , la cual comúnmente es la medida de Lebesgue cuando Q es algún espacio euclidiano y una medida de conteo cuando es un espacio discreto.

Si denotamos por $\bar{h}_n(t|\mathcal{F}_{(n-1)})$ a la función de densidad marginal de T_n condicionada a $\mathcal{F}_{(n-1)}$, y $f_{Z_n}(z|\mathcal{F}_{t-})$ a la función de densidad condicional de la marca Z_n cuando esta se localiza en t . Entonces

$$\lambda_t(z) = \frac{\bar{h}_n(t|\mathcal{F}_{(n-1)}) f_{Z_n}(z|\mathcal{F}_{t-})}{\int_t^\infty \bar{h}_n(s|\mathcal{F}_{(n-1)}) ds},$$

y tomando \bar{h}_n^* como la función de riesgo condicional, dada por

$$\bar{h}_n^*(t|\mathcal{F}_{(n-1)}) = \frac{\bar{h}_n(t|\mathcal{F}_{(n-1)})}{\int_t^\infty \bar{h}_n(s|\mathcal{F}_{(n-1)}) ds},$$

obtenemos otra expresión para denotar $\lambda_t(z)$, la cual será

$$\lambda_t(z) = \bar{h}_n^*(t|\mathcal{F}_{(n-1)}) f_{Z_n}(z|\mathcal{F}_{t-}), \quad T_{n-1} < t \leq T_n,$$

con $T_0 \equiv 0$.

1.9.3. Clasificación de procesos puntuales marcados

Consideremos un proceso de conteo marcado N con espacio de marcas Q y definido en el espacio de probabilidad (Ω, \mathcal{A}, P) . Denotemos por \mathcal{F}_t^T a la historia del proceso base subyacente y \mathcal{F}_t^Z a la historia de la secuencia de marcas, es decir, las marcas que han ocurrido hasta el tiempo t sin especificar el instante en el que ocurrieron. Además, representemos la historia completa del proceso hasta el tiempo t mediante la filtración $\mathcal{F}_t = (\mathcal{F}_t^T \cup \mathcal{F}_t^Z)$.

Sea λ_t la función de intensidad del proceso y F_{Z_n} la distribución de la n -ésima marca.

Definición 1.29. Diremos que el proceso cuenta con:

- i) *Marcas dependientes* si F_{Z_n} depende del instante en el que ocurrieron las marcas anterior pero no necesariamente del valor de estas, es decir, si depende de $\mathcal{F}_{T_{n-1}}^T$.
- ii) *Marcas auto-dependientes* si esta distribución depende del valor de las marcas anteriores, $\mathcal{F}_{T_{n-1}}^Z$, pero no necesariamente de su ubicación.
- iii) *Marcas independientes* si la distribución de la n -ésima marca no depende ni de $\mathcal{F}_{T_{n-1}}^T$ ni de $\mathcal{F}_{T_{n-1}}^Z$.

Adicionalmente, si λ_t depende de $\mathcal{F}_{T_{n-1}}^Z$, entonces diremos que el proceso se encuentra *controlado* por las marcas.

1.9.4. El proceso de Poisson compuesto

Así como el proceso de Poisson resulta ser el proceso puntual más simple, para el caso de procesos puntuales marcados es el proceso de Poisson compuesto el ejemplo más sencillo de brindar. Este se forma al tomar como proceso base a un proceso de Poisson con tasa λ , $T = \{T_n, n \in \mathbb{N}\}$, y asignarle a cada evento una determinada marca, Z_n , independiente de T . Dichas marcas formarán una secuencia de variables aleatorias independientes entre sí e idénticamente distribuidas, que siguen cierta distribución F_Z conocida.

1.10. Procesos puntuales multivariados

Hasta este momento hemos trabajamos con procesos que modelan la dispersión de puntos, los cuales representaban el momento en que ocurría cierto tipo de eventos y sólo eran distinguibles por su ubicación en la recta. Ahora consideraremos la existencia de dos o más tipos diferentes de eventos.

Definición 1.30. Un *proceso puntual multivariado* es un proceso puntual marcado, $T = \{(T_n, Z_n); n \geq 1\}$, con espacio de marcas finito o infinito numerable $Q = \{1, 2, \dots\}$.

Asimismo, denotamos los procesos de conteo marginales como

$$N_t^{(i)} = \mathcal{N}([0, t] \times \{i\}) = \sum_n \mathbb{1}_{\{T_n \leq t, Z_n = i\}},$$

y las funciones de intensidad asociadas

$$\lambda_t^{(i)} = \frac{\mathbf{E} \left(dN_t^{(i)} \mid \mathcal{F}_{t-} \right)}{dt},$$

para $i \in Q$.

Definición 1.31. El *proceso de conteo multivariado* asociado a T estará dado por

$$N = \left\{ N_t^{(i)}; i \in Q, t \geq 0 \right\},$$

Además, definimos el *proceso de conteo base* mediante $\bar{N} = \{\bar{N}_t; t \geq 0\}$, con

$$\bar{N}_t = \mathcal{N}([0, t] \times Q) = \sum_{i \in Q} N_t^{(i)}.$$

Sin pérdida de generalidad, consideremos que solo existen dos tipos diferentes de eventos, $Q = \{1, 2\}$, por lo cual tenemos el proceso de conteo *bivariado*

$$N = \left\{ N_t^{(1)}, N_t^{(2)}; t \geq 0 \right\},$$

donde $N_t^{(i)}$ representa el número de eventos del tipo i que han ocurrido en $[0, t]$, para $i = 1, 2$.

Sea $\mathcal{F}_t^{(i)}$ la historia del proceso de tipo i , y $\mathcal{F}_t = \mathcal{F}_t^{(1)} \cup \mathcal{F}_t^{(2)}$ la historia del proceso completo. Supongamos además que N es simple, entonces su función de intensidad estará dada por $\lambda = \left\{ \lambda_t^{(1)}, \lambda_t^{(2)}; t \geq 0 \right\}$, con

$$\lambda_t^{(1)} = \lim_{h \downarrow 0} \frac{1}{h} \mathbf{P} \left(N_{(t+h)-}^{(1)} - N_{t-}^{(1)} = 1 \mid \mathcal{F}_{t-} \right),$$

$$\lambda_t^{(2)} = \lim_{h \downarrow 0} \frac{1}{h} \mathbf{P} \left(N_{(t+h)-}^{(2)} - N_{t-}^{(2)} = 1 \mid \mathcal{F}_{t-} \right).$$

Proposición 1.32. Sea N un proceso de conteo marcado con espacio de marcas (Q, \mathcal{Q}) . Entonces, para conjuntos disjuntos a pares $A_1, \dots, A_d \in \mathcal{Q}$, tenemos que el proceso $\{N_t(A_1), \dots, N_t(A_d)\}$ forma un proceso de conteo multivariado de dimensión d .

2

Estadística aplicada a los procesos puntuales

En éste capítulo abordaremos los problemas de inferencia estadística aplicados a los procesos puntuales. Para ello, trataremos temas como estimación, pruebas de hipótesis y simulación, entre otros.

En los problemas de estimación, trabajaremos para el caso paramétrico, por lo cual supondremos que las leyes de los procesos puntuales se encuentran completamente especificadas, excepto por una cantidad finita de parámetros.

2.1. Estimadores de máxima verosimilitud

2.1.1. Estimación para procesos univariados

Sea $T = \{T_n; n \in \mathbb{N}\}$ un proceso puntual regular con función de intensidad λ_t . Supongamos que esta función depende de cierto vector de parámetros $\theta = (\theta_1, \dots, \theta_r) \in \Theta$, de modo que podemos representar $\lambda_t = \lambda(t; \theta)$. Supongamos además que observamos cierta realización del proceso en el intervalo de tiempo finito $[0, \tau]$.

Teorema 2.1. *Sea $\{t_1, \dots, t_k\}$ cierta realización de T , con $0 < t_1 \leq \dots \leq t_k \leq \tau$. La función de verosimilitud del proceso estará dada por*

$$L(\theta; t_1, \dots, t_k) = \left[\prod_{n=1}^k \lambda(t_n; \theta) \right] \exp[-\Lambda(\tau; \theta)], \quad (2.1)$$

donde Λ representa el compensador del proceso definido por $\Lambda(t; \theta) = \int_0^t \lambda(s; \theta) ds$.

Demostración. La función L es obtenida a partir de la función de densidad conjunta de N_τ y T ,

$$L(\theta; t_1, \dots, t_k) = f_{N_\tau, T}(k, t_1, \dots, t_k; \theta),$$

donde $f_{N_\tau, T}$ está dada por (1.18). □

Asimismo, la función de log-verosimilitud se expresa a través de la fórmula

$$\begin{aligned} l(\theta; t_1, \dots, t_k) &= \sum_{i=1}^k \log[\lambda(t_i; \theta)] - \Lambda(\tau; \theta). \\ &= \int_0^\tau \log[\lambda(t; \theta)] dN_t - \Lambda(\tau; \theta). \end{aligned}$$

Por lo tanto, el estimador de máxima verosimilitud de θ estará dado por

$$\hat{\theta} = \underset{\theta \in \Theta}{\text{máx arg}} l(\theta; t_1, \dots, t_k). \quad (2.2)$$

Más aún, contrastando con un proceso de Poisson de tasa uno definido en $[0, \tau]$, la función log-razón de verosimilitud del proceso será

$$\log\left(\frac{L^1}{L(\theta)}\right) = \int_0^\tau [\lambda(t; \theta) - 1] dt - \int_0^\tau \log \lambda(t; \theta) dN(t),$$

donde L^1 representa la función de verosimilitud evaluada en $\lambda_t = 1$.

De esta manera, si nosotros estamos interesados en realizar una prueba de hipótesis en la cual $H_0 : \lambda_t = 1$ contra $H_1 : \lambda_t \neq 1$, entonces podemos hacer uso de una prueba por log-razón de verosimilitud, tomando el estadístico:

$$r(t_1, \dots, t_k) = \int_0^\tau [\lambda(t; \hat{\theta}) - 1] dt - \int_0^\tau \log \lambda(t; \hat{\theta}) dN(t),$$

donde $\hat{\theta} = \hat{\theta}(t_1, \dots, t_k)$ está dado mediante (2.2).

2.1.2. Estimación para el proceso de Poisson homogéneo

En la Sección 1.3 definimos al proceso de Poisson homogéneo, ahora abordaremos los problemas de inferencia referentes a éste modelo.

Sea $T = \{T_n; n \in \mathbb{N}\}$ un proceso de Poisson de tasa desconocida $\lambda > 0$, y N el proceso de conteo asociado a T . Denotamos por $\mathcal{F} = \{\mathcal{F}_t, t \geq 0\}$ a la historia del proceso, con $\mathcal{F}_t = \sigma(N_s, 0 \leq s \leq t)$. Supongamos además que el proceso es observado hasta un tiempo τ , generando la realización $\{t_1, \dots, t_k\}$.

Si utilizamos la intensidad del proceso para obtener la función de log-verosimilitud, tal y como se describió en la sección anterior, tenemos que esta estará dada por

$$l(\lambda; t_1, \dots, t_k) = k \log \lambda - \lambda \tau,$$

con lo cual, al maximizar en función de λ , obtenemos el estimador

$$\hat{\lambda} = \frac{k}{\tau}. \quad (2.3)$$

Cabe destacar que es posible realizar pruebas de hipótesis considerando, por ejemplo, las hipótesis simples

$$H_0 : \lambda = \lambda_0$$

$$H_1 : \lambda = \lambda_1,$$

mediante la razón de log-verosimilitud:

$$\log \left(\frac{L(\lambda_0)}{L(\lambda_1)} \right) = k \log \left(\frac{\lambda_0}{\lambda_1} \right) - \tau (\lambda_0 - \lambda_1).$$

O tomar hipótesis compuestas tales como

$$\begin{aligned} H_0 : \lambda &= \lambda_0 \\ H_1 : \lambda &\neq \lambda_0, \end{aligned}$$

considerando el estadístico

$$\log \left(\frac{L(\lambda_0)}{L(\hat{\lambda})} \right) = k \log \left(\frac{\lambda_0}{\hat{\lambda}} \right) - \tau (\lambda_0 - \hat{\lambda}),$$

con $\hat{\lambda}$ el estimador de máxima verosimilitud propuesto en (2.3).

No obstante, el proceso de Poisson homogéneo posee características específicas que podemos emplear para encontrar caminos alternativos que nos lleven al estimador de máxima verosimilitud de λ . Estas características se describen a continuación.

Primero, supongamos que τ corresponde al momento en que ocurre el k -ésimo evento, $\tau = T_k$, en éste caso diremos que nuestros datos son *sincrónicos*. Recordando que los tiempos entre llegadas W_1, \dots, W_k , donde $W_n = T_n - T_{n-1}$ para $n \geq 1$ y con $T_0 \equiv 0$, corresponden a variables aleatorias independientes con distribución común exponencial de tasa λ , entonces tenemos la siguiente proposición:

Proposición 2.2. *El estimador de máxima verosimilitud de λ (dado \mathcal{F}_τ) será*

$$\hat{\lambda} = \frac{k}{\sum_{n=1}^k W_n}. \quad (2.4)$$

Se puede demostrar que éste estimador converge a λ casi seguramente, mientras que $\sqrt{k} (\lambda - \hat{\lambda}) / \lambda$ converge en distribución a una normal estándar. Además, dado que $\sum_{n=1}^k W_n$ se distribuye Gama de parámetros k y $1/\lambda$, será posible crear intervalos de confianza para el verdadero valor de λ .

Por otro lado, supongamos que nuestro tiempo de observación τ no corresponde al momento de ocurrencia de algún evento, en éste caso diremos que se tienen datos *asincrónicos*. Recordando que el número de eventos ocurridos en $[0, t]$ sigue una distribución Poisson de parámetro λt para todo $t > 0$, entonces:

Proposición 2.3. *El estimador de máxima verosimilitud de λ (dado \mathcal{F}_τ) será*

$$\hat{\lambda} = \frac{k}{\tau}. \quad (2.5)$$

Dicho estimador también convergerá a λ casi seguramente, mientras que $\sqrt{\tau} (\lambda - \hat{\lambda}) / \lambda$ converge en distribución a una variable normal estándar. Notemos que el estimador es idéntico al obtenido en (2.3).

2.1.3. Estimación para procesos marcados y multivariados

Sea T un proceso puntual marcado con espacio de marcas Q y función de intensidad $\lambda(t, z; \theta)$, donde $\theta \in \Theta$ es un vector de parámetros. Suponiendo que observamos el proceso en el intervalo $[0, \tau]$, entonces tenemos el siguiente teorema:

Teorema 2.4. *Sean $(t_1, z_1), \dots, (t_k, z_k)$ los tiempos y marcas en cierta realización de T , con $0 < t_1 \leq \dots \leq t_k \leq \tau$. La función de verosimilitud del proceso estará dada por*

$$L(\theta; t_1, \dots, t_k, z_1, \dots, z_k) = \left[\prod_{n=1}^k \lambda(t_n, z_n; \theta) \right] \exp \left[- \int_0^\tau \int_Q \lambda(t, z; \theta) \ell_Q(dz) dt \right],$$

donde ℓ_Q representa la una medida de referencia para Q .

Por lo que aplicando logaritmo, su función de log-verosimilitud es

$$l(\theta; t_1, \dots, t_k, z_1, \dots, z_k) = \sum_{n=1}^k \log[\lambda(t_n, z_n; \theta)] - \int_0^\tau \int_Q \lambda(t, z; \theta) \ell_Q(dz) dt.$$

Además, sea T^* un proceso de Poisson compuesto en el intervalo $[0, \tau]$, de tasa μ y marcas auto-independientes con distribución común y función de densidad π . La función de log-razón de verosimilitud relativa a T^* estará dada por

$$\log \left(\frac{L(\theta)}{L^*} \right) = \sum_{n=1}^k \log \left[\frac{\lambda(t_n, z_n; \theta)}{\mu \pi(z_n)} \right] - \int_0^\tau \int_Q [\lambda(t, z; \theta) - \mu \pi(z)] \ell_Q(dz) dt.$$

Corolario 2.5. *Sea T un proceso puntual multivariado de dimensión d , y cuya función de intensidad está dada por $\lambda = \{\lambda^{(i)}(t; \theta); i \in \{1, \dots, d\}, t \geq 0\}$. Sea $\{(t_1, z_1), \dots, (t_k, z_k)\}$ una realización de T , entonces su función de verosimilitud estará dada por*

$$L(\theta; t_1, \dots, t_k, z_1, \dots, z_k) = \left[\prod_{n=1}^k \lambda^{(z_n)}(t_n; \theta) \right] \prod_{i=1}^d \exp \left[- \int_0^\tau \lambda^{(i)}(t; \theta) dt \right].$$

2.2. Transformación de tiempos aleatorios

A continuación presentaremos un método para convertir cierto proceso puntual regular T en un proceso de Poisson homogéneo, y viceversa. Esto será útil para simular realizaciones de dicho proceso T , así como para realizar pruebas de bondad de ajuste para nuestras estimaciones.

Primero comenzaremos dando el siguiente lema, el cual es conocido también como el método de la transformada inversa, y que es utilizado para generar variables aleatorias con cierta distribución (exponencial en nuestro caso), a través de variables continuas que siguen una distribución F conocida.

Lema 2.6. (i) *Sea X una variable aleatoria con función de distribución continua F y $H(x) = -\log(1 - F(x))$ una función monótona invertible. Entonces $Y = H(X)$ sigue una distribución exponencial de parámetro uno. (ii) De manera inversa, si Y es una variable aleatoria con distribución exponencial de parámetro uno, entonces $X = H^{-1}(Y)$ tiene distribución F .*

Demostración. Para demostrar (i) basta notar que $H^{-1}(y) = F^{-1}(1 - e^{-y})$, y por lo tanto

$$\begin{aligned} \Pr(Y \leq y) &= \Pr(H(X) \leq y) = \Pr(X \leq H^{-1}(y)) \\ &= F[H^{-1}(y)] = F[F^{-1}(1 - e^{-y})] = 1 - e^{-y}, \end{aligned}$$

la cual es la función de distribución de una variable exponencial de media uno. Procediendo de manera inversa se puede demostrar (ii). \square

Teorema 2.7 (Cambio de tiempo aleatorio). *Sea $\{T_n; n \in \mathbb{N}\}$ un proceso puntual regular adaptado a la filtración $\mathcal{F} = \{\mathcal{F}_t; t \geq 0\}$, con función de intensidad acotada λ_t y \mathcal{F} -compensador $\Lambda(t) = \int_0^t \lambda_s ds$, tal que $\Lambda(t) \rightarrow \infty$ cuando $t \rightarrow \infty$ (c.s.). Entonces, con probabilidad uno, la secuencia*

$$T^* = \{\Lambda(T_n); n \in \mathbb{N}\},$$

corresponderá a tiempos aleatorios de un proceso de Poisson de tasa uno.

De manera inversa, ahora consideramos un proceso de Poisson de tasa uno, $\{T_n^; n \in \mathbb{N}\}$, con historia dada por la filtración $\mathcal{F}^* = \{\mathcal{F}^*(t); t \geq 0\}$. Sea $\Lambda = \{\Lambda(t); t \geq 0\}$ un proceso \mathcal{F}^* -adaptado, finito, monotamente creciente y con trayectorias continuas. Entonces*

$$T = \{\Lambda^{-1}(T_n^*); n \in \mathbb{N}\}$$

será un proceso puntual simple adaptado a la filtración $\mathcal{F} = \{\mathcal{F}^(\Lambda(t)); t \geq 0\}$ y con \mathcal{F} -compensador Λ .*

Demostración. Consideremos los tiempos entre llegadas, W_1, W_2, \dots , definidos por

$$W_n = T_n - T_{n-1}, \quad n \in \mathbb{N},$$

con $T_0 \equiv 0$.

La función de distribución de cada una de estas variables, condicionada a su pasado, estará dada por

$$\begin{aligned} F_{W_n}(w) &= \Pr(W_n \leq w | \mathcal{F}_{t_{n-1}}) = \Pr(T_n \leq t_{n-1} + w | \mathcal{F}_{t_{n-1}}) \\ &= 1 - \Pr(T_n > t_{n-1} + w | \mathcal{F}_{t_{n-1}}) = 1 - G_n(t_{n-1} + w) \\ &= 1 - \exp\left[-\int_{t_{n-1}}^{t_{n-1}+w} \lambda_s ds\right], \end{aligned}$$

con $t_0 \equiv 0$, y donde G_n representa la función de supervivencia condicionada a la información hasta t_{n-1} y definida en (1.9).

Sea

$$H_n(w) = -\log(1 - F_{W_n}(w)) = \int_{t_{n-1}}^{t_{n-1}+w} \lambda_s ds = \Lambda(w + t_{n-1}) - \Lambda(t_{n-1}), \quad (2.6)$$

aplicando (i) del Lema 2.6, se puede demostrar que las variables W_1^*, W_2^*, \dots , con

$$\begin{aligned} W_n^* &= H_n(W_n) = \Lambda(W_n + t_{n-1}) - \Lambda(t_{n-1}) \\ &= \Lambda(T_n) - \Lambda(T_{n-1}), \quad n \in \mathbb{N}, \end{aligned}$$

forman una secuencia de variables aleatorias con distribución exponencial de parámetro uno, y que además son independientes.

De esto se sigue que la sucesión

$$\{\Lambda(T_1), \Lambda(T_2), \dots\}$$

represente los instantes de tiempo en que ocurren los eventos en un proceso de Poisson de tasa uno. \square

Sea $N = \{N(t); t \geq 0\}$ el proceso de conteo asociado a T , del teorema anterior se sigue que

$$N^* = \{N(\Lambda^{-1}(t)); t \geq 0\}$$

forma un proceso de conteo asociado a un proceso de Poisson de tasa uno.

Concluimos que todo proceso puntual simple con compensador continuo puede ser generado a partir de un proceso de Poisson de tasa uno. Es por esto que a éste tipo de procesos suele llamarse *procesos de tipo Poisson*. Es posible extender este método para aplicarlo tanto a procesos puntuales marcados como a multivariados. Esta extensión se describe en el siguiente teorema y su corolario.

Teorema 2.8. *Sea $\{(T_n, Z_n); n \in \mathbb{N}\}$ un proceso puntual marcado en el intervalo $[0, \tau]$ y con espacio de marcas (Q, \mathcal{Q}) , donde ℓ_Q representa una medida de probabilidad de referencia para dicho espacio, y denotemos la historia del proceso mediante la filtración \mathcal{F} . Suponiendo que existe la función de intensidad del proceso, $\lambda = \{\lambda(t, z); t \geq 0, z \in Q\}$, tal que sea \mathcal{F} -predecible, estrictamente positiva y continua por la izquierda. Y considerando que el compensador $\Lambda(t, z) = \int_0^t \lambda(s, z) ds \rightarrow \infty$ cuando $t \rightarrow \infty$ (c.s.). Entonces bajo la transformación $(t, z) \rightarrow \Lambda(t, z)$, el proceso*

$$\{(\Lambda(T_n, Z_n), Z_n); n \in \mathbb{N}\}$$

forma un proceso de Poisson compuesto de tasa base uno y cuyas marcas tienen distribución ℓ_Q .

Corolario 2.9. *Sea $\{N^{(i)}(t); i \in \{1, \dots, d\}, 0 \leq t \leq \tau\}$ un proceso de conteo multivariado de dimensión finita d en el intervalo $[0, \tau]$. Sea $\lambda^{(i)}(t)$ la función de intensidad asociada al i -ésimo componente, y suponiendo que $\Lambda^{(i)}(t) = \int_0^t \lambda^{(i)}(s) ds \rightarrow \infty$ cuando $t \rightarrow \infty$ (c.s.), para todo $i \in \{1, \dots, d\}$. Entonces*

$$\left\{ N^{(i)} \left(\left(\Lambda^{(i)} \right)^{-1} (t) \right); i \in \{1, \dots, d\}, 0 \leq t \leq \tau \right\}$$

forma un proceso de conteo asociado a un proceso de Poisson multivariado con componentes independientes y de tasa uno.

2.3. Bondad de ajuste

El Teorema 2.7 será esencial en el denominado *análisis de residuos para procesos puntuales* que presentaremos en esta sección. La idea de dicho análisis es reducir un proceso

puntual regular a un proceso de Poisson de tasa uno y realizar pruebas de ajuste a este modelo, el cual es más simple y ha sido extensamente estudiado.

Partimos de una colección de puntos $\{t_1, \dots, t_k\}$ en el intervalo $[0, \tau]$, tal que $0 < t_1 \leq \dots \leq t_k \leq \tau$. Supongamos que estamos interesados en saber si estos puntos corresponden a los eventos generados por un proceso puntual con compensador $\Lambda(t)$. Aplicando el cambio de tiempos aleatorios descrito en la sección anterior, y si el proceso en efecto tiene compensador $\Lambda(t)$, entonces la secuencia $\{t_1^*, \dots, t_k^*\}$, con $t_n^* = \Lambda(t_n)$ para $n \in \{1, \dots, k\}$, son los tiempos generados por un proceso de Poisson de tasa uno en el intervalo $[0, \tau^*]$, con $\tau^* = \Lambda(\tau)$.

2.3.1. Pruebas de uniformidad

Supongamos que nuestros datos son asincrónicos, y por lo tanto $\tau \neq T_k$, un primer análisis que podemos realizar será revisar el conjunto $\{V_1, \dots, V_k\}$, con $V_i = T_i^*/\tau^*$. Supongamos además que V_1, \dots, V_k representan las estadísticas de orden de cierta muestra aleatoria, $\{U_1, \dots, U_k\}$, de una variable U con distribución F_U .

Si los tiempos T_1^*, \dots, T_k^* son realmente generados por un proceso de Poisson de tasa uno, entonces U_1, \dots, U_k serán variables aleatorias independientes cuya distribución está dada por F , donde

$$F(u) = \begin{cases} 0, & \text{si } u < 0, \\ u, & \text{si } 0 \leq u \leq 1, \\ 1, & \text{si } u > 1, \end{cases}$$

es decir, $U \sim \text{Uniforme}(0, 1)$.

De esta forma, nos interesa el problema de prueba de hipótesis cuya hipótesis nula será

$$H_0 : F_U = F$$

contra la alternativa

$$H_1 : F_U \neq F.$$

Ahora, siendo F_k la función de distribución empírica de U , podemos utilizar el estadístico de Kolmogorov-Smirnov dado por

$$D_k = \sup_{0 \leq u \leq 1} |F_k(u) - F(u)| = \max_{i \in \{1, \dots, k\}} \left| \frac{i}{k} - V_i \right|,$$

el estadístico de Cramér-von Mises

$$\omega_k^2 = k \int_0^1 [F_k(u) - F(u)]^2 du = \frac{1}{12k} + \sum_{i=1}^k \left[\frac{2i-1}{2k} - V_i \right]^2,$$

o el de Anderson-Darling

$$\begin{aligned} A_k^2 &= k \int_0^1 \left[\frac{F_k(u) - F(u)}{\sqrt{u(1-u)}} \right]^2 du \\ &= -k - \frac{1}{k} \sum_{i=1}^k [(2i-1) \log V_i + [2(k-i)+1] \log(1-V_i)], \end{aligned}$$

para completar nuestra prueba. La distribución de cualquiera de estos tres estadísticos es libre de parámetros, con lo cual podemos crear intervalos de aceptación y rechazo para H_0 en función del valor que tomen los estadísticos. Sin embargo, se sabe que para distribuciones uniformes, la prueba que tiene mayor potencia entre estas tres es la de Anderson-Darling, por lo cual es la primera que se sugeriría hacer. Para consultar mayor información referente a estas pruebas, se puede consultar [2].

Por otro lado, si los datos que tenemos son sincrónicos, i.e. $\tau = T_k$, entonces podemos analizar la sucesión de variables $\{V_1, \dots, V_{k-1}\}$ con $V_i = T_i^*/\tau^*$. Dicha sucesión también representará una muestra aleatoria ordenada de cierta variable uniforme en $[0, 1]$ si los tiempos $\{T_1^*, \dots, T_k^*\}$ forman un proceso de Poisson de tasa uno. De esta manera, se repite la metodología anterior pero para una secuencia con $k - 1$ elementos.

2.3.2. Pruebas de exponencialidad

Alternativamente, se puede revisar el conjunto de los tiempos entre eventos $W^* = \{W_1^*, \dots, W_k^*\}$, con $W_n^* = T_n^* - T_{n-1}^*$ y donde $T_0 \equiv 0$. Si los tiempos T_1^*, \dots, T_k^* provienen de un proceso de Poisson de tasa uno, entonces W_1^*, \dots, W_k^* será una muestra de variables aleatorias independientes e idénticamente distribuidas con distribución exponencial y media uno. Se pueden realizar las pruebas de Kolmogorov-Smirnov, Crámer-von Mises y Anderson-Darling para hacer pruebas de hipótesis respecto a este escenario.

2.3.3. Prueba ji-cuadrada

Antes de presentar esta prueba, empezaremos enunciando el siguiente lema:

Lema 2.10. Sean X_1, \dots, X_r variables aleatorias independientes con distribución de Poisson y media μ_1, \dots, μ_r respectivamente. Si denotamos con Y a la suma de dichas variables, $Y = \sum_{i=1}^r X_i$, entonces el vector aleatorio (X_1, \dots, X_r) condicionado a $Y = y$ sigue una distribución multinomial de parámetros $\{y, p_1, \dots, p_r\}$, donde $p_i = \mu_i / \sum_{j=1}^r \mu_j$ para $i = 1, \dots, r$.

Demostración. Es bien sabido que la suma de variables Poisson también sigue una distribución Poisson, cuyo parámetro será la suma de los parámetros de las distribuciones de los sumandos. De esta manera, calculando la función de probabilidad de (X_1, \dots, X_k) condicionada a Y , tenemos que

$$\begin{aligned} \Pr(X_1 = x_1, \dots, X_k = x_k | Y = y) &= \frac{\Pr(X_1 = x_1, \dots, X_k = x_k, Y = y)}{\Pr(Y = y)} \\ &= \frac{\Pr(X_1 = x_1) \cdots \Pr(X_k = x_k) \mathbb{1}_{\{\sum_{i=1}^k x_i = y\}}}{\Pr(Y = y)} \\ &= \left(\frac{\mu_1^{x_1} e^{-\mu_1}}{x_1!} \right) \cdots \left(\frac{\mu_k^{x_k} e^{-\mu_k}}{x_k!} \right) \left(\frac{y!}{\mu^{(x_1 + \dots + x_k)} e^{-\mu}} \right) \mathbb{1}_{\{\sum_{i=1}^k x_i = y\}} \\ &= \left(\frac{\mu_1}{\mu} \right)^{x_1} \cdots \left(\frac{\mu_k}{\mu} \right)^{x_k} \left(\frac{y!}{x_1! \cdots x_k!} \right) \mathbb{1}_{\{\sum_{i=1}^k x_i = y\}}, \end{aligned}$$

donde $\mu = \sum_{j=1}^r \mu_j$. Observamos que esta expresión corresponde a la función de probabilidad de una variable multinomial de parámetros y y $(\mu_1/\mu, \dots, \mu_k/\mu)$, completando así la demostración. \square

Ahora, sea $T = \{T_n; n \in \mathbb{N}\}$ un proceso de Poisson dentro del intervalo $[0, \tau^*]$ y definido sobre el espacio de probabilidad (Ω, \mathcal{A}, P) . Si particionamos al intervalo $(0, \tau^*]$ en r subintervalos disjuntos A_1, \dots, A_r , con

$$A_i = \left(\frac{(i-1)\tau^*}{r}, \frac{i\tau^*}{r} \right], \quad i = 1, \dots, r,$$

y tomamos la función $\mathcal{N}(A) = \sum_n \mathbb{1}_{\{T_n \in A\}}$, para todo $A \in \mathcal{B}([0, \tau^*])$. Entonces, sabemos que las variables aleatorias K_1, \dots, K_r , con $K_i = \mathcal{N}(A_i)$, son independientes y siguen una distribución de Poisson. Denotamos por $\lambda_1, \dots, \lambda_r$ al parámetro asociado a la distribución de cada una de estas variables.

Por el Lema 2.10, condicionado a que se tiene un total de K eventos en el intervalo $[0, \tau^*]$, el vector (K_1, \dots, K_r) sigue una distribución multinomial de parámetros K y (p_1, \dots, p_r) , con $p_i = \lambda_i / \sum_{j=1}^r \lambda_j$.

Supongamos que (k_1, \dots, k_r) corresponde a una realización de (K_1, \dots, K_r) , con $\sum_{j=1}^r k_j = k$, y que desconocemos el verdadero valor de los parámetros p_1, \dots, p_r . En este caso, considerando que la muestra es multinomial, tenemos que su función de verosimilitud estará dada por

$$\begin{aligned} L(p_1, \dots, p_r; k_1, \dots, k_r) &= p_1^{k_1} \cdots p_r^{k_r} \left(\frac{k!}{k_1! \cdots k_r!} \right) \\ &\propto p_1^{k_1} \cdots p_r^{k_r}, \end{aligned}$$

condicionada a que $\sum_{j=1}^r p_j = 1$.

Notemos que si S proviniera de un proceso de Poisson homogéneo de parámetro λ , entonces $p_i = 1/r$ para todo i . De esta forma, estaremos interesados en aplicar una prueba de hipótesis considerando como hipótesis nula:

$$H_0 : p_1 = \cdots = p_r = \frac{1}{r},$$

la cual será una hipótesis simple. La hipótesis alternativa estará compuesta por todos los casos en que no se cumpla H_0 .

Consideremos la función de razón de verosimilitud, dada por

$$R(k_1, \dots, k_r) = \frac{L\left(\frac{1}{r}, \dots, \frac{1}{r}; k_1, \dots, k_r\right)}{L(\hat{p}_1, \dots, \hat{p}_r; k_1, \dots, k_r)},$$

donde \hat{p}_i representa el estimador de máxima verosimilitud de p_i en todo el espacio paramétrico.

Como sólo se cuenta con una muestra, se puede demostrar que $\hat{p}_i = k_i/k$ para todo

$i = 1, \dots, r$. Por lo tanto,

$$\begin{aligned} R(k_1, \dots, k_r) &= \frac{\left(\frac{1}{r}\right)^{k_1} \cdots \left(\frac{1}{r}\right)^{k_r}}{\left(\frac{k_1}{k}\right)^{k_1} \cdots \left(\frac{k_r}{k}\right)^{k_r}} \\ &= \left(\frac{k}{k_1 r}\right)^{k_1} \cdots \left(\frac{k}{k_r r}\right)^{k_r}, \end{aligned}$$

mientras que la función de log-razón de verosimilitud será

$$\log R(k_1, \dots, k_r) = \sum_{i=1}^r k_i \log \left(\frac{k}{k_i r} \right).$$

Obtener la distribución de $R(K_1, \dots, K_r)$ no es un problema trivial, sin embargo podemos hacer uso de la distribución asintótica del estadístico $-2 \log R$, la cual se presenta a continuación.

Teorema 2.11. *El estadístico $-2 \log R(K_1, \dots, K_r)$ tiene distribución ji-cuadrada con $r - 1$ grados de libertad siempre que $k \rightarrow \infty$.*

No presentaremos la demostración de este teorema, para mayores referencias se puede consultar [7].

De esta forma, para nuestra muestra (k_1, \dots, k_r) , tomaremos un p -valor igual a

$$p = P \left(\chi_{(r-1)}^2 \geq -2 \sum_{i=1}^r k_i \log \left(\frac{k}{k_i r} \right) \right), \quad (2.7)$$

donde $\chi_{(r-1)}^2$ representa una variable aleatoria ji-cuadrada con $r - 1$ grados de libertad. Una prueba aproximadamente al $100(1 - \alpha)\%$ de confianza estará dada por

$$\phi(k_1, \dots, k_r) = \begin{cases} 1, & \text{si } p \geq \alpha, \\ 0, & \text{en otro caso,} \end{cases} \quad (2.8)$$

es decir, rechazamos la hipótesis nula en caso de que $p = p(k_1, \dots, k_r) < \alpha$.

Más aún, denotamos por $e_i = E(K_i)$. Bajo H_0 , tenemos que $e_i = \left(\sum_{j=1}^r K_j \right) / r = e$ para todo i .

Teorema 2.12. *El estadístico*

$$\sum_{j=1}^r \frac{(K_j - e)^2}{e} \quad (2.9)$$

también tiene distribución ji-cuadrada con $r - 1$ grados de libertad cuando $k \rightarrow \infty$.

Demostración. Esta demostración se basa en el hecho de que la expresión (2.9) puede ser generada a partir del estadístico $-2 \log R(K_1, \dots, K_r)$. Denotemos por $J_i = K_i - e_i$,

entonces

$$\begin{aligned}
 -2 \log R(K_1, \dots, K_r) &= -2 \sum_{i=1}^r K_i \log \left(\frac{e}{K_i} \right) \\
 &= 2 \sum_{i=1}^r (J_i + e_i) \log \left(\frac{J_i + e}{e} \right) \\
 &= 2 \sum_{i=1}^r (J_i + e_i) \log \left(1 + \frac{J_i}{e} \right).
 \end{aligned}$$

Utilizando la expansión de Taylor del logaritmo, tenemos que

$$\begin{aligned}
 -2 \log R(K_1, \dots, K_r) &= 2 \sum_{i=1}^r (J_i + e_i) \sum_{j=1}^{\infty} \frac{(-1)^{j+1}}{j} \left(\frac{J_i}{e} \right)^j \\
 &= 2 \sum_{i=1}^r \left[\sum_{j=1}^{\infty} \frac{(-J_i)^{j+1}}{j e^j} + J_i - \sum_{j=1}^{\infty} \frac{(-J_i)^{j+1}}{(j+1) e^j} \right] \\
 &= 2 \sum_{i=1}^r \left[J_i + \sum_{j=1}^{\infty} \frac{(-J_i)^{j+1}}{j(j+1) e^j} \right],
 \end{aligned}$$

donde $\sum_{i=1}^r J_i = \sum_{i=1}^r K_i - r e = 0$. Así

$$-2 \log R(K_1, \dots, K_r) = \sum_{i=1}^r \sum_{j=1}^{\infty} \left[\frac{2}{j(j+1)} \left(-\frac{J_i}{e} \right)^{j-1} \right] \frac{J_i^2}{e}. \quad (2.10)$$

Ahora calculamos la distribución de J_i/e . Recordemos que (K_1, \dots, K_r) , condicionado a que $\sum K_i = k$, tienen distribución multinomial de parámetros k y (p_1, \dots, p_r) , esto implica que K_i se distribuya binomial de parámetros con media $\mu_i = k p_i = e$ y varianza $\sigma^2 = k p_i (1 - p_i)$. Por lo tanto

$$\frac{J_i}{e} = \frac{K_i - e}{e} = \frac{K_i - \mu_i}{\sigma_i^2 / (1 - p_i)} = \frac{K_i - \mu_i}{\sigma_i} \cdot \sqrt{\frac{(1 - p_i)}{k p_i}}$$

Por el teorema de límite central, cuando k es grande, el término $(K_i - \mu_i)/\sigma_i$ converge en distribución a una variable aleatoria normal estándar. No obstante, el otro factor converge a cero con orden $1/\sqrt{k}$. Por lo tanto, el producto de ambas se acercará a cero con orden de $1/\sqrt{k}$. Esto permite que de la ecuación (2.10) obtengamos la aproximación

$$-2 \log R(K_1, \dots, K_r) \approx \sum_{i=1}^r \frac{J_i^2}{e} = \sum_{i=1}^r \frac{(K_i - e)^2}{e} \sim \chi_{(r-1)}^2,$$

concluyendo así la demostración. \square

Por lo tanto, el p -valor (2.7) considerado para la prueba de hipótesis puede ser cambiado por

$$p = P \left(\chi_{(r-1)}^2 \geq \sum_{j=1}^r \frac{(k_j - e)^2}{e} \right),$$

y la prueba en función de este estadístico será la misma expresión que en (2.8).

2.3.4. Aproximación al movimiento browniano

Supongamos que en efecto, $\{t_i^*; i \in \{1, \dots, k\}\}$ es un proceso de Poisson de tasa 1. Entonces, $\{t_i^*/\tau^*; i \in \{1, \dots, k\}\}$ forma también la realización de un proceso de Poisson pero de tasa τ^* en el intervalo $[0, 1]$.

Teorema 2.13. *Sea N_t un proceso de Poisson de tasa μ , y denotamos*

$$M_t = \frac{N_t - \mu t}{\sqrt{\mu}}, \quad t \in [0, 1].$$

Cuando hacemos tender a $\mu \rightarrow \infty$, $\{M_t; 0 \leq t \leq 1\}$ converge en distribución a un movimiento Browniano estándar.

Por lo tanto, podemos aplicar una prueba de bondad de ajuste de la siguiente forma:

1. Graficamos la función escalonada $Y(x)$ definida a través de los puntos $(t_i^*/\tau^*, i/k)$, para $i \in \{1, \dots, k\}$, en el cuadrado unitario.
2. Para cierto valor α , asignamos las bandas de confianza

$$y = x \pm \frac{Z_{1-\alpha/2}}{\sqrt{\tau^*}},$$

donde Z_p es una constante tal que $\Phi(Z_p) = p$, con Φ la función de distribución de una variable aleatoria normal estándar.

3. Una prueba de hipótesis aproximadamente al $100(1 - \alpha)$ % de nivel de confianza para la hipótesis nula

$$H_0 : \{t_1^*, \dots, t_k^*\} \text{ proviene de un proceso de Poisson de tasa uno,}$$

se obtiene rechazando H_0 cuando el proceso $Y(x)$ se sale de las bandas.

Un método alternativo de rechazo se basa en tomar la distribución del tiempo en el que un browniano estándar alcanza su máximo valor dentro del intervalo $[0, 1]$. Para plantear este método, empezamos enunciando el siguiente teorema.

Teorema 2.14. *Sea B_t un movimiento browniano estándar, definimos*

$$T^* = \arg \sup_{t \in [0, 1]} \{B_t\}.$$

T^ seguirá una distribución Beta de parámetros $1/2$ y $1/2$.*

De esta forma, tomando $t^* = t_{n^*}/\tau^*$, donde n^* está dado por

$$n^* = \underset{n \in \{1, \dots, k\}}{\operatorname{arg\,sup}} n - t_n^* k,$$

entonces se sigue que una prueba de hipótesis al $100(1 - \alpha)\%$ de nivel de confianza estará dada por rechazar H_0 cuando t^* esta por debajo del cuantil $(\alpha/2)$ o por encima del cuantil $(1 - \alpha/2)$ de una distribución $\text{Beta}(1/2, 1/2)$.

2.4. Simulación de procesos puntuales

2.4.1. Simulación del proceso de Poisson homogéneo

El problema de simular la realización de un proceso de Poisson homogéneo en el intervalo $[0, \tau]$ es trivial, basta con generar una sucesión de variables aleatorias exponenciales con tasa λ , $\{E_n; n \geq 1\}$, hasta alcanzar una cantidad $k + 1$ de variables simuladas tal que que $\sum_{n=1}^{k+1} E_n \geq \tau$. Entonces la sucesión $\{T_1, \dots, T_k\}$ con

$$T_n = \sum_{i=1}^n E_i, \quad n = 1, \dots, k,$$

serán dicho proceso de Poisson con intensidad λ en tal intervalo.

2.4.2. Método de la inversa del compensador

Además de emplearse para llevar a cabo pruebas de bondad de ajuste, otra aplicación del método de transformación de tiempos aleatorios es generar simulaciones de un cierto proceso puntual regular a través de un proceso de Poisson de tasa uno. Supongamos que queremos generar un proceso puntual T con función de intensidad condicional $\lambda(t)$. El siguiente algoritmo describe el método de simulación:

Algoritmo 2.1 Simulación utilizando la función inversa del compensador.

- 1: Generar un conjunto, $\{e_1, e_2, \dots\}$, de variables aleatorias exponenciales de parámetro uno.
 - 2: Utilizando la transformación de tiempo aleatorios, obtener secuencialmente los tiempos $T = \{t_1, t_2, \dots\}$, con $t_n = \Lambda^{-1}(\sum_{i=1}^n e_i)$.
 - 3: **Regresar** T .
-

El problema con este método es que usualmente se necesita hacer uso de métodos numéricos para evaluar la función inversa del compensador, lo cual puede hacer a este método poco eficaz.

2.4.3. Método de adelgazamiento

Es una alternativa al método anterior, la cual no necesita calcular la función inversa del compensador. Este método se fundamenta en la siguiente propiedad.

Lema 2.15. Sea $T^* = \{(T_n^*, Z_n^*); n \in \mathbb{N}\}$ un proceso puntual regular marcado con espacio de marcas \mathbb{R}_0 , función de intensidad uno y adaptado a la filtración $\mathcal{F} = \{\mathcal{F}_t; t \geq 0\}$, y dada la función

$$\mathcal{N}^*(C) = \sum_n \mathbb{1}_{\{(T_n^*, Z_n^*) \in C\}}, \quad C \in \mathcal{B}_0 \otimes \mathcal{B}_0,$$

con $\mathcal{B}_0 = \mathcal{B}(\mathbb{R}_0)$ el conjunto de Borel de \mathbb{R}_0 . Además, sea $\lambda(t)$ una función no negativa, continua por la izquierda y \mathcal{F} -adaptada. Entonces, el proceso $N = \{N(t); t \geq 0\}$, con t

$$N(t) = \begin{cases} \mathcal{N}((0, t]), & t > 0, \\ 0, & t = 0 \end{cases}$$

y donde $\mathcal{N}(dt) = \mathcal{N}^*(dt \times (0, \lambda(t)))$, será un proceso de conteo con función de intensidad $\lambda(t)$ y \mathcal{F} -adaptado.

Demostración. Esta demostración está basada en el hecho de que

$$\begin{aligned} \mathbb{E}(\mathcal{N}(dt) | \mathcal{F}_{t-}) &= \lim_{h \downarrow 0} \frac{1}{h} \mathbb{E}(\mathcal{N}^*(dt \times (0, \lambda(t))) | \mathcal{F}_{t-}) \\ &= \lim_{h \downarrow 0} \frac{1}{h} \mathbb{E}(\mathcal{N}^*([t, t+h) \times (0, \lambda(t))) | \mathcal{F}_{t-}) \\ &= \lambda(t) dt. \end{aligned}$$

□

Los algoritmos de adelgazamiento siguen la idea del lema anterior, sin embargo añaden la condición adicional de que la intensidad λ sea acotada. En este caso, podemos elegir a T^* como un proceso de Poisson compuesto de tasa base c , con $\lambda(t) \leq c$ para todo $t > 0$. Así, el algoritmo de simulación quedará descrito de la siguiente manera:

Algoritmo 2.2 Simulación mediante adelgazamiento.

- 1: Generar una realización $\{t_1^*, \dots, t_k^*\}$ de un proceso de Poisson de tasa c . Parar en cuanto $t_{k+1}^* > \tau$.
 - 2: Generar variables aleatorias independientes $u_1, \dots, u_k \sim \text{unif}(0, c)$.
 - 3: Tomar $i = 1, T = \emptyset$.
 - 4: Evaluar $\lambda(t_i^*)$.
 - 5: **Si** $u_i \leq \lambda(t_i^*)$ **entonces:**
 - 6: $T = T \cup \{t_i^*\}$.
 - 7: Avanzar $i = i + 1$ y repetir el paso 4.
 - 8: **Regresar** T .
-

2.4.4. Simulación de un proceso puntual con función de intensidad acotada a trozos

El siguiente algoritmo fue descrito por Shedler y Lewis en 1979, [11], aunque originalmente fue diseñado para generar puntos que seguían un proceso de Poisson no

Algoritmo 2.3 Algoritmo de adelgazamiento de Shedler-Lewis.

Entrada: $\lambda(t)$, función de intensidad no creciente durante los tiempos entre llegadas.

- 1: Tomar $\lambda^* = \mu$, $T = \emptyset$.
- 2: Generar $e_0 \sim \exp(\lambda^*)$.
- 3: **Si** $e_0 > \tau$ **entonces:**
- 4: Salir.
- 5: Tomar $t_1^* = e_0$, $T = \{t_1^*\}$, $\lambda^* = \lambda(t_1^*)$, $i = 1$.
- 6: Generar $e_i \sim \exp(\lambda^*)$.
- 7: **Si** $t_i^* + e_i > \tau$ **entonces:**
- 8: Salir.
- 9: Tomar $i = i + 1$, $t_i^* = t_{i-1}^* + e_{i-1}$.
- 10: Generar $u_i \sim \text{unif}(0, \lambda^*)$.
- 11: $\lambda^* = \lambda(t_i^* | T)$.
- 12: **Si** $u_i \leq \lambda^*$ **entonces:**
- 13: $T = T \cup \{t_i^*\}$.
- 14: Generar $e_i \sim \exp(\lambda^*)$ y regresar al paso 7.
- 15: **Regresar** T .

homogéneo. Este método es similar al algoritmo visto en la Subsección 2.4.2, pero condicionando y empleando el método de aceptación y rechazo para generar las variables aleatorias $T_n - T_{n-1}$.

Consideremos un proceso de conteo regular $\{N_t; 0 \leq t \leq \tau\}$ en el espacio de probabilidad (Ω, \mathcal{A}, P) , y $\mathcal{F} = \{\mathcal{F}_t\}$ la historia del proceso, con $\mathcal{F}_t = \sigma(\{N_s; s \leq t\})$. Sea $\lambda(t)$ su función de intensidad asociada, supongamos que podemos crear una función constante a trozos $\lambda^* : [0, \tau] \rightarrow \mathbb{R}$ tal que

$$\lambda^*(t) \geq \lambda(t), \quad c.s. \text{ para toda } t \in [0, \tau].$$

Tomemos $\lambda^*(t)$ como la función de intensidad de un proceso de Poisson homogéneo a trozos $\{N_t^*; 0 \leq t \leq \tau\}$.

Proposición 2.16. Sean $t_1^* < \dots < t_k^*$ los puntos observados del proceso $\{N_t^*\}$ en $[0, \tau]$. Si asociamos una variable aleatoria Y_i con distribución Bernoulli de parámetro $\lambda(t_n^*) / \lambda^*(t_n^*)$ a cada punto t_n^* , con $n = 1, \dots, k$. Entonces el conjunto de puntos

$$\{t_n^* : y_n = 1, n = 1, \dots, k\}$$

formarán la realización de un proceso puntual con función de intensidad $\lambda(t)$.

Supongamos que podemos escribir a la función de intensidad en términos de t y la secuencia de puntos $\{t_i : t_i \leq t\}$, es decir,

$$\lambda(t) = \lambda(t | t_1, \dots, t_n).$$

Además, supongamos que $\lambda(t) = \mu$ para $0 \leq t < t_1$.

Suponiendo que esta función sea decreciente siempre que no ocurren nuevos eventos, y si además, el salto de la función no es mayor a α en los instantes en los que cada evento

ocurre. Entonces el proceso $\{\lambda^*(t)\}$ tal que $\lambda(t|t_1, \dots, t_n) \leq \lambda^*(t)$ será construido siguiendo el Algoritmo 2.3.

Cabe destacar que este algoritmo se vuelve poco eficiente cuando la función de intensidad decrece rápido dado que esto implicará un mayor número de valores rechazados.

Ogata en 1981 [13] propuso una variación de este algoritmo para no necesitar que la función de intensidad sea acotada en todo el intervalo $[0, \tau]$. En su lugar, solo se requiere que esta función sea acotada localmente.

Sea $N = \{N_t; 0 \leq t \leq \tau\}$ un proceso de conteo regular con \mathcal{F} -función de intensidad asociada $\lambda(t)$. Supongamos que existen funciones $\lambda^*(t)$ y $\tau^*(t)$ tales que, para toda secuencia $0 < t_1 < \dots < t_n < t$, satisfacen

$$\lambda^*(t) \geq h_{n+1}^*(t + u | \mathcal{F}_{(n)}), \quad \text{para toda } u \in [0, \tau^*(t)],$$

donde h^* es la función de riesgo definida en el capítulo anterior por (1.14).

En este caso, el método para simular una realización de N se presenta a continuación:

Algoritmo 2.4 Algoritmo de adelgazamiento modificado de Ogata.

Entrada: $\lambda(t)$, función de intensidad localmente acotada.

- 1: Tomar $t = 0$, $i = 1$, $T = \emptyset$.
 - 2: Calcular $\lambda^*(t)$ y $\tau^*(t)$.
 - 3: Generar $e_i \sim \exp(\lambda^*(t))$.
 - 4: **Si** $t + e_i > \tau$ **entonces:**
 - 5: Salir.
 - 6: **Si** $e_i > \tau^*(t)$ **entonces:**
 - 7: Hacer $t = t + \tau^*(t)$.
 - 8: Regresar al paso 2.
 - 9: Generar $u_i \sim \text{unif}(0, \lambda^*(t))$.
 - 10: **Si** $u_i \leq \lambda(t + e_i)$ **entonces:**
 - 11: $t_i^* = t_i^* + e_i$, $T = T \cup \{t_i^*\}$, $i = i + 1$, $t = t + \tau^*(t)$.
 - 12: Regresar al paso 2.
 - 13: **Regresar** T .
-

2.4.5. Simulación de un proceso puntual multivariado

La generalización del algoritmo anterior se da mediante la siguiente proposición.

Proposición 2.17. *Consideremos un proceso puntual regular multivariado con espacio de marcas $Q = \{1, \dots, d\}$ definido en el intervalo $[0, \tau]$, $T = \{(T_n, Z_n); n \in \mathbb{N}\}$, con historia dada por la filtración \mathcal{F} y función de intensidad \mathcal{F} -predecible asociada $\lambda = \{\lambda^{(i)}(t); i \in Q, 0 \leq t \leq \tau\}$. Supongamos que podemos encontrar un proceso \mathcal{F} -predecible unidimensional $\lambda^*(t)$ tal que*

$$\sum_{i=1}^d \lambda^{(i)}(t) \leq \lambda^*(t), \quad t \in [0, \tau] \quad (\text{c.s.}),$$

y definamos

$$\lambda^{(0)}(t) = \lambda^*(t) - \sum_{i=1}^d \lambda^{(i)}(t).$$

Sea $\{t_n^*, n \in \mathbb{N}\}$ la realización de un proceso puntual con función de intensidad $\lambda^*(t)$ en el intervalo $[0, \tau]$. Si asociamos a cada tiempo t_n^* una variable aleatoria Y_n donde

$$P(Y_n = i) = \begin{cases} \frac{\lambda^{(i)}(t_n^*)}{\lambda^*(t_n^*)}, & i \in Q \cup \{0\}, \\ 0, & \text{en otro caso.} \end{cases}$$

Entonces la secuencia

$$\{(t_n^*, y_n) : n \in \mathbb{N}, y_n \in Q\}$$

corresponde a una realización de T .

2.5. Predicción

Con frecuencia, el tema de predicción de un proceso puntual se encuentra fuertemente vinculado al problema de simulación. Sea $T = \{T_n\}$ un proceso puntual, supongamos que deseamos predecir cierta cantidad V que depende del futuro del proceso, por ejemplo, el tiempo en que ocurre determinado número de eventos, la probabilidad de que ocurra un evento en un intervalo de tiempo futuro, etcétera. Obtener una expresión analítica para V no es un problema sencillo en la mayoría de las ocasiones, sin embargo, este valor puede ser estimado por medio del método de Bootstrap realizando múltiples simulaciones, una vez que se conoce la estructura del proceso.

Consideremos que el proceso T es observado en el intervalo $[a, b]$ y supongamos que conocemos su estructura. Empezamos realizando simulaciones del proceso, tomando como origen al tiempo b , por lo que la información que conocemos hasta ese instante la describiremos por medio de la filtración \mathcal{F}_0 . Fijamos un horizonte $\tau > 0$ hasta donde queremos realizar nuestras simulaciones, y obtenemos la muestra v_1, \dots, v_b de nuestra variable de interés V . La función

$$\hat{f}(v) = \sum_{i=1}^b \mathbb{1}_{\{v_i \leq v\}}$$

será la función de distribución empírica de V .

Así, valores importantes de V tales como su media, varianza o cuantiles, podrán ser estimados en función de $\hat{f}(v)$. El tamaño de muestra b tendrá que ser lo suficientemente grande considerando la precisión que queremos obtener.

Por otro lado, si no conocemos la estructura del proceso pero si sus observaciones, entonces podemos estimar su función de intensidad y a partir de ella simular los valores futuros v'_1, \dots, v'_b para calcular nuestro estadístico de interés. Sin embargo, esta estimación conllevará un error de ajuste que habría que considerar al momento de querer hacer intervalos de predicción a un cierto nivel.

2.6. Ejemplo

Consideremos un proceso de Poisson no homogéneo, $N = \{N(t); t \geq 0\}$, cuya función de intensidad está dada por la siguiente ecuación:

$$\lambda(t) = \frac{3}{4} \sin\left(\frac{\pi t}{20}\right) + \frac{5}{6}, \quad t \geq 0. \quad (2.11)$$

La forma senoidal de λ causará un agrupamiento periódico en los puntos generados, los cuales tenderán a estar más juntos conforme la función de intensidad aumenta, y se separan cuando esta decrece. Un modelo con estas características es útil para simular los tiempos de llegada de sucesos periódicos tales como: lluvias, las cuales se concentran en meses específicos del año, o bien, el instante en que vehículos atraviesan cierta vialidad, congestionada en determinadas horas del día, entre otros casos.

2.6.1. Simulación

Se observa como el contradominio de λ es el conjunto acotado $[\frac{1}{12}, \frac{19}{12}]$. Aplicando el Algoritmo 2.2, podemos construir una realización de nuestro proceso de conteo hasta un tiempo final $\tau = 120$, rechazando algunos de los puntos generados por un proceso de Poisson de tasa $c = \frac{19}{12}$. La Figura 2.1 muestra en la parte inferior la forma de la función de intensidad, y del lado superior una realización de N obtenida con este método.

Para este ejemplo fueron requeridos 186 puntos generados por el proceso de Poisson, y se obtuvieron 96 del proceso de conteo N , es decir, se rechazaron poco menos de la mitad de los puntos. Aunque en este caso fue fácil elegir una cota adecuada utilizando el contradominio de λ , en general es un tema de gran importancia ya que de esto dependerá la eficiencia con que trabaje nuestro algoritmo.

Corriendo 10 000 simulaciones de N , la siguiente tabla muestra el porcentaje de puntos rechazados del proceso de Poisson de tasa c , para diferentes valores de c .

c	Porcentaje de rechazos
19/12	47.1147 %
2	58.1117 %
4	78.9895 %
8	89.4829 %
16	94.7341 %

Observamos como los rechazos aumentan notablemente. En el último caso, generar 100 valores del proceso de Poisson solo nos proporcionará alrededor de 6 puntos del proceso de conteo de nuestro interés.

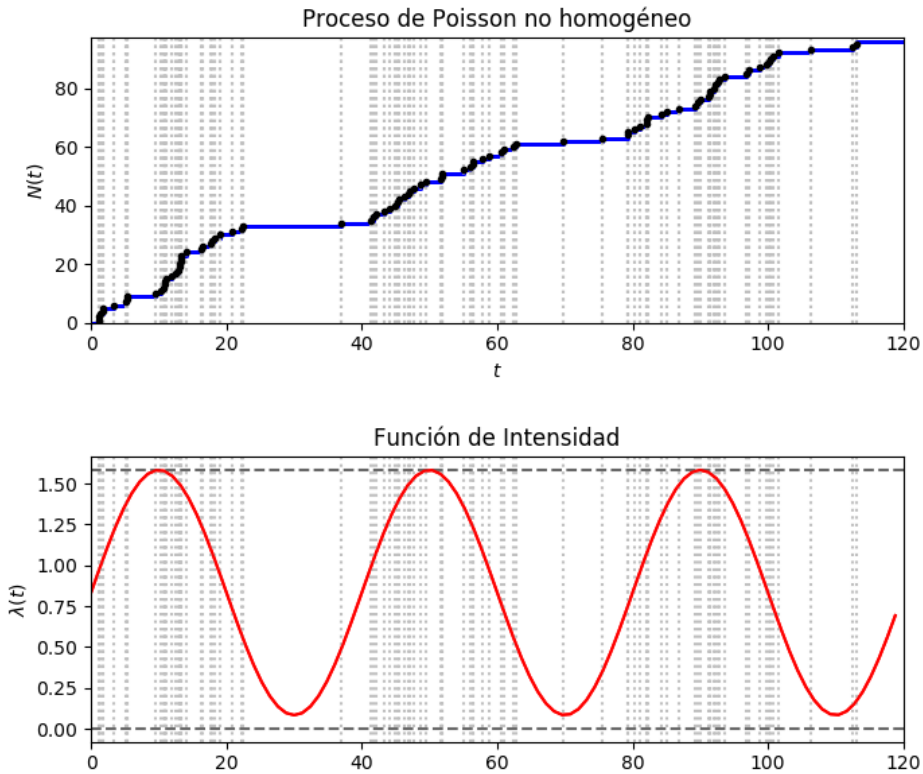


Figura 2.1: Proceso de Poisson no homogéneo con función de intensidad periódica

2.6.2. Bondad de ajuste

Por otro lado, el compensador de la función de intensidad dada por (2.11), es

$$\begin{aligned} \Lambda(t) &= \int_0^t \lambda(s) ds = \int_0^t \left[\frac{3}{4} \sin\left(\frac{\pi s}{20}\right) + \frac{5}{6} \right] ds \\ &= \frac{5}{6}t + \frac{3}{4} \int_0^t \sin\left(\frac{\pi s}{20}\right) ds = \frac{5}{6}t + \frac{15}{\pi} \left[1 - \cos\left(\frac{\pi t}{20}\right) \right], \end{aligned} \tag{2.12}$$

para $t \geq 0$.

De la Sección 2.3, sabemos que la sucesión $\{u_1, \dots, u_k\}$, con $u_n = \Lambda(t_n) / \Lambda(t_k)$ para $n = 1, \dots, k - 1$, corresponde a las estadísticas de orden de una muestra aleatoria de cierta variable uniforme $(0, 1)$. Así, en la Figura 2.2 se observa esta transformación aplicada a los tiempos $\{t_1, \dots, t_k\}$ obtenidos en la realización mostrada en la Figura 2.1.

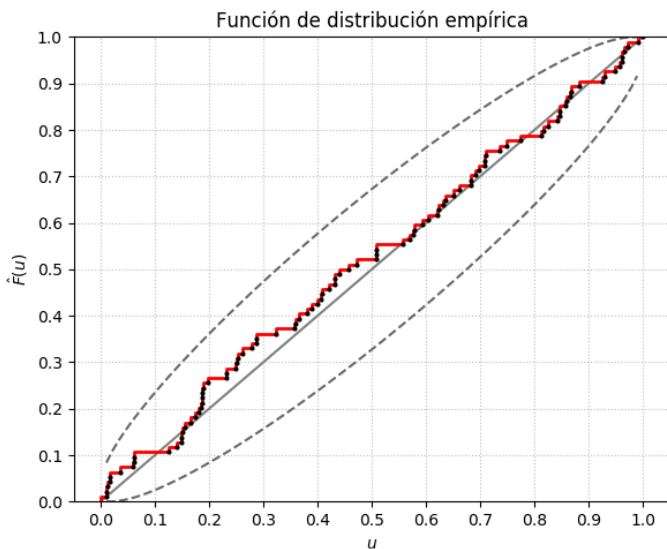


Figura 2.2: Función de distribución empírica de $\{u_1, \dots, u_{k-1}\}$.

Como era de esperarse, esta función se ajusta bien a la recta de pendiente 1.

Las bandas de confianza que se observan, se obtienen al considerar a cada estadística de orden de la uniforme como una variable aleatoria con distribución beta,

$$U_n \sim \text{Beta}(n, k - n), \quad n = 1, \dots, k - 1.$$

De esta forma, se puede crear un intervalo de predicción de probabilidad γ para cada estadística de orden uniforme. Al considerar a los k intervalos simultáneos, se puede emplear la desigualdad de Bonferroni para garantizar que la probabilidad global de todos ellos sea al menos de $(1 - \alpha)$. La relación que guardan las probabilidades para tal fin es que $\gamma = 1 - \alpha/k$. De manera que si $\alpha = 0.05$ y $k = 30$, por ejemplo, entonces $\gamma \approx 0.9983$. Esto quiere decir que si se obtienen 30 intervalos de predicción, cada uno de ellos con una probabilidad individual de 0.9983, entonces se tendrá que la probabilidad global conjunta simultánea de todos ellos de incluir a las k variables aleatorias consideradas será de al menos 0.95. Para consultar más detalles de esta prueba, se pueden consultar [16] y [4].

Por último, observamos como el compensador, dado por (2.12), no resultó ser una función inversible en este ejemplo. Por lo tanto, aplicar el método de simulación de la inversa del compensador no sería recomendable ya que, como se mencionó anteriormente, esto conllevaría a resolver ecuaciones no lineales numéricamente, lo cual significa un gran gasto de recursos computacionales.

3

El Proceso de Hawkes

En el ejemplo presentado al final del capítulo anterior, se planteó un modelo que trataba de replicar la llegada de eventos, los cuales con cierta regularidad se agrupaban o separaban dependiendo del instante en el que se encontraban. Pero ¿cómo podríamos modelar fenómenos en los cuales estos agrupamientos no se vieran dominados por el tiempo sino que, en cambio, estuvieran influidos por los mismo individuos?

Para estos casos, en la década de los 70's, Alan Geoffrey Hawkes [6] propuso un tipo particular de proceso puntual conocido hoy como modelo de Hawkes. La característica principal de este modelo es la denominada propiedad de *autoexcitación* del proceso, la cual consiste en el aumento instantáneo de su intensidad con cada llegada de un nuevo evento, misma que decae gradualmente en ausencia de más sucesos conforme pasa el tiempo.

Este proceso fue aplicado originalmente para replicar los tiempos en que ocurrían sismos. Sin embargo, sus características le proporcionaron una gran versatilidad, por lo que en los últimos años se ha utilizado para simular el comportamiento de mercados financieros, brotes de algún tipo de enfermedad contagiosa, modelar el incremento de actividades delictivas en cierta ciudad, etc.

3.1. Caso unidimensional

Definición 3.1 (Proceso de Hawkes unidimensional). Sea $N = \{N_t; t \geq 0\}$ un proceso de conteo regular en $(\Omega, \mathcal{A}, \mathbb{P})$. Decimos que N es un *proceso de Hawkes* unidimensional si su función de intensidad asociada puede ser descrita por la ecuación:

$$\lambda_t = \mu + \int_0^t \phi(t-s) dN_s, \quad (3.1)$$

para alguna constante $\mu > 0$, y con $\phi : [0, \infty) \rightarrow [0, \infty)$ una función tal que satisface la condición de regularidad

$$\int_0^\infty \phi(s) ds < 1. \quad (3.2)$$

La razón por la cual necesitamos de esta condición es para evitar que nuestro proceso explote con probabilidad 1. Su demostración no se efectuará de momento, pero este tema se abordará nuevamente en la sección 3.4 cuando discutamos la relación que existe entre un proceso de Hawkes y uno de ramificación. Por el momento, en las secciones siguientes solo supondremos que se satisface.

El valor de μ se puede considerar como la *intensidad exógena* al proceso. Por otro lado, ϕ es llamada la *función de excitación* o *de respuesta*, la cual usualmente depende de un vector de parámetros $\theta = (\theta_1, \dots, \theta_r)$, y por lo tanto, podemos denotar $\phi(s) = \phi(s; \theta)$. Consecuentemente

$$\lambda_t = \lambda(t; \mu, \theta) = \mu + \int_0^t \phi(t-s; \theta) dN_s.$$

Sean t_1, t_2, t_3, \dots los tiempos en que ocurren los eventos para cierta realización del proceso de Hawkes, podemos escribir (3.1) como función de t y el conjunto de tiempos $\{t_i : t_i \leq t\}$. De esta forma obtenemos

$$\lambda_t = \mu + \sum_{t_i < t} \phi(t - t_i; \theta). \quad (3.3)$$

De esta última ecuación, se deduce que el proceso de Hawkes no pertenece a la clase de procesos markovianos.

Denotaremos por $\mathcal{F} = \{\mathcal{F}_t; t \geq 0\}$ a la filtración que representa la historia de N . Una propiedad destacable del Capítulo 1 sobre la función de intensidad es que

$$P(N_{(t+h)-} - N_{t-} = n | \mathcal{F}_{t-}) = \begin{cases} 1 - \lambda_t h + o(h), & n = 0, \\ \lambda_t h + o(h), & n = 1, \\ o(h), & n > 1, \end{cases}$$

de lo cual se sigue que

$$\begin{aligned} P(N_t - N_s = 0 | \mathcal{F}_s) &= \exp \left[- \int_s^t \lambda_u du \right] = \exp \left[- \int_s^t \left[\mu + \sum_{t_i < u} \phi(u - t_i; \theta) \right] du \right] \\ &= \exp \left[-\mu(t-s) - \sum_{i=1}^{N_s} \int_s^t \phi(u - t_i; \theta) du \right]. \end{aligned}$$

De esta última expresión se genera la propiedad de excitación del proceso: mientras más eventos hayan ocurrido hasta un tiempo s , la probabilidad de que no vaya a ocurrir ninguno más en el intervalo $(s, t]$ decrece exponencialmente.

A continuación, aplicaremos algunas otras ideas expuestas de forma general para los procesos puntuales en los dos capítulos anteriores, para el caso particular del proceso de Hawkes unidimensional.

3.1.1. Verosimilitud del proceso de Hawkes

El tema abordado en esta sección fue trabajado originalmente por Ozaki en [15]. Supongamos que observamos el proceso hasta un tiempo $\tau > 0$, y los tiempos en que ocurren

los eventos son t_1, \dots, t_k , con $t_n \leq \tau$ para todo $n = 1, \dots, k$. Supongamos además que la función de intensidad del proceso depende de un vector de parámetros $\theta = (\theta_1, \dots, \theta_r)$. Por el Teorema 2.1, sabemos que la función de log-verosimilitud del modelo estará dada por

$$l(\theta; t_1, \dots, t_k) = - \int_0^\tau \lambda(t; \mu, \theta) dt + \int_0^\tau \log \lambda(t; \mu, \theta) dN_t. \quad (3.4)$$

Aunque depende de la forma de λ_t , usualmente la función de log-verosimilitud es no lineal respecto a sus parámetros. En tal caso, requeriremos hacer uso de técnicas especiales de optimización no lineal para encontrar el estimador de máxima verosimilitud (EMV) para θ . Una forma de resolver este problema será aplicando el método de Newton-Raphson de dimensión r para maximizar (3.4), para ello, calcularemos el gradiente y el Hessiano de nuestra función.

El gradiente, ∇l , estará formado por las componentes

$$\frac{\partial l}{\partial \theta_j} = - \int_0^\tau \frac{\partial}{\partial \theta_j} \lambda(t; \mu, \theta) dt + \int_0^\tau \frac{\frac{\partial}{\partial \theta_j} \lambda(t; \mu, \theta)}{\lambda(t; \mu, \theta)} dN_t, \quad (3.5)$$

por su parte, el Hessiano, $D^2 l$, se formará por

$$\begin{aligned} \frac{\partial^2 l}{\partial \theta_j \partial \theta_i} = & - \int_0^\tau \frac{\partial^2}{\partial \theta_j \partial \theta_i} \lambda(t; \mu, \theta) dt \\ & + \int_0^\tau \frac{\lambda(t; \mu, \theta) \frac{\partial^2 \lambda(t; \mu, \theta)}{\partial \theta_j \partial \theta_i} - \frac{\partial \lambda(t; \mu, \theta)}{\partial \theta_j} \frac{\partial \lambda(t; \mu, \theta)}{\partial \theta_i}}{[\lambda(t; \mu, \theta)]^2} dN_t, \end{aligned} \quad (3.6)$$

para $i, j = 1, 2, \dots, r$.

El método de maximización de una función por Newton-Raphson se basa en la expansión de Taylor de segundo orden de la función objetivo. El siguiente algoritmo describe su implementación en este caso:

Algoritmo 3.1 EMV obtenido con el método de maximización de Newton-Raphson.

- 1: Proponer un valor inicial θ_0 para la solución, y una tolerancia ϵ .
 - 2: Tomar $i = 0$
 - 3: **Mientras** $\|\nabla l(\theta_i)\| \geq \epsilon$ **hacer**:
 - 4: $i = i + 1$.
 - 5: $\theta_i = \theta_{i-1} - (D^2 l(\theta_{i-1}))^{-1} \nabla l(\theta_i)$
 - 6: Tomar $\hat{\theta} = \theta_i$.
 - 7: **Regresar** $\hat{\theta}$.
-

3.1.2. Simulación del proceso de Hawkes

Para generar una sucesión de tiempos aleatorios, $\{T_1, T_2, \dots\}$, que siga un proceso de Hawkes, podemos utilizar el método de la inversa del compensador descrito en el Algoritmo 2.1 del capítulo anterior. El compensador de nuestro proceso estará dado por:

$$\Lambda(t) = \int_0^t \left[\mu + \sum_{t_i < s} \phi(s - t_i; \theta) \right] ds,$$

la idea entonces será generar variables aleatorias independientes con distribución exponencial de parámetro 1, $\{E_1, E_2, \dots\}$, y resolver el sistema:

$$T_i = \Lambda^{-1} \left(\sum_{j=1}^i E_j \right).$$

Podemos demostrar que lo anterior se reduce a encontrar los valores t_1, t_2, \dots que satisfagan recursivamente el sistema de ecuaciones:

$$E_i = \int_{t_{i-1}}^{t_i} \left[\mu + \sum_{t_i < s} \phi(s - t_i; \theta) \right] ds, \quad (3.7)$$

para $i = 1, 2, \dots$, y tomando $t_0 \equiv 0$. De esta manera los tiempos de ocurrencia de nuestro proceso de Hawkes estarán dados por $T_i = t_i$, para $i \geq 1$.

Asimismo, dependiendo de la forma de la función de excitación, es posible emplear los Algoritmos 2.2, 2.3 y 2.4 de simulación por el método de adelgazamiento para generar datos de nuestro proceso, sin la necesidad de resolver el sistema de ecuaciones (3.7).

3.2. Selección de la función de excitación

Hasta ahora solo se ha pedido que la función de excitación setisfaga $\int_0^\infty \phi(s) ds < 1$. Pero ¿qué forma debe de tener dicha función? En la presente sección se discutirá sobre el uso de diferentes tipos de funciones de excitación y, en particular, el uso de una función con decrecimiento exponencial. Asimismo, se expondrán ejemplos con simulaciones para cada caso.

3.2.1. Función con decrecimiento exponencial

Desde la introducción de su modelo en [6], Hawkes propone emplear una función de este tipo, tal función la representaremos de la siguiente manera:

$$\phi(s) = \alpha e^{-\beta s}, \quad s \geq 0,$$

para valores constantes $\alpha, \beta > 0$.

Entonces, su función de intensidad quedará

$$\lambda_t = \mu + \alpha \int_0^t e^{-\beta(t-s)} dN_s,$$

o equivalentemente, denotando con t_1, \dots, t_k a los instantes en que los eventos ocurren, entonces

$$\lambda_t = \mu + \alpha \sum_{t_n < t} e^{-\beta(t-t_n)}.$$

Supongamos que observamos una realización de nuestro proceso hasta el tiempo $\tau > 0$, y por simplicidad, tomaremos $\tau = t_k$. La función de log verosimilitud asociada estará dada por

$$l(\mu, \alpha, \beta; t_1, \dots, t_k) = \int_0^{t_k} \log \lambda(t; \mu, \theta) dN_t - \Lambda(t_k),$$

donde $\Lambda(\cdot)$ representa el compensador de nuestro proceso, definido en este caso como

$$\Lambda(t) = \mu t + \alpha \int_0^t \sum_{t_i < s} e^{-\beta(s-t_i)} ds.$$

Notemos que $\Lambda(t_n) = \mu t_n$ para $n = 1$. Por otro lado,

$$\begin{aligned} \Lambda(t_n) &= \mu t_n + \alpha \sum_{j=1}^{n-1} \int_{t_j}^{t_{j+1}} \sum_{t_i < s} e^{-\beta(s-t_i)} ds = \mu t_n + \alpha \sum_{j=1}^{n-1} \int_{t_j}^{t_{j+1}} \sum_{i=1}^j e^{-\beta(s-t_i)} ds \\ &= \mu t_n + \alpha \sum_{j=1}^{n-1} \sum_{i=1}^j \int_{t_j}^{t_{j+1}} e^{-\beta(s-t_i)} ds \\ &= \mu t_n - \frac{\alpha}{\beta} \sum_{j=1}^{n-1} \sum_{i=1}^j \left[e^{-\beta(t_{j+1}-t_i)} - e^{-\beta(t_j-t_i)} \right] \\ &= \mu t_n - \frac{\alpha}{\beta} \sum_{j=1}^{n-1} \left[e^{-\beta(t_n-t_j)} - 1 \right], \quad n = 2, \dots, k. \end{aligned} \tag{3.8}$$

De esta manera, la función de log-verosimilitud queda descrita de la siguiente manera

$$\begin{aligned} l(\mu, \alpha, \beta; t_1, \dots, t_k) &= \sum_{n=1}^k \log \left[\mu + \alpha \sum_{t_m < t_n} e^{-\beta(t_n-t_m)} \right] - \left[\mu t_k - \frac{\alpha}{\beta} \sum_{j=1}^{k-1} \left[e^{-\beta(t_k-t_j)} - 1 \right] \right] \\ &= \sum_{n=1}^k \log [\mu + \alpha S(n)] - \mu \tau + \frac{\alpha}{\beta} \sum_{n=1}^k \sum_{m=1}^n \left[e^{-\beta(t_{n+1}-t_m)} - e^{-\beta(t_n-t_m)} \right], \end{aligned} \tag{3.9}$$

donde $S(n) = \sum_{t_m < t_n} e^{-\beta(t_n-t_m)}$ para $n \geq 2$, y $S(1) = 0$.

De (3.5), podemos calcular las componentes de su gradiente:

$$\begin{aligned} \frac{\partial l}{\partial \alpha} &= \frac{1}{\beta} \sum_{n=1}^k \left[e^{-\beta(\tau-t_n)} - 1 \right] + \sum_{n=1}^k \frac{S(n)}{\mu + \alpha S(n)}, \\ \frac{\partial l}{\partial \beta} &= \frac{\alpha}{\beta} \sum_{n=1}^k \left[-(\tau - t_n) e^{-\beta(\tau-t_n)} \right] - \frac{\alpha}{\beta^2} \sum_{n=1}^k \left[e^{-\beta(\tau-t_n)} - 1 \right] + \sum_{n=1}^k \frac{S_\beta(n)}{\mu + \alpha S(n)}, \\ &= -\frac{\alpha}{\beta} \sum_{n=1}^k \left[\left(\tau - t_n + \frac{1}{\beta} \right) e^{-\beta(\tau-t_n)} - \frac{1}{\beta} \right] + \sum_{n=1}^k \frac{S_\beta(n)}{\mu + \alpha S(n)}, \end{aligned}$$

$$\frac{\partial l}{\partial \mu} = -\mu + \sum_{n=1}^k \frac{1}{\mu + \alpha S(n)},$$

con $S_\beta(n) = \frac{\partial S(n)}{\partial \beta} = -\sum_{t_m < t_n} (t_n - t_m) e^{-\beta(t_n - t_m)}$ para $n \geq 2$, y $S_\beta(1) = 0$.

Mientras que con (3.6), formamos el hessiano:

$$\frac{\partial^2 l}{\partial \alpha^2} = -\sum_{n=1}^k \left[\frac{S(n)}{\mu + \alpha S(n)} \right]^2, \quad \frac{\partial^2 l}{\partial \alpha \partial \mu} = -\sum_{n=1}^k \frac{S(n)}{(\mu + \alpha S(n))^2},$$

$$\begin{aligned} \frac{\partial^2 l}{\partial \alpha \partial \beta} = & -\sum_{n=1}^k \left[\frac{1}{\beta} (\tau - t_n) e^{-\beta(\tau - t_n)} + \frac{1}{\beta^2} (e^{-\beta(\tau - t_n)} - 1) \right] \\ & + \sum_{n=1}^k \left[\frac{\alpha S(n)}{(\mu + \alpha S(n))^2} - \frac{1}{\mu + \alpha S(n)} \right] S_\beta(n), \end{aligned}$$

$$\begin{aligned} \frac{\partial^2 l}{\partial \beta^2} = & \frac{\alpha}{\beta} \sum_{n=1}^k \left[(\tau - t_n)^2 e^{-\beta(\tau - t_n)} + \frac{2}{\beta} (\tau - t_n) e^{-\beta(\tau - t_n)} + \frac{2}{\beta^2} (e^{-\beta(\tau - t_n)} - 1) \right] \\ & + \sum_{n=1}^k \left[\frac{\alpha S_{\beta\beta}(n)}{\mu + \alpha S(n)} - \left(\frac{\alpha S_\beta(n)}{\mu + \alpha S(n)} \right)^2 \right], \end{aligned}$$

$$\frac{\partial^2 l}{\partial \beta \partial \mu} = \sum_{n=1}^k \frac{\alpha S_\beta(n)}{(\mu + \alpha S(n))^2}, \quad \frac{\partial^2 l}{\partial \mu^2} = -\sum_{n=1}^k \frac{1}{(\mu + \alpha S(n))^2},$$

con $S_{\beta\beta}(n) = \frac{\partial^2 S(n)}{\partial \beta^2} = \sum_{t_m < t_n} (t_n - t_m)^2 e^{-\beta(t_n - t_m)}$ para $n \geq 2$, y $S_{\beta\beta}(1) = 0$.

Con esto, la maximización de (3.9) puede efectuarse siguiendo el Algoritmo 3.1 o algún otro método de optimización.

Con respecto a las simulaciones, un modo de generar valores que sigan este proceso es emplear el Algoritmo 2.3 de adelgazamiento, valiéndose de que la función de intensidad en este caso será acotada a trozos: Sean t_1, \dots, t_k los tiempos en que los eventos ocurren en un proceso de Hawkes con función de excitación con decrecimiento exponencial, entonces

$$\lambda(t) \leq \lambda(t_k), \quad \text{para todo } t_k < t < t_{k+1}.$$

3.2.2. Ejemplos

Se simula una realización de un proceso de Hawkes, considerando una función de excitación con decrecimiento exponencial de parámetros $\mu = 0.2$, $\alpha = 0.5$ y $\beta = 0.7$, y dejando correr el proceso hasta un tiempo final $\tau = 200$. Durante este intervalo, el proceso, mostrado en la parte superior de la Figura 3.1, generó 192 datos en total. Su función de intensidad asociada es graficada en la parte inferior de la misma figura.

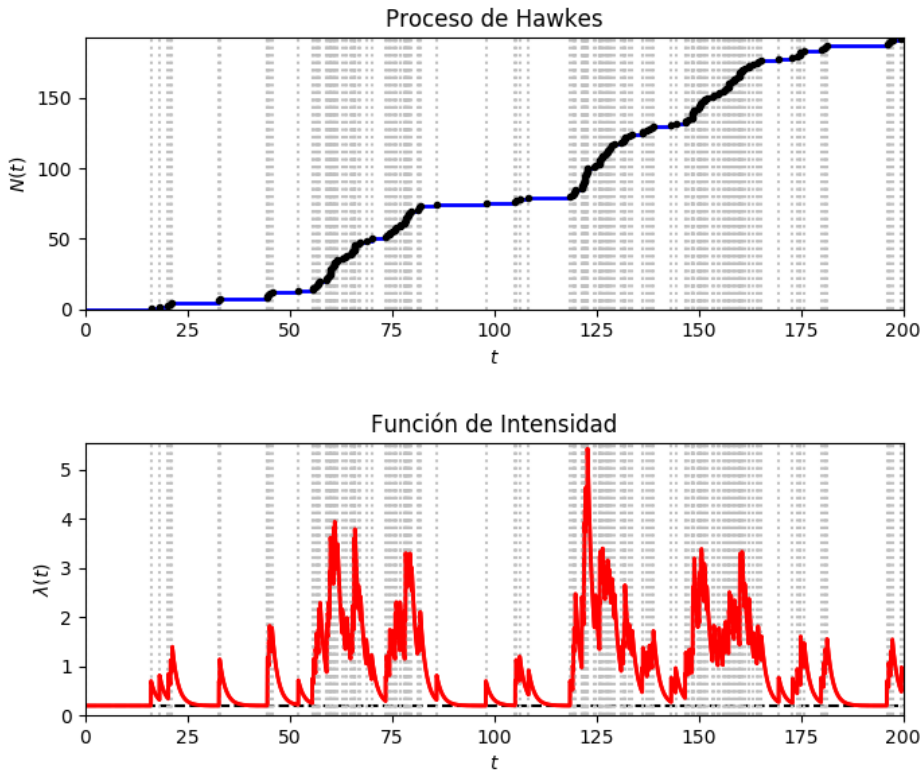


Figura 3.1: Simulación de un proceso de Hawkes unidimensional y su función de intensidad asociada.

De esta gráfica, se puede observar la denominada propiedad de autoexcitación del proceso cuando la función de intensidad da un salto cada que un nuevo evento llega. Esto ocasionó la aparición de agrupamientos alrededor de los instantes 70, 130 y 160.

Con estos datos, trataremos de estimar los mismos parámetros de la función de excitación maximizando la función de log-verosimilitud del proceso. La siguiente tabla muestra los resultados:

Parámetro	Valor real	Valor estimado
μ	0.2	0.179733
α	0.5	0.660526
β	0.7	0.806402

En la figura 3.2, se muestra la función de intensidad real y la estimada a lo largo de nuestra ventana de tiempo. Se observa una gran similitud, sin embargo, dado que el valor estimado de α fue mayor al real, la intensidad estimada tiende a estar por encima de la real cada vez que un suceso ocurre.

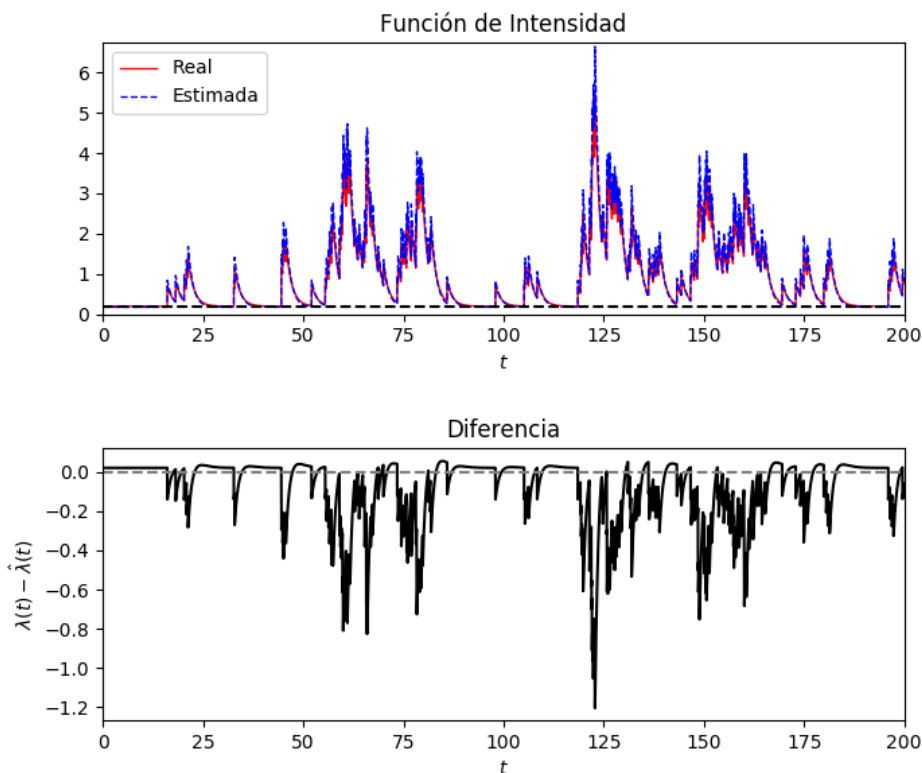


Figura 3.2: Comparativo entre la función de intensidad real y la estimada por máxima verosimilitud.

A continuación, con el propósito de evaluar la estimación obtenida, aplicaremos las pruebas de hipótesis de uniformidad presentadas en el Capítulo 2. Para ello, realizaremos la transformación de tiempos aleatorios utilizando el compensador de nuestro proceso $\Lambda(\cdot)$, definido en (3.8), y tomando

$$u_n = \Lambda(t_n) / \Lambda(t_k), \quad n = 1, \dots, k - 1.$$

La Figura 3.3 muestra la función de distribución empírica de los tiempos transformados, así como las bandas de confianza obtenidas de la distribución beta asociada a las estadísticas de orden, análogo a como se hizo con los datos del proceso de Poisson no homogéneo de la Sección 2.6. Notamos como todos los puntos caen dentro de dichas bandas, por lo que no podríamos rechazar la hipótesis de uniformidad de estos datos.

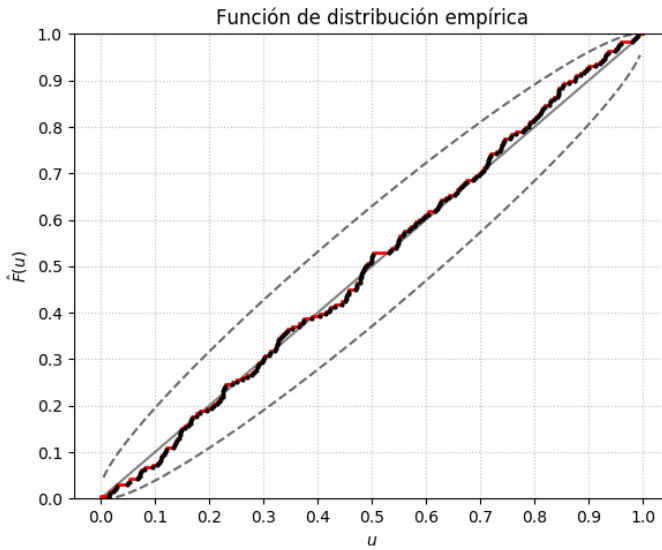


Figura 3.3: Función de distribución empírica de $\{u_1, \dots, u_{k-1}\}$.

Ahora, usaremos el proceso del ejemplo de la Sección 2.6 para generar datos hasta un tiempo $\tau = 200$ y ajustarles un proceso de Hawkes. Recordemos que ese proceso contaba con una función de intensidad senoidal, por lo cual, presentaba agrupamientos siguiendo cierta periodicidad (ver Figura 2.1). Proponiendo una función de excitación con decrecimiento exponencial, la siguiente gráfica muestra los parámetros obtenidos por medio de la estimación de máxima verosimilitud:

Parámetro	Valor estimado
μ	0.365782
α	0.352175
β	0.560325

En la Figura 3.4 podemos observar la función de intensidad estimada por el proceso de Poisson, y la real.

Observamos como, aunque no con gran exactitud, el modelo logra captar razonablemente bien la figura de la función de intensidad real. En la gráfica de la Figura 3.5, se obtiene la función de distribución empírica de los tiempos transformados $\{u_1, \dots, u_{k-1}\}$, se observa como la gráfica se ajusta muy bien a la recta de pendiente 1, y en ningún instante se sale de las bandas de confianza al 95%, por lo cual tampoco podríamos rechazar la hipótesis de normalidad.

Lo anterior nos muestra la versatilidad de uso del modelo de Hawkes, siendo capaz de usarse en diferentes tipos de datos que presenten agrupamientos.

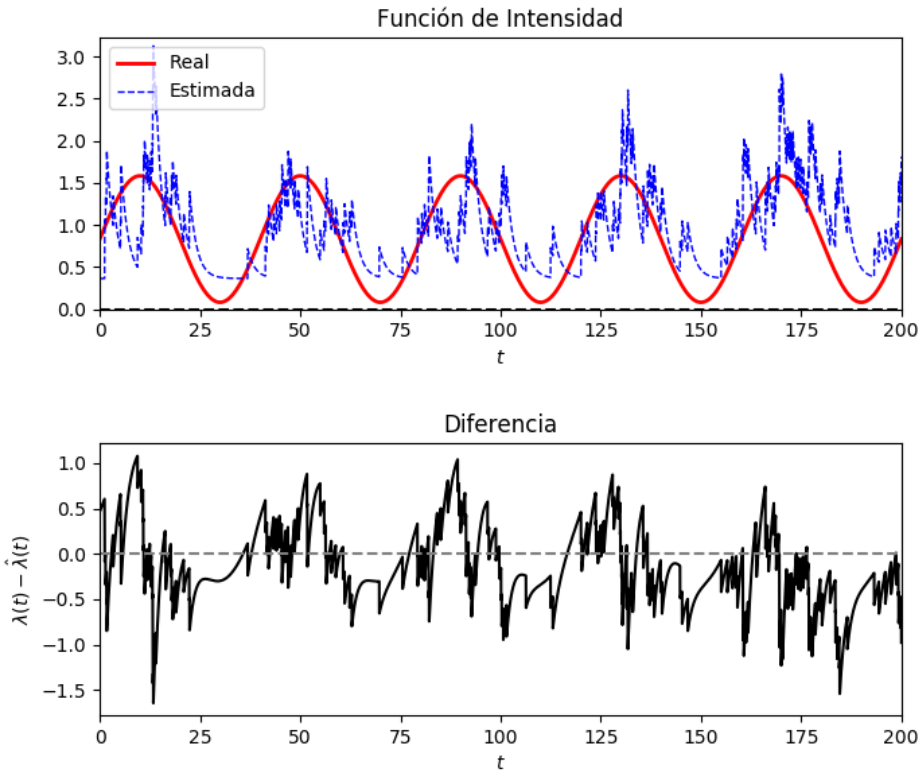


Figura 3.4: Comparativo entre la función de intensidad real y la estimada por máxima verosimilitud.

3.2.3. Otras funciones de excitación

Aunque los ejemplos mostrados se han centrado en el uso de la función de excitación con decrecimiento condicional, existe una amplia variedad de otras funciones que pueden ser igualmente seleccionadas. A continuación nombraremos solo un par de ellas.

En el mismo artículo de Hawkes se utiliza una forma más general de función de excitación, la cual consisten en tomar una mezcla de funciones exponenciales. De esta manera,

$$\phi(s) = \sum_{j=1}^p \alpha_j e^{-\beta_j s},$$

con $\sum_{j=1}^p \alpha_j / b_j < 1$. Sin embargo, esto también incrementa la complejidad del modelo, al tener un mayor número de parámetros que estimar.

En la literatura, también se ha empleado

$$\phi(s) = \frac{a}{(b + (t - s))^c},$$

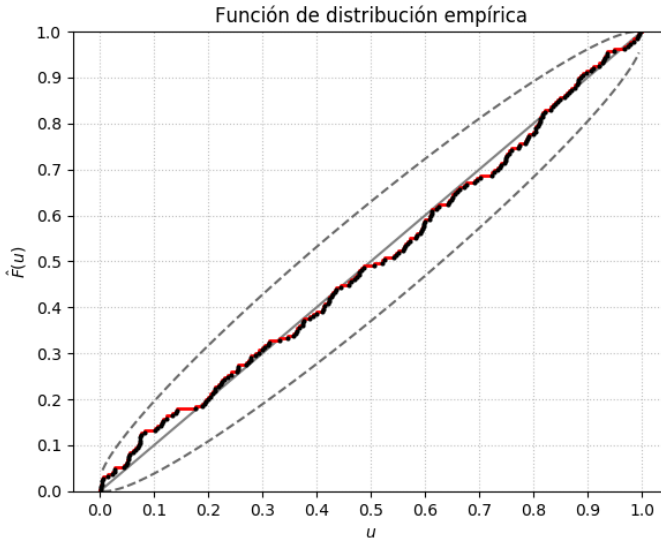


Figura 3.5: Función de distribución empírica de $\{u_1, \dots, u_{k-1}\}$.

para valores $a, b, c > 0$. La elección de esta función ha permitido la creación de modelos en el campo de la sismología, para mayor información se puede consultar [14].

3.3. El proceso de Hawkes multidimensional

De manera similar a como definimos un proceso puntual multidimensional en la Sección 1.10, ahora daremos una introducción al proceso de Hawkes multidimensional definido a través de su función de intensidad.

Definición 3.2 (Proceso de Hawkes multidimensional). Un proceso puntual regular d -dimensional $\{N_t^{(i)}; i \in \{1, \dots, d\}\}_{t \geq 0}$ es un *proceso de Hawkes multidimensional* si su función de intensidad está dada por $\lambda_t = \{\lambda_t^{(i)}; i \in \{1, \dots, d\}\}$, con

$$\lambda_t^{(i)} = \mu^i + \sum_{j=1}^d \int_{-\infty}^t \phi^{i,j}(t-s) dN_s^{(j)}, \quad i = 1, \dots, d, \quad (3.10)$$

donde $\mu = (\mu^i)_{1 \leq i \leq d}$, con $\mu^i \geq 0$, es un vector que representab la intensidad exógena al proceso, y $\phi = [\phi^{i,j}]_{i,j \in \{1, \dots, d\}}$ es llamada la matriz de kernels de Hawkes, con $\phi^{i,j}(t)$ una función real positiva.

Si además, ϕ tiene la forma

$$\phi^{ij}(s) = \alpha^{ij} e^{-\beta_j s},$$

entonces diremos que tenemos un proceso de Hawkes con decrecimientos exponenciales.

La función de intensidad con componentes dados por (3.10) puede ser escrita en forma matricial como

$$\lambda_t = \mu + \phi * dN_s,$$

donde $*$ denota la convolución de matrices.

Su simulación se puede hacer empleando el algoritmo de adelgazamiento, a partir de la realización de un proceso de Poisson de tasa $\lambda^*(t) \geq \sum_{i=1}^d \lambda_t^{(i)}$ y asociando una marca a cada punto generado, tal como se describió en la Sección 2.4.

3.3.1. Ejemplo

Consideremos un proceso de Hawkes bidimensional, $\{N_t^{(i)}; i = 1, 2, t \geq 0\}$, con función de excitación decreciente exponencial, esto es,

$$\begin{aligned}\lambda_t^{(1)} &= \mu_1 + \alpha_{11} \int_0^t e^{-\beta_1(t-s)} dN_s^{(1)} + \alpha_{12} \int_0^t e^{-\beta_1(t-s)} dN_s^{(2)}, \\ \lambda_t^{(2)} &= \mu_2 + \alpha_{21} \int_0^t e^{-\beta_2(t-s)} dN_s^{(1)} + \alpha_{22} \int_0^t e^{-\beta_2(t-s)} dN_s^{(2)},\end{aligned}$$

donde $\mu_1, \mu_2, \alpha_{11}, \alpha_{12}, \alpha_{21}, \alpha_{22}, \beta_1, \beta_2$ son todas constantes positivas.

Supongamos que $T^{(1)} = (t_1^{(1)}, \dots, t_{k_1}^{(1)})$, $T^{(2)} = (t_1^{(2)}, \dots, t_{k_2}^{(2)})$ son las observaciones en el intervalo $[0, \tau]$ del evento 1 y 2 respectivamente. Entonces, la función de log-verosimilitud de nuestro proceso queda, al desarrollar,

$$\begin{aligned}L(\mu_1, \mu_2, \alpha_{11}, \alpha_{12}, \alpha_{21}, \alpha_{22}, \beta_1, \beta_2; T^{(1)}, T^{(2)}) &= L^{(1)}(\mu_1, \alpha_{11}, \alpha_{12}, \beta_1) \\ &\quad + L^{(2)}(\mu_2, \alpha_{21}, \alpha_{22}, \beta_2),\end{aligned}$$

con

$$\begin{aligned}L^{(1)}(\mu_1, \alpha_{11}, \alpha_{12}, \beta_1) &= \sum_{i=2}^{k_1} \log[\mu_1 + \alpha_{11}R_{11}(i) + \alpha_{12}R_{12}(i)] - \mu_1\tau \\ &\quad - \frac{\alpha_{11}}{\beta_1} \sum_{i=1}^{k_1} \left(1 - e^{-\beta_1(\tau - t_i^{(1)})}\right) \\ &\quad - \frac{\alpha_{12}}{\beta_1} \sum_{j=1}^{k_2} \left(1 - e^{-\beta_1(\tau - t_j^{(2)})}\right),\end{aligned}$$

y

$$\begin{aligned}L^{(2)}(\mu_2, \alpha_{21}, \alpha_{22}, \beta_2) &= \sum_{j=2}^{k_2} \log[\mu_2 + \alpha_{21}R_{21}(j) + \alpha_{22}R_{22}(j)] - \mu_2\tau \\ &\quad - \frac{\alpha_{21}}{\beta_2} \sum_{j=1}^{k_2} \left(1 - e^{-\beta_2(\tau - t_j^{(2)})}\right) \\ &\quad - \frac{\alpha_{22}}{\beta_2} \sum_{i=1}^{k_1} \left(1 - e^{-\beta_2(\tau - t_i^{(1)})}\right),\end{aligned}$$

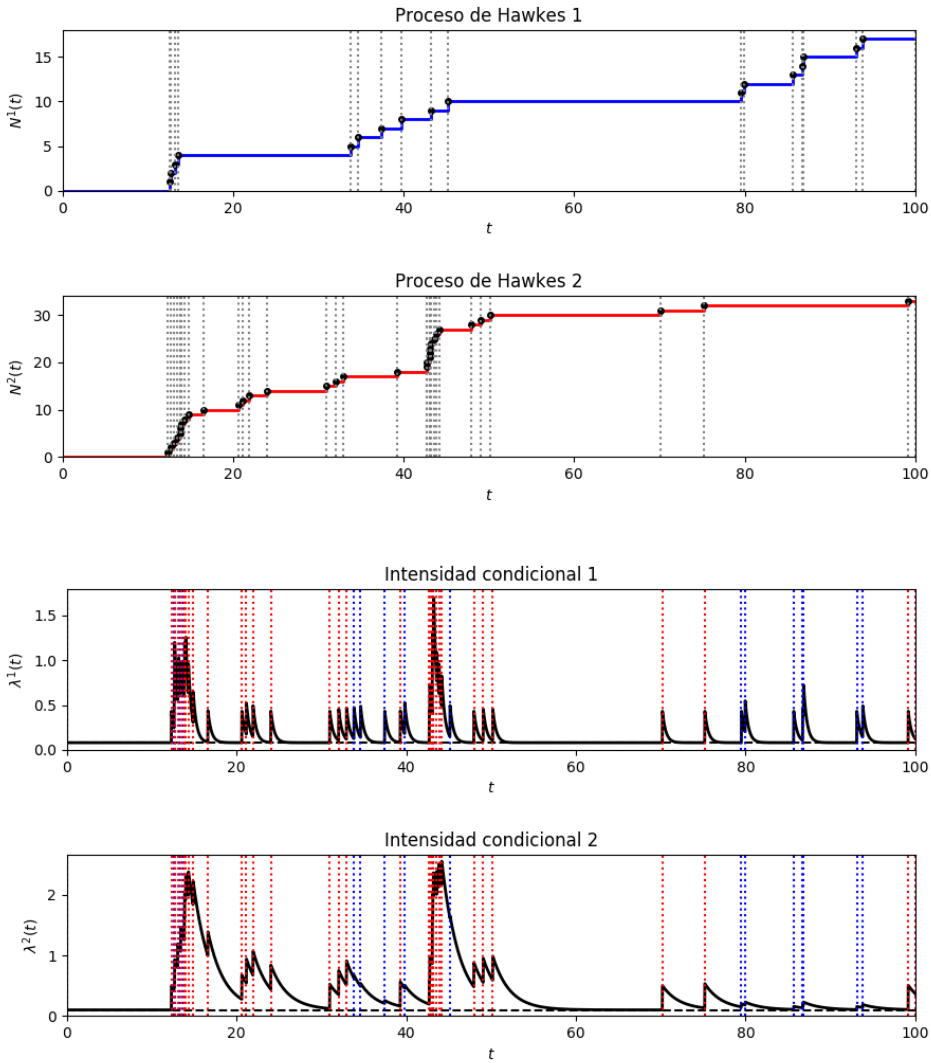


Figura 3.6: Proceso de Hawkes bidimensional, en la parte superior se observan los dos procesos de conteo asociados al modelo, mientras que en la parte inferior se muestra la función de intensidad de cada uno.

donde

$$\begin{aligned}
 R_{11}(1) &= R_{12}(1) = R_{21}(1) = R_{22}(1) = 0, \\
 R_{11}(i) &= (1 + R_{11}(i-1)) e^{-\beta_1(t_i^{(1)} - t_{i-1}^{(1)})}, \\
 R_{12}(i) &= (R_{12}(i-1)) e^{-\beta_1(t_i^{(1)} - t_{i-1}^{(1)})} + \sum_{\{j: t_{i-1}^{(1)} \leq t_j^{(2)} < t_i^{(1)}\}} e^{-\beta_1(t_i^{(1)} - t_j^{(2)})},
 \end{aligned}$$

$$R_{21}(j) = (R_{21}(j-1)) e^{-\beta_2(t_j^{(2)} - t_{j-1}^{(2)})} + \sum_{\{i: t_{j-1}^{(2)} \leq t_i^{(1)} < t_j^{(2)}\}} e^{-\beta_2(t_j^{(2)} - t_i^{(1)})},$$

$$R_{22}(j) = (1 + R_{22}(j-1)) e^{-\beta_2(t_j^{(2)} - t_{j-1}^{(2)})},$$

para $i = 2, \dots, k_1$ y $j = 2, \dots, k_2$. Más detalles de estos cálculos se pueden encontrar en [13].

Consideremos los siguientes valores para nuestros parámetros

$$\begin{array}{llll} \mu_1 = 0.08 & \alpha_{11} = 0.35 & \alpha_{12} = 0.35 & \beta_1 = 2.5 \\ \mu_2 = 0.1 & \alpha_{21} = 0.05 & \alpha_{22} = 0.4 & \beta_2 = 0.5 \end{array}$$

Tener $\alpha_{11} = \alpha_{12}$, provoca que la intensidad de $N^{(1)}$ se vea igualmente afectada al ocurrir un evento de 1 o de 2, al menos instantáneamente ya que un mayor valor en β_1 provoca una caída de la intensidad a su estado inicial a mayor velocidad que en los eventos de 2. Todo esto se puede apreciar en la Figura 3.6, donde se muestra una realización del proceso de Hawkes bidimensional aquí planteado.

3.4. Relación del proceso de Hawkes con los procesos de ramificación

Otra forma útil de pensar en un proceso de Hawkes es a través de un proceso de ramificación a tiempo continuo, más específicamente, un proceso de Poisson con agrupamientos o *clusters*. Para esto, clasifiquemos los eventos en dos tipos: inmigrantes y descendientes. Supongamos que los inmigrantes llegan a nuestro sistema siguiendo un proceso de Poisson con función de intensidad localmente integrable μ_t , para $t \geq 0$.

Al conjunto de todos los inmigrantes lo llamaremos Generación 0. Cada uno de ellos generará un número aleatorio de descendientes que conformarán la Generación 1, que a su vez tendrán descendientes formando la Generación 2, y así sucesivamente. Los descendientes de las diferentes generaciones del i -ésimo inmigrante forman un *cluster* al que denotaremos C_i , y cada elemento $t_j \in C_i$ generará a sus descendientes a través de un proceso de Poisson con función de intensidad $\phi(t - t_j)$, para $t > t_j$.

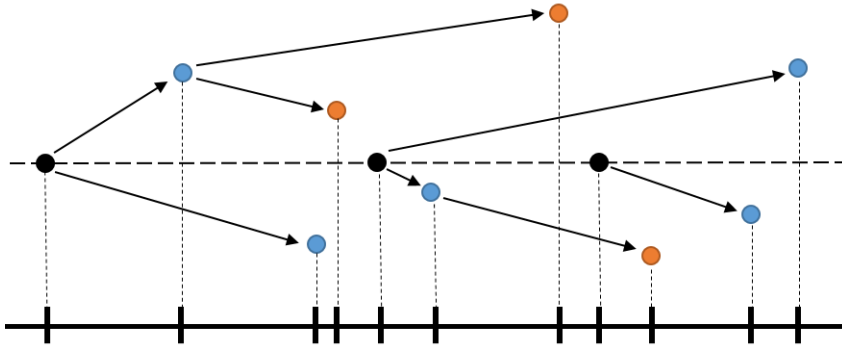


Figura 3.7: Esquema del proceso de Hawkes visto como un proceso de ramificación.

Los clusters, dados los inmigrantes, serán independientes y seguirán la misma estructura relativa al inmigrante que los generó.

Proposición 3.3. Cuando $\mu_t = \mu$ para todo $t > 0$. Sea T la unión de todos los clusters que generaron los inmigrantes,

$$T = \bigcup_i C_i,$$

entonces T forma un proceso de Hawkes unidimensional con función de intensidad dada por la ecuación (3.1).

Esta similitud nos permite aplicar al proceso de Hawkes ciertos resultados conocidos en la teoría de procesos de ramificación. Uno de los aspectos a destacar es que esta teoría nos brinda una justificación a la condición (3.2) de la Definición 3.1, ya que ϕ ya no solo representa la función de excitación, sino que también corresponde a la tasa de fertilidad de un proceso de ramificación. Es bien sabido que si la función de fertilidad de un proceso de ramificación no satisface

$$\int_0^\infty \phi(s) ds < 1,$$

entonces el proceso de ramificación asociado explotará con probabilidad 1.

Más resultados relacionados al proceso de Hawkes como proceso de ramificación, puede ser consultados a detalle en el Capítulo 4 de [12].

Bibliografía

- [1] D.R. Cox and Valerie Isham. *Point Processes*. Monographs on Statistics & Applied Probability. Chapman and Hall, 1980.
- [2] R.B. D'Agostino. *Goodness-of-Fit-Techniques*. Statistics: A Series of Textbooks and Monographs. Taylor & Francis, 1986.
- [3] Daryl J. Daley and David Vere-Jones. *An Introduction to the Theory of Point Processes*. Springer Series in Statistics. Springer, 1998.
- [4] Eloisa Díaz-Francés. Intervalos de Predicción Simultáneos de Cuantiles para Validación de Modelos Estadísticos. *Comunicación del CIMAT: I-18-01/07.02.2018 (PE/CIMAT)*.
- [5] Paul Embrechts, Thomas Liniger, and Lu Lin. Multivariate Hawkes Processes: An Application to Financial Data. *Journal of Applied Probability*, 48(A):367–378, 2011.
- [6] Alan G. Hawkes. Spectra of Some Self-Exciting and Mutually Exciting Point Processes. *Biometrika*, 58(1):83–90, 1971.
- [7] Paul G. Hoel. *Introduction to Mathematical Statistics*. Wiley Series in Probability and Statistics. John Wiley and Sons, 1954.
- [8] Martin Jacobsen. *Point Process Theory and Applications*. Probability and Its Applications. Birkhäuser, 2006.
- [9] Thibault Jaisson. *Market Activity and Price Impact Throughout Time Scales*. Theses, Ecole Polytechnique, October 2015.
- [10] Patrick J. Laub, Thomas Taimre, and Philip K. Pollett. Hawkes Processes.
- [11] P. A. W Lewis and G. S. Shedler. Simulation of Nonhomogeneous Poisson Processes by Thinning. *Naval Research Logistics Quarterly*, 26(3):403–413, 1979.
- [12] Thomas Liniger. *Multivariate Hawkes Processes*. PhD thesis, Department of Mathematics, ETH Zürich, 2009.
- [13] Y. Ogata. On Lewis' Simulation Method for Point Processes. *IEEE Transactions on Information Theory*, 27(1):23–31, Jan 1981.

-
- [14] Y. Ogata. Seismicity Analysis Through Point-Process Modeling: A Review. *Pure and Applied Geophysics*, 155(2):471–507, Aug 1999.
 - [15] T. Ozaki. Maximum Likelihood Estimation of Hawkes' Self-exciting Point Processes. *Annals of the Institute of Statistical Mathematics*, 31(1):145–155, 1979.
 - [16] C.P. Quesenberry and Craig Hales. Concentration Bands for Uniformity Plots. *Journal of Statistical Computation and Simulation*, 11(1):41–53, 1980.
 - [17] Rolf-Dieter Reiss. *A Course on Point Processes*. Springer Series in Statistics. Springer, 1993.
 - [18] I. Rubin. Regular Point Processes and Their Detection. *IEEE Transactions on Information Theory*, 18(5):547–557, Sep 1972.