

~~T. P. 001 33/73~~  
5/75

BIBLIOTECA UNIVERSITARIA
GRANADA
Nº Documento 1013384363
Nº Copia 15555604

UNIVERSIDAD DE GRANADA

FACULTAD DE CIENCIAS



UNIVERSIDAD DE GRANADA
Facultad de Ciencias
Fecha 21/5/99
ENTRADA NUM. 1705

DEPARTAMENTO DE ESTADÍSTICA E I.O.

PREDICIÓN CON MODELOS ESTOCÁSTICOS DE  
TIPO POISSONIANO

TESIS DOCTORAL

Paula Rodríguez Bouzas

Granada, mayo de 1999



TESIS DOCTORAL

PREDICCIÓN CON MODELOS  
ESTOCÁSTICOS DE TIPO POISSONIANO

Memoria que para optar al grado de doctor en Ciencias,  
Sección Matemáticas, presenta:



Fdo.: Paula Rodríguez Bouzas

Granada, mayo de 1999



UNIVERSIDAD DE GRANADA  
DEPARTAMENTO DE ESTADÍSTICA E I.O.

**D. MARIANO J. VALDERRAMA BONNET**

Catedrático de Universidad del área de Estadística e Investigación Operativa en la Universidad de Granada.

**DÑA. ANA M<sup>a</sup>AGUILERA DEL PINO**

Profesora titular del área de Estadística e Investigación Operativa en la Universidad de Granada.

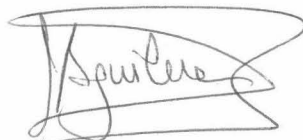
**CERTIFICAN:**

Que la presente Tesis Doctoral titulada "**Predicción con Modelos Estocásticos de Tipo Poissoniano**" ha sido realizada por Dña. Paula Rodríguez Bouzas, bajo nuestra inmediata dirección, durante el período 1996-99, siendo expresión de la capacidad investigadora e interpretativa de su autor, en condiciones que, a nuestro juicio, la hacen merecedora del título de Doctora en Ciencias, sección Matemáticas.

Granada, 12 de mayo de 1999



Dr. Mariano J. Valderrama Bonnet



Dra. Ana M<sup>a</sup> Aguilera del Pino



# Índice General

Prefacio	v
<b>1 Proceso de Poisson</b>	<b>1</b>
1.1 Definición y conceptos básicos . . . . .	1
1.2 Caracterizaciones de un proceso de Poisson . . . . .	7
1.2.1 Función y funcional característico . . . . .	7
1.2.2 Teoremas de caracterización de un proceso de Poisson . . . . .	8
1.3 Estadísticos de tiempo . . . . .	18
1.3.1 Distribución de los tiempos de ocurrencias . . .	19
1.3.2 Distribución de los tiempos entre ocurrencias . .	23
1.3.3 Función de densidad muestral y de verosimilitud	25
1.4 Proceso de Poisson según la Teoría de Martingalas . . .	29
1.4.1 Conceptos básicos . . . . .	30
1.4.2 El proceso de Poisson es una semimartingala . .	36
<b>2 Generalizaciones del proceso de Poisson</b>	<b>41</b>
2.1 Proceso de Poisson multidimensional . . . . .	42
2.2 Proceso de Poisson trasladado . . . . .	44
2.3 Proceso de Poisson compuesto . . . . .	51
2.3.1 Estadísticos de un proceso de Poisson compuesto	53

2.3.2	Casos particulares notables . . . . .	58
2.4	Proceso de Poisson doblemente estocástico . . . . .	61
2.4.1	Estadísticos de recuento . . . . .	63
2.4.2	Teoremas límite para los estadísticos de recuento	69
2.4.3	Estadísticos de tiempo . . . . .	72
2.4.4	El P.P.D.E. es una semimartingala . . . . .	73
2.5	P.P.D.E. con media aleatoria Normal truncada . . . . .	74
2.5.1	Función masa de probabilidad . . . . .	74
2.5.2	Función característica . . . . .	76
2.5.3	Estimación de la moda . . . . .	79
2.5.4	Aplicación: número de efectos de comercio de- vuelto o impagados . . . . .	84
<b>3</b>	<b>Predicción de procesos de Poisson mediante ACP</b>	<b>89</b>
3.1	Regresión multivariante en componentes principales . .	89
3.2	Regresión funcional en componentes principales . . . .	91
3.2.1	Estimación . . . . .	94
3.2.2	Aproximación de estimadores en el ACP de pro- cesos con funciones muestrales escalonadas . . .	98
3.3	Modelos RMCP y PCP para procesos de Poisson . . .	100
3.3.1	Descripción de los datos . . . . .	102
	<b>Apéndice</b>	<b>111</b>

# Prefacio

Esta memoria pretende dar una visión general de un tipo fundamental de proceso de recuento, el *proceso de Poisson*. Fundamental, por la razón de que su caso más simple, el *proceso de Poisson homogéneo*, es siempre referencia y punto de partida, para el estudio de cualquier proceso de recuento. Además, tanto éste como sus extensiones, son modelos matemáticos que se ajustan bien a numerosos fenómenos reales en variados campos de las ciencias. Debido a esto, es un proceso de recuento muy estudiado, lo que no mermó nuestro interés por él, pues la mayoría de los trabajos acerca de este tipo de proceso estocástico, hasta hace un par de décadas, versó casi exclusivamente en el caso de proceso de *Poisson homogéneo* y *no homogéneo*.

Asimismo se encuentra numerosa bibliografía clásica que aborda algunas extensiones del proceso de Poisson, como es el caso del *proceso de Poisson mixto* o el *compuesto*. También Cox (1955) trató el caso de proceso de Poisson doblemente estocástico, y a él se debe que se le denomine también *proceso de Cox*. Tal vez se deba a la complejidad de cálculo, más que teórica, que conllevan los casos más generales de procesos de tipo Poissoniano, el hecho de que no se haya profundizado más en su estudio.

De entre las generalizaciones del proceso de Poisson, esta memoria se detiene en la exposición del *proceso de Poisson doblemente estocástico*, adentrándonos especialmente en un caso particular de este pro-

ceso Poissoniano, aquél cuya media es aleatoria con distribución una Normal truncada para cada instante de tiempo.

Hemos organizado los objetivos de esta memoria de la siguiente manera. En el Capítulo 1, se introduce el proceso de recuento base de este trabajo, el proceso de Poisson, y se estudian sus características. Puesto que el *proceso de Poisson homogéneo* (con intensidad constante) es caso particular del *no homogéneo* (cuya intensidad es una función del tiempo), que se define con los axiomas clásicos del proceso de Poisson, nos referiremos, a partir de ahora, al proceso de *Poisson no homogéneo* simplemente como *proceso de Poisson*. En este capítulo se hace hincapié en los teoremas de caracterización del proceso de Poisson, especialmente en el teorema menos restrictivo, así como en ciertas propiedades, pues hemos intentado, o bien hacer una adaptación a los axiomas que tomamos en la definición del proceso de Poisson, o bien dar una demostración más detallada por no ser habitual en la literatura. Como última sección del Capítulo 1, se hace una aproximación al proceso de Poisson a través de la teoría de martingalas por ser el punto de vista más actual de los procesos de recuento y potente herramienta para su estudio como explicamos en esa sección, así como en la sección análoga del Capítulo 2, dedicado al *proceso de Poisson doblemente estocástico*.

Dedicamos el Capítulo 2 de este trabajo al estudio de procesos de tipo Poissoniano, es decir, a generalizaciones del proceso de Poisson. En cada nuevo tipo de proceso que se trata en este capítulo se relaja algunas de las propiedades del proceso de Poisson. En la primera sección dejamos de restringirnos a la recta real como espacio en el que se consideran las ocurrencias del proceso puntual, dando lugar al *proceso de Poisson multidimensional*. En la sección segunda se estudia qué ocurre cuando los sucesos de un proceso de Poisson, incluso en el caso multidimensional, son trasladados a otro espacio, apareciendo entonces el que llamaremos *proceso de Poisson trasladado*. El mayor interés de este nuevo proceso es que se demuestra que es a su vez

un proceso de Poisson, así como que cualquier proceso de Poisson se puede interpretar como trasladado si tenemos en cuenta que al trabajar con un proceso real, siempre es susceptible de tener errores en su observación y por tanto, de ser recogido no el proceso original sino otro similar con alguna que otra inserción o supresión de puntos. Si relajamos la propiedad de regularidad del proceso de Poisson, obtenemos el *proceso de Poisson compuesto* que recogemos en la sección tercera de este segundo capítulo. Se denomina también a este proceso poissoniano, *proceso de Poisson marcado*. Este nombre es quizás más intuitivo pues se trata de un proceso de Poisson cuyas ocurrencias nos interesan no sólo como tales, sino que nos interesa además alguna característica que llevan asociada, que será la marca para cada punto. Como caso particular del *proceso de Poisson compuesto* encontramos aquel con ocurrencias simultáneas o el resultante de la mezcla de varios procesos de Poisson con diferentes intensidades.

El *proceso de Poisson doblemente estocástico*, que abordamos en la sección cuarta del Capítulo 2, es el proceso poissoniano que analizamos con mayor profundidad. Es la generalización del proceso de Poisson en la que la intensidad es a su vez un proceso estocástico por la influencia de algún proceso externo al proceso de recuento. En la literatura se conoce también como *proceso de Cox* como hemos explicado antes, pero también se encuentra con el nombre de proceso de Poisson condicionado ya que condicionando al proceso externo que determina la intensidad, se tiene un proceso de Poisson. Es decir, conocido ese proceso externo, tenemos un proceso de Poisson, de ahí que se denomine a ese proceso externo proceso información. Las características de un *proceso de Poisson doblemente estocástico* son fáciles de exponer teóricamente debido a esa propiedad de condicionamiento de la que hablábamos, pero intratables, por lo general, en la práctica. Se encuentra en la literatura gran escasez de ejemplos concretos de procesos de Poisson doblemente estocásticos; suele reducirse al caso en el que la intensidad es una variable aleatoria, el producto de una función

del tiempo por una variable aleatoria o el proceso información es de tipo gaussiano o markoviano. Nosotros vemos como ejemplo el *proceso de Poisson doblemente estocástico cuya intensidad es un proceso de banda estrecha*, hallando sus características gracias a su desarrollo en serie de Karhunen-Loève.

Ocurre también que en el marco del *proceso de Poisson doblemente estocástico* se trabaja siempre desde la asunción de algún conocimiento de la intensidad del proceso. En esta memoria proponemos la aproximación al proceso a través de su función paramétrica, es decir, la integral de la intensidad (por lo que trivialmente es a su vez un proceso estocástico), que no es más que la media del *proceso de Poisson doblemente estocástico*. De esta forma vemos como tanto la función masa de probabilidad como la función característica del proceso pueden ser más abordables. De hecho, en el caso de un *proceso de Poisson doblemente estocástico* cuya media sea un proceso estocástico de forma que para cada instante de tiempo se tenga una distribución Normal truncada (ejemplo que estudiamos y comentaremos más adelante) pueden obtenerse a pesar de su complejidad, la función masa de probabilidad y función característica del *proceso de Poisson doblemente estocástico*.

Otra medida que se toma para evitar la complejidad de las expresiones de las características del *proceso de Poisson doblemente estocástico* es el estudio de su comportamiento cuando el tiempo tiende a infinito. Los teoremas dedicados a ello se han demostrado partiendo de la demostración encontrada en Snyder y Miller (1991) y el caso particular encontrado en Grandell (1971).

Comentábamos anteriormente que la teoría de martingalas es una potente herramienta en el tratamiento de procesos de recuento; esto se ve claramente en el hecho de que el punto donde tratamos al *proceso de Poisson doblemente estocástico* con esta teoría es tan escueto debido a que basta con extender de forma sencilla lo visto para el proceso de Poisson.

El mayor aporte de este trabajo se encuentra en el estudio de un

*proceso de Poisson doblemente estocástico cuya media es una Normal truncada*, resumiendo lo dicho anteriormente, por lo que hemos habilitado una nueva sección, la quinta, para este proceso aunque sea un caso particular del *proceso de Poisson doblemente estocástico*. La idea de estudiar el *proceso de Poisson doblemente estocástico* considerando que tenemos algún tipo de información de su media con lo cual sus características podían ser más abordables desde el punto de vista de su cálculo, nos incitó a pensar en qué tipo de distribución podía ser lógica en algunos casos reales y pensamos en la Normal truncada. Decíamos que a menudo se recoge en la literatura el ejemplo de *proceso de Poisson doblemente estocástico* cuya intensidad en una variable aleatoria o una función del tiempo por una variable aleatoria. Estos casos no deben confundirse con el tratado en la sección quinta del capítulo 2. En nuestro caso se tiene que la media es una variable aleatoria cuya distribución es una Normal truncada, pero no necesariamente la misma para cada instante de tiempo. Se dan expresiones explícitas de características de este proceso como son su función masa de probabilidad, su función característica y a partir de una descomposición de ésta, pueden encontrarse también las expresiones de la esperanza y la varianza del proceso.

La meta final del estudio de un determinado modelo estadístico dinámico es la estimación en el futuro. Con ese fin, introducimos un método de estimación de la moda del proceso de *Poisson doblemente estocástico con media aleatoria Normal truncada* que consiste en estimar cuál sería la distribución de la media del proceso en un instante futuro, por lo que se podría conocer la función masa de probabilidad del *proceso de Poisson doblemente estocástico* en ese instante. Demostramos que su máximo pertenece a un cierto intervalo acotado y así puede encontrarse el valor de este *proceso de Poisson doblemente estocástico* con mayor probabilidad. Por último, aplicamos este método a un caso real, el número de efectos de comercio devueltos o impagados en España.



De forma complementaria, en el Capítulo 3 se aplicará la técnica de reducción de dimensión Análisis en Componentes Principales, tanto en sus versiones multivariante como funcional, para predecir la evolución temporal de procesos Poissonianos en un período de tiempo futuro a partir de su pasado. Para ello, se propone una adaptación a este tipo de procesos de los modelos de Regresión Multivariante en Componentes Principales y de los modelos de Regresión Funcional en Componentes Principales, cuya capacidad predictiva será evaluada con los mismos datos de efectos de comercio devueltos o impagados utilizados en el anterior capítulo. Las predicciones así obtenidas nos permiten, además, proponer un paso alternativo en el método de predicción de la moda desarrollado en el Capítulo 2.

Desearía, antes de finalizar, manifestar mi profundo agradecimiento a los directores de esta memoria, el Profesor Dr. D. Mariano J. Valderrama Bonnet y la Profesora Dra. Dña. Ana M<sup>a</sup> Aguilera del Pino, por su ayuda no sólo profesional sino también por sus desvelos y ánimo constantes.

Granada, mayo de 1999.

# Capítulo 1

## Proceso de Poisson

### 1.1 Definición y conceptos básicos

Un fenómeno físico, económico, etc. caracterizado por sucesos localizados aleatoriamente en un espacio continuo,  $\mathcal{X}$ , ya sea el tiempo, plano o incluso un espacio de mayor dimensión, se modeliza matemáticamente mediante un proceso puntual. Cada suceso se representa por un punto identificado con una posición en ese conjunto  $\mathcal{X}$ .

El número de fenómenos que pueden modelizarse mediante un proceso puntual es inmenso. Algunos ejemplos en diferentes espacios pueden ser:

- el número de emisiones de electrones desde un cátodo, en el que  $\mathcal{X}$  es un intervalo de la recta real;
- número de descargas eléctricas en las fibras nerviosas, con el mismo espacio de localización de los sucesos que el ejemplo anterior;
- ruido atmosférico (descargas eléctricas que ocurren en la atmósfera produciendo interferencias en ondas de radio), en el que nuevamente tenemos el mismo espacio;

- número de estrellas en una región del espacio, en el que  $\mathcal{X}$  es una región es el espacio euclídeo  $\mathbb{R}^3$ ;
- estudio de movimientos sísmicos; si el interés es además del número de sucesos, el lugar geométrico del epicentro, el espacio de realización del proceso puntual es el producto cartesiano de un intervalo de la recta real representando al tiempo y una región de  $\mathbb{R}^3$  correspondiente con la Tierra;
- número de reclamaciones en una compañía de seguros, en el que el espacio es nuevamente un intervalo de  $\mathbb{R}$ , etc.;

se pueden encontrar más ejemplos y detalles en numerosa bibliografía, ( Snyder y Miller, 1991) , ( Grigoriu, 1995) y ( Andersen et al. (1993)).

El proceso de Poisson es el más simple de los procesos puntuales, lo que no significa que no resuelva satisfactoriamente muchos problemas reales. Este modelo es usado ampliamente para la modelización en fiabilidad de sistemas reparables o de software considerados complejos ( Härtler, 1989), ( Zhao y Xie, 1996) e incluso enfatizando sus virtudes en una carta al editor de *Microelectronics* ( Ascher, 1989).

El caso de proceso de Poisson más extensa y habitualmente estudiado es aquel cuyo espacio  $\mathcal{X}$  en el que ocurren los sucesos es el tiempo. En este caso, el *proceso de Poisson puntual* ( Last y Brandt, 1995) se define como el conjunto de variables aleatorias positivas  $\{w_n; n \geq 1\}$ , donde  $w_n \in [t_0, \infty)$  se interpreta como el instante en el que ocurre el  $n$ -ésimo suceso verificando que  $w_1, w_2 - w_1, w_3 - w_2, \dots$  son independientes y distribuidos mediante una distribución exponencial con parámetro  $\lambda(t) \geq 0$ , denominada intensidad del proceso puntual.

Asociado a un proceso puntual siempre tenemos el llamado *proceso de recuento* que no es más que un acercamiento al mismo fenómeno desde otro punto de vista. Éste es el recuento del número de sucesos que ocurren en subconjuntos del espacio  $\mathcal{X}$  en el que está definido el

proceso puntual. En la práctica, puede ser muy conveniente mirar a un proceso puntual de esta manera.

Tomando nuevamente como  $\mathcal{X}$  el tiempo, es decir la semirrecta  $[t_0, \infty)$ , el proceso de Poisson visto como proceso de recuento, estudiará el número de sucesos que ocurren desde  $t_0$  hasta un instante de tiempo  $t$ . Se puede definir según diferentes conjuntos de axiomas; nosotros usaremos la siguiente definición (Snyder y Miller, 1991).

**Definición 1.1.1** *Un proceso de Poisson temporal,  $\{N(t), t \geq t_0\}$ , es aquel proceso de recuento que verifica:*

- i.  $P[N(t_0) = 0] = 1$ ;
- ii. para  $t_0 \leq s \leq t$ , el incremento  $N(s, t) = N(t) - N(s)$  se distribuye mediante una distribución de Poisson de parámetro  $\Lambda(t) - \Lambda(s)$ ,

$$P[N(s, t) = n] = \frac{1}{n!} [\Lambda(t) - \Lambda(s)]^n e^{-[\Lambda(t) - \Lambda(s)]}, \quad (1.1)$$

para  $n = 0, 1, 2, \dots$ , donde  $\Lambda(t)$  es finita valuada, no negativa y no decreciente;

- iii.  $\{N(t), t \geq t_0\}$  tiene incrementos independientes.

Aunque hemos definido el proceso de Poisson con la condición (i) de que  $P[N(t_0) = 0] = 1$ , no es necesario. Si  $N(t_0)$  es una variable aleatoria entera y no negativa,  $N(t_0) = k$ , entonces:

$$P[N(t) = n] = \sum_{k=0}^{\infty} P[N(t_0, t) = n - k] P[N(t_0) = k]$$

pues  $N(t) = N(t_0) + N(t_0, t)$  y el proceso tiene incrementos independientes.

La propiedad (iii) es la propiedad distintiva. Nos dice que en un proceso de recuento de Poisson, el número de puntos en intervalos disjuntos, es decir, que no se solapen, es estadísticamente independiente sin importar su longitud o su cercanía; de forma rigurosa, si  $[t_i, u_i)$ , para  $i = 1, 2, \dots, k$  son intervalos disjuntos en  $[t_0, \infty)$ , entonces:

$$P[N(t_1, u_1) = n_1, N(t_2, u_2) = n_2, \dots, N(t_k, u_k) = n_k] = \prod_{i=1}^k P[N(t_i, u_i) = n_i]$$

para  $k = 1, 2, \dots$

El conjunto de las tres propiedades determinan completamente al proceso de Poisson, es decir, a su probabilidad conjunta; ésta resulta

$$P[N(t_1) = n_1, N(t_2) = n_2, \dots, N(t_k) = n_k] = \prod_{i=1}^k P[N(t_{i-1}, t_i) = n_i - n_{i-1}]$$

para cualquier conjunto de instantes  $t_1, t_2, \dots, t_k$  en el intervalo  $[t_0, \infty)$  y enteros no negativos  $n_1, n_2, \dots, n_k$  donde  $n_0 = 0$  y  $k = 1, 2, \dots$

La función  $\Lambda$  de la propiedad (ii) se denomina *función paramétrica* del proceso de Poisson. Las únicas restricciones que se le han exigido son que sea no negativa y no decreciente. No es necesario que  $\Lambda$  sea continua, pero aunque lo sea, tampoco se le exige que sea diferenciable.

Del estudio detallado de la propiedad (ii) se obtienen las siguientes conclusiones bien conocidas.

1. Si  $\Lambda(t) - \Lambda(s)$  es finito,

$$P[N(s, t) \geq 1] = 1 - P[N(s, t) = 0] = 1 - e^{[\Lambda(t) - \Lambda(s)]} < 1;$$

por tanto, en este caso los puntos no ocurren con certeza en el intervalo  $[s, t)$ .

Se concluye también fácilmente que

$$P[N(s, t) = n] \xrightarrow{n \rightarrow \infty} 0,$$

lo que significa que no pueden ocurrir infinitos puntos en  $[s, t)$ .

Otra conclusión conocida es que algunos puntos del proceso de Poisson ocurren en instantes predeterminados y es cuando hay una discontinuidad de salto de la función paramétrica. Como  $\Lambda(t)$  es finita y no decreciente, el conjunto de estos instantes predeterminados en  $[s, t)$  es numerable y la suma de los saltos es finita. Khinchin (1956), demostró que  $N(s, t)$  se puede descomponer en dos componentes estadísticamente independientes:  $N_d(s, t)$ , la componente singular y  $N_c(s, t)$ , la no singular. La primera es la contribución a  $N(s, t)$  de los puntos que ocurren en instantes predeterminados debido a las discontinuidades y  $N_c(s, t) = N(s, t) - N_d(s, t)$ , es la contribución a  $N(s, t)$  en los instantes de  $[s, t)$  donde  $\Lambda$  es continua.

2. Si la función paramétrica,  $\Lambda(t)$ , es continua se demuestra que:

$$\lim_{\delta \searrow 0} P[N(t, t + \delta) = 0] = 1, \forall t \geq t_0$$

y los puntos no ocurren en instantes predeterminados. Por tanto, en este caso, el proceso de Poisson  $\{N(t); t \geq t_0\}$  es estocásticamente continuo.

Si existe la derivada de la función paramétrica, es decir el límite

$$\lim_{\delta \searrow 0} \frac{\Lambda(t + \delta) - \Lambda(t)}{\delta} = \lambda(t)$$

que hemos notado con  $\lambda(t)$ , éste es el promedio de ocurrencias para el instante  $t$ , pues el número esperado de puntos en  $[s, t)$  es

(como volveremos a comentar más adelante),

$$E[N(s, t)] = \sum_{n=0}^{\infty} n P[N(s, t) = n] = \Lambda(t) - \Lambda(s) \quad (1.2)$$

sin más que sustituir  $P[N(s, t) = n]$  por su valor.

En caso de tener una función paramétrica continua, si ésta es además absolutamente continua, puede expresarse como

$$\Lambda(t) = \int_{t_0}^t \lambda(\sigma) d\sigma, \forall t \geq t_0,$$

donde  $\lambda(t)$ , definida anteriormente, es no negativa. Esta función es entonces denominada *función intensidad* del proceso de Poisson.

Se diferencian dos casos según sea la función intensidad. Si  $\lambda(t)$  es constante, es decir, no depende del tiempo, entonces se dice que el proceso de Poisson es *homogéneo*. Así, tendríamos que

$$\Lambda(t) - \Lambda(s) = \int_s^t \lambda(\sigma) d\sigma$$

es proporcional a  $t - s$ , por lo que es estadístico de recuento  $N(s, t)$  es el mismo que el  $N(s + \tau, t + \tau)$ , para todo  $\tau$  tal que  $s + \tau \geq t_0$ . En consecuencia, un proceso de Poisson homogéneo tiene incrementos independientes y estacionarios.

Si la intensidad del proceso,  $\lambda(t)$ , no es constante, el proceso se denomina *no homogéneo* (también se encuentra en la literatura como *proceso de Poisson general*, nosotros no usaremos ese nombre por evitar equívocos con otros procesos todavía más generales, resultantes de la relajación de propiedades del proceso de Poisson). En este caso, como hemos observado anteriormente, el proceso es estocásticamente continuo.

## 1.2 Caracterizaciones de un proceso de Poisson

### 1.2.1 Función y funcional característico

Como proceso estocástico que es, el proceso de Poisson se caracteriza mediante la función característica para el número de puntos en un intervalo cualquiera  $[s, t]$ .

Por definición de función característica,  $M_{s,t}(iv) \equiv E[e^{ivN(s,t)}]$ , donde  $i$  es la unidad imaginaria, y usando (1.1):

$$\begin{aligned} M_{s,t}(iv) &= \sum_{n=0}^{\infty} e^{ivn} P[N(s,t) = n] = \\ &= \sum_{n=0}^{\infty} \frac{1}{n!} [\Lambda(t) - \Lambda(s)]^n e^{-[\Lambda(t) - \Lambda(s)]} e^{ivn} = e^{-[\Lambda(t) - \Lambda(s)]} \exp[e^{iv}(\Lambda(t) - \Lambda(s))] \\ &= \exp[(e^{iv} - 1)(\Lambda(t) - \Lambda(s))]. \end{aligned}$$

Dada la propiedad de generación de momentos de la función característica, podemos calcular entre otros la esperanza y varianza del proceso de Poisson:

$$E[N(s,t)] = M_{s,t}^1(i0) = \Lambda(t) - \Lambda(s); \quad (1.3)$$

así mismo, el momento no centrado de segundo orden es

$$E[N^2(s,t)] = M_{s,t}^2(i0) = [\Lambda(t) - \Lambda(s)] + [\Lambda(t) - \Lambda(s)]^2$$

y por tanto, la varianza del número de puntos en el intervalo  $[s, t]$  es:

$$\text{Var}[N(s,t)] = E[N^2(s,t)] - E^2[N(s,t)] = \Lambda(t) - \Lambda(s) = E[N(s,t)], \quad (1.4)$$



así pues, la varianza de un proceso de Poisson en  $[s, t)$  coincide con el número de puntos esperados en ese intervalo.

Ya en 1935, Kolmogorov sugirió una generalización de la función característica, el *funcional característico*. Dada una variable aleatoria  $X$  de un espacio lineal  $L$  y  $\xi$  un elemento cualquiera de ese espacio, el funcional característico de  $X$  lo definió como

$$\Phi(\xi) = E[e^{i\langle X, \xi \rangle}]. \quad (1.5)$$

Recogiendo la sugerencia de Kolmogorov, el funcional característico de un proceso de Poisson  $\{N(s, t); t_0 \leq \sigma < T\}$  se define como la esperanza

$$\Phi_N(iv) = E \left[ \exp \left( i \int_{t_0}^T v(\sigma) N(d\sigma) \right) \right], \quad (1.6)$$

donde  $v$  es una función real. La integral es una integral llamada de recuento que se evalúa de la siguiente forma:

$$\int_{t_0}^T v(\sigma) N(d\sigma) = \begin{cases} 0, & N(T) = 0 \\ \sum_{i=1}^{N(T)} v(w_i), & N(T) \geq 1. \end{cases}$$

donde  $w_i$  son los tiempos de ocurrencia de los sucesos o puntos del proceso de Poisson. Veremos en la siguiente sección la expresión concreta que puede tomar el funcional característico al estudiar la propiedad de la distribución condicional de los tiempos de ocurrencia de los puntos del proceso, como consecuencia de la expresión de la función de densidad muestral; ver página 29.

### 1.2.2 Teoremas de caracterización de un proceso de Poisson

Los axiomas con los que definimos el proceso de Poisson no son prácticos a la hora de identificar un proceso de recuento observado, como

proceso de Poisson. Por este motivo buscaremos condiciones que lo caractericen y como último objetivo, condiciones cualitativas para reconocer un proceso arbitrario como proceso de Poisson.

Definamos en primer lugar tres conceptos que usaremos para lograr nuestro objetivo.

**Regularidad.** Un proceso de recuento  $\{N(t), t \geq t_0\}$  se dice *regular* en el instante  $t \geq t_0$ , si para un  $\varepsilon$  dado, existe  $\delta \equiv \delta(t, \varepsilon) > 0$  tal que

$$P[N(t, t + \delta') > 1] \leq \varepsilon P[N(t, t + \delta') = 1], \quad \forall \delta' \in (0, \delta). \quad (1.7)$$

Un proceso es *regular en un intervalo* de tiempo, si lo es en todos sus instantes y es *uniformemente regular* en el intervalo, si un único  $\delta \equiv \delta(\varepsilon)$  verifica la propiedad para todos los instantes del intervalo.

Si  $P[N(t, t + \delta') = 1]$  no pudiese anularse, la propiedad de regularidad podría escribirse:

$$\lim_{\delta \searrow 0} \frac{P[N(t, t + \delta) > 1]}{P[N(t, t + \delta) = 1]} = 0.$$

Intuitivamente, esta propiedad nos dice que dos puntos no ocurren simultáneamente; sin embargo, que los puntos sean distintos no implica regularidad.

En general el proceso de Poisson es regular cuando la función paramétrica  $\Lambda(t)$  es continua.

**Regularidad condicionada.** Notemos por  $P$  un suceso arbitrario determinado por las variables aleatorias  $\{N(\sigma); t_0 \leq \sigma < t\}$ . Así,  $P$  es cualquier suceso del pasado. Un proceso de recuento

$\{N(t), t \geq t_0\}$  es *condicionadamente regular* en el instante  $t \geq t_0$ , si para cualquier  $P$  y  $\varepsilon > 0$ , existe  $\delta \equiv \delta(t, \varepsilon) > 0$  tal que

$$P[N(t, t + \delta') > 1/P] \leq \varepsilon P[N(t, t + \delta') = 1/P], \quad \forall \delta' \in (0, \delta). \quad (1.8)$$

El *proceso será condicionadamente regular* si lo es para todo  $t \geq t_0$ .

Un proceso condicionadamente regular es regular, pero el inverso no es cierto.

Un proceso de Poisson, como tiene incrementos independientes, si su función paramétrica es continua, es un proceso condicionadamente regular.

**Proceso sin memoria.** Un proceso puntual en  $[t_0, \infty)$  se dice que no tiene memoria si para cualquier  $t \geq t_0$ , la ocurrencia de puntos durante  $[t, \infty)$  no depende de los que ocurrieron en  $[t_0, t)$ . Esto es, que el futuro no depende del pasado; rigurosamente expresado:

Sea  $P$  un suceso arbitrario determinado por las variables aleatorias  $\{N(\sigma); t_0 \leq \sigma < t\}$ . Asimismo, sea  $F$  un suceso arbitrario asociado a las variables aleatorias  $\{N(\sigma); \sigma \geq t\}$ , donde  $t$  denota el instante presente. Entonces,

$$P[F/P] = P[F], \quad \forall t \geq t_0 \quad (1.9)$$

es decir,  $P$  y  $F$  son independientes.

Un proceso con incrementos independientes no tiene memoria y viceversa. Según esto, un proceso de Poisson no tiene memoria.

Ya definidas las propiedades anteriores podemos enunciar el siguiente teorema de caracterización.

**Teorema 1.2.1** Sea  $\{N(t), t \geq t_0\}$  un proceso de recuento que verifica  $P[N(s, t) = n] \xrightarrow{n \rightarrow \infty} 0$  y  $P[N(t, t + \delta) = 0] = 1$ . Entonces:

$N(t)$  es un proceso de Poisson si y sólo si es regular y sin memoria.

**Demostración.** Si el proceso es sin memoria es equivalente a que tenga incrementos independientes. Se demuestra además:

$$\left. \begin{array}{l} i) N(\cdot) \text{ verifica } P[N(s, t) = n] \xrightarrow{n \rightarrow \infty} 0 \text{ y} \\ P[N(t, t + \delta) = 0] = 1 \\ ii) N(\cdot) \text{ tiene incrementos independientes} \end{array} \right\} \iff \text{ existe } \Lambda \text{ con-}$$
 tinua y acotadamente finita tal que  $P[N(s, t) = 0] = e^{-[\Lambda(t) - \Lambda(s)]}$ , para el intervalo  $[s, t)$ .

También se tiene que, siendo el proceso regular, existe  $\Lambda$  en las condiciones anteriores si y sólo si el proceso es de Poisson. ■

El teorema anterior ha sido adaptado a los axiomas que nosotros hemos tomado como definición de proceso de Poisson. Un estudio en profundidad de la demostración puede verse en ( Daley y Vere-Jones, 1988).

Con condiciones más generales y dando condiciones cualitativas, sin presunciones previas acerca del proceso, podemos demostrar el siguiente teorema de caracterización enunciado en ( Snyder y Miller, 1991).

**Teorema 1.2.2** Sea  $\{N(t) : t \geq t_0\}$  un proceso de recuento asociado a un proceso puntual en  $[t_0, \infty)$ .

$\{N(t) : t \geq t_0\}$  es un proceso de Poisson con función paramétrica absolutamente continua

$$\Lambda(t) = \int_{t_0}^t \lambda(\sigma) d\sigma.$$

si y sólo si verifica las siguientes condiciones:

- a. el proceso puntual es condicionadamente regular;
- b. para todo  $t \geq t_0$  y cualquier suceso arbitrario  $P$  asociado con las variables aleatorias  $\{N(\sigma) : t_0 \leq \sigma < t\}$ , el límite cuando  $\delta$  tiende a cero de  $\delta^{-1}P[N(t, t + \delta) = 1/P]$  existe, es finito e integrable en función de  $t$ . Notando esa función por  $\lambda(t)$ , tenemos:

$$\lambda(t) = \lim_{\delta \searrow 0} \frac{1}{\delta} P[N(t, t + \delta) = 1/P], \quad (1.10)$$

y la integral

$$\int_s^t \lambda(\sigma) d\sigma$$

existe y es finita para todo intervalo finito  $[s, t], t_0 \leq s \leq t$ ;

- c.  $P[N(t_0) = 0] = 1$ .

La hipótesis (b) implica no sólo que el límite existe, sino también que  $\{N(t) : t \geq t_0\}$  es un proceso sin memoria en un intervalo infinitamente pequeño. En la demostración del teorema concluiremos que el proceso no tiene memoria de forma global.

**Demostración.** La implicación hacia la derecha es trivial después de los comentarios hechos en cada una de las definiciones de las propiedades vistas en esta misma sección.

Para demostrar la implicación a la izquierda debemos probar que se cumplen las propiedades de la definición dada para un proceso de recuento de Poisson, por tanto debemos probar que los incrementos  $N(s, t) = N(t) - N(s)$  tienen distribución de Poisson con parámetro  $\Lambda(t)$  y que  $\{N(t); t \geq t_0\}$  tiene incrementos independientes. En primer lugar, hallaremos una ecuación en diferencias para  $P[N(s, t) = n]$  y luego veremos que la solución es una distribución de Poisson. Para

$n \geq 1$ , el suceso  $\{N(s, t + \delta) = n\}$  puede ocurrir de  $n + 1$  maneras mutuamente excluyentes, que son:

1.  $N(s, t) = n$  y  $N(t, t + \delta) = 0$
2.  $N(s, t) = n - 1$  y  $N(t, t + \delta) = 1$
- 
- 
- $n + 1$ .  $N(s, t) = 0$  y  $N(t, t + \delta) = n$ .

Por tanto,

$$\begin{aligned}
 P[s, t + \delta = n] &= \sum_{k=0}^n P[N(s, t) = n - k, N(t, t + \delta) = k] \quad (1.11) \\
 &= P[N(t, t + \delta) = 0/N(s, t) = n] P[N(s, t) = n] + \\
 &P[N(t, t + \delta) = 1/N(s, t) = n - 1] P[N(s, t) = n - 1] + \\
 &\sum_{k=0}^n P[N(t, t + \delta) = k/N(s, t) = n - k] P[N(s, t) = n - k]
 \end{aligned}$$

Debido a que el proceso es regularmente condicionado, dado un  $\varepsilon > 0$  y con un  $\delta > 0$ , tenemos que:

$$\begin{aligned}
 &P[N(t, t + \delta) > 1/N(s, t) = n - k] \quad (1.12) \\
 &\leq \varepsilon P[N(t, t + \delta) = 1/N(s, t) = n - k]
 \end{aligned}$$

y podemos escribir (1.10) de la siguiente manera:

$$-\varepsilon \leq \frac{1}{\delta} P[N(t, t + \delta) = 1/N(s, t) = P] - \lambda(t) \leq \varepsilon \quad (1.13)$$

y por tanto:

$$(\lambda(t) - \varepsilon)\delta \leq P[N(t, t + \delta) = 1/N(s, t) = P] \leq \delta(\varepsilon + \lambda(t)). \quad (1.14)$$

El último término de la ecuación (1.11) se puede acotar superiormente usando las conclusiones (1.12) y (1.14) de las hipótesis del teorema, usando como sucesos del pasado  $N(s, t) = n - k$ , con  $k = 2, \dots, n$  en cada sumando de la sumatoria, de la siguiente forma:

$$\begin{aligned}
& \sum_{k=2}^n P[N(t, t + \delta) = k / N(s, t) = n - k] P[N(s, t) = n - k] \\
& \leq \sum_{k=2}^n P[N(t, t + \delta) > 1 / N(s, t) = n - k] P[N(s, t) = n - k] \\
& \leq \varepsilon \sum_{k=2}^n P[N(t, t + \delta) = 1 / N(s, t) = n - k] P[N(s, t) = n - k] \\
& \leq \varepsilon \sum_{k=2}^n \delta(\varepsilon + \lambda(t)) P[N(s, t) = n - k] \\
& = [\delta\varepsilon^2 + \delta\varepsilon\lambda(t)] \sum_{k=2}^n P[N(s, t) = n - k] \\
& \leq [\delta\varepsilon^2 + \delta\varepsilon\lambda(t)] 1 = \delta\varepsilon^2 + \delta\varepsilon\lambda(t).
\end{aligned}$$

Ahora acotaremos inferior y superiormente  $P[N(t, t + \delta) = 0 / N(s, t) = n]$ . Para la cota inferior, haremos uso de las conclusiones (1.12) y (1.14), con el suceso del pasado  $N(s, t) = n$ , como sigue:

$$\begin{aligned}
& P[N(t, t + \delta) = 0 / N(s, t) = n] \\
& = 1 - P[N(t, t + \delta) = 1 / N(s, t) = n] - P[N(t, t + \delta) > 1 / N(s, t) = n] \\
& \geq 1 - \varepsilon P[N(t, t + \delta) = 1 / N(s, t) = n] - P[N(t, t + \delta) = 1 / N(s, t) = n] \\
& \geq 1 - (1 + \varepsilon)[\delta\varepsilon + \delta\lambda(t)].
\end{aligned}$$

Para la cota superior, tendremos en cuenta (1.14):

$$\begin{aligned} & P[N(t, t + \delta) = 0/N(s, t) = n] \\ &= 1 - P[N(t, t + \delta) = 1/N(s, t) = n] - P[N(t, t + \delta) > 1/N(s, t) = n] \\ &\leq 1 - P[N(t, t + \delta) = 1/N(s, t) = n] \leq 1 - \delta(\lambda(t) - \varepsilon). \end{aligned}$$

Resumiendo,

$$1 - (1 + \varepsilon)\delta[\varepsilon + \lambda(t)] \leq P[N(t, t + \delta) = 0/N(s, t) = n] \leq 1 - \delta(\lambda(t) + \varepsilon). \quad (1.15)$$

Estamos ya en condiciones de acotar  $P[N(s, t + \delta) = n]$ , usando las acotaciones que acabamos de obtener para cada uno de los términos de la expresión (1.11). Resulta por tanto:

$$\begin{aligned} & \{1 - (1 + \varepsilon)\delta[\varepsilon + \lambda(t)]\}P[N(s, t) = n] + \delta(\lambda(t) - \varepsilon)P[N(s, t) = n - 1] \\ &\leq P[N(s, t + \delta) = n] \leq \\ &(1 - \lambda(t)\delta + \varepsilon\delta)P[N(s, t) = n] + \delta(\varepsilon + \lambda(t))P[N(s, t) = n - 1] + \varepsilon\lambda(t)\delta + \varepsilon^2\delta. \end{aligned}$$

Operando y despejando, llegamos a la siguiente acotación:

$$\begin{aligned} -\varepsilon(2 + \varepsilon + \lambda(t)) &\leq \frac{1}{\delta}\{P[N(t, t + \delta) = n] - P[s, t]\} \quad (1.16) \\ &+ \lambda(t)P[N(s, t) = n] - \lambda(t)P[N(s, t) = n - 1] \\ &\leq \varepsilon(2 + \lambda(t) + \varepsilon). \end{aligned}$$

Por fin estamos en condiciones de hallar  $\frac{d P[N(s, t) = n]}{d t}$ . Dado que  $\lambda(t)$  es finita y  $\varepsilon$  es arbitrario, por (1.16), tenemos:

$$\begin{aligned} \frac{d P[N(s, t) = n]}{d t} &= \lim_{\delta \rightarrow 0} \frac{P[N(t, t + \delta) = n] - P[N(s, t) = n]}{\delta} \\ &= -\lambda(t)P[N(s, t) = n] + \lambda(t)P[N(s, t) = n - 1]; \quad n \geq 1 \text{ y } t_0 \leq s \leq t. \end{aligned}$$



De forma análoga, podemos calcular:

$$\frac{d P[N(s, t) = 0]}{d t} = -\lambda(t)P[N(s, t) = 0]; \quad t_0 \leq s \leq t. \quad (1.17)$$

De las expresiones (1.12), (1.14) y (1.15) es claro que las condiciones iniciales de las ecuaciones anteriores son:

$$\lim_{t \searrow s} P[N(s, t) = n] = \delta_{0,n}$$

donde  $\delta_{0,n}$  son las funciones delta de Kronecker

$$\delta_{0,n} = \begin{cases} 1 & ; \quad n = 0 \\ 0 & ; \quad n \neq 0. \end{cases}$$

Las ecuaciones diferenciales a las que hemos llegado se pueden resolver de forma secuencial como vemos a continuación.

Definamos para un  $s$  fijo

$$f_n(t) \equiv P[N(s, t) = n] e^{\int_s^t \lambda(\sigma) d\sigma}, \quad n = 0, 1, \dots$$

Fácilmente observamos que estas funciones cumplen

$$\begin{aligned} \frac{df_0(t)}{dt} &= 0, & f_0(s) &= 1 \\ \frac{df_n(t)}{dt} &= \lambda(t)f_{n-1}(t) & \text{y} & \quad f_n(s) = 0, \quad n \geq 1. \end{aligned}$$

También se cumple que

$$f_0(t) = 1, \quad t \geq s \quad \text{y} \quad f_n(t) = \int_s^t \lambda(\sigma) f_{n-1}(\sigma) d\sigma.$$

Hallemos por inducción otra expresión para  $f_n(t)$ :

$$f_1 = \int_s^t \lambda(\sigma) f_0(t) d\sigma = \int_s^t \lambda(\sigma) d\sigma = \Lambda(t) - \Lambda(s)$$

$$f_2 = \int_s^t \lambda(\sigma) f_1(t) d\sigma = \int_s^t \lambda(\sigma) (\Lambda(\sigma) - \Lambda(s)) d\sigma = \int_s^t (\Lambda(\sigma) - \Lambda(s)) d\Lambda(\sigma)$$

$$= \frac{1}{2} (\Lambda(t) - \Lambda(s))^2$$

$$f_3 = \int_s^t \lambda(\sigma) f_2(t) d\sigma = \int_s^t \frac{1}{2} (\Lambda(\sigma) - \Lambda(s))^2 d\Lambda(\sigma) = \frac{1}{3!} (\Lambda(t) - \Lambda(s))^3$$

supongamos que con  $n - 1$  se cumple

$$f_{n-1} = \frac{1}{(n-1)!} (\Lambda(s) - \Lambda(t))^{n-1}$$

entonces para  $n$ , claramente tenemos que

$$f_n(t) = \frac{1}{n} (\Lambda(t) - \Lambda(s))^n.$$

Por tanto y sin más que despejar hemos obtenido, como pretendíamos, que  $N(s, t)$  tiene distribución de Poisson,

$$P[N(s, t) = n] = \frac{1}{n!} \left( \int_s^t \lambda(\sigma) d\sigma \right)^n e^{-\int_s^t \lambda(\sigma) d\sigma}, \quad n = 0, 1, \dots \quad (1.18)$$

Lo único que falta para demostrar el teorema que nos ocupa, es deducir que el proceso tiene incrementos independientes, para esto se observa que el mismo argumento para hallar (1.18) sirve para  $P[N(s, t) = n/P]$ , con  $P$  un suceso del pasado, resultando igualmente que

$$P[N(s, t) = n/P] = \frac{1}{n!} \left( \int_s^t \lambda(\sigma) d\sigma \right)^n e^{-\int_s^t \lambda(\sigma) d\sigma}, \quad n = 0, 1, \dots,$$

por tanto  $P[N(s, t) = n/P] = P[N(s, t) = n]$ , es decir, el proceso tiene incrementos independientes y con esto concluye la demostración del teorema. ■

### 1.3 Estadísticos de tiempo

Hasta ahora hemos trabajado con distintos estadísticos del proceso de Poisson, pero todos ellos eran estadísticos de recuento, es decir, acerca del número de puntos. Nos ocuparemos ahora de otros estadísticos interesantes del proceso de Poisson, éstos son los estadísticos de *localización y espacios entre puntos*. Restringiéndonos por el momento al proceso de Poisson en una dimensión, seguiremos refiriéndonos al espacio como el tiempo y por tanto los estadísticos que vamos a estudiar serán los *tiempos de ocurrencia* y *tiempos entre ocurrencias*, respectivamente. Estos estadísticos son especialmente interesantes, por ejemplo en teoría de colas, siendo de interés cuándo ocurren los sucesos (llegadas de clientes, fallos en un sistema, etc.) o detección de señales ( Bar-David, 1969). Incluso en numerosa literatura se define al proceso de Poisson desde el punto de vista de los instantes de ocurrencia de los sucesos.

Notaremos por  $\{w_n\}_{n=1,2,\dots}$  a los tiempos de ocurrencia de los puntos y por  $\{t_n\}_{n=1,2,\dots}$  a los tiempos transcurridos entre dos sucesos consecutivos, de forma que  $t_n$  es el tiempo transcurrido entre la ocurrencia del punto  $w_{n-1}$  y el  $w_n$ . Usando esta notación, claramente

$$w_n = t_0 + t_1 + \dots + t_n, \quad \text{con } n = 1, 2, \dots$$

o bien

$$\begin{cases} t_1 = w_1 - t_0 \\ t_n = w_n - w_{n-1}; \quad n \geq 2. \end{cases}$$

### 1.3.1 Distribución de los tiempos de ocurrencias

Sea  $\{N(t); t \geq t_0\}$  un proceso de Poisson con función paramétrica  $\Lambda(t)$  absolutamente continua, es decir, que puede expresarse como

$$\Lambda(t) = \int_{t_0}^t \lambda(\sigma) d\sigma; \quad t \geq t_0,$$

por lo que  $\lambda(t)$  es la función de intensidad del proceso.

Así como teníamos la densidad conjunta para el número de puntos en intervalos de tiempo disjuntos, vamos a calcular la *densidad conjunta de los  $n$  primeros tiempos de ocurrencia*  $w = (w_1, \dots, w_n)$ ; a esta densidad conjunta la notaremos como  $p_w^{(n)}(W)$ . Para ello consideraremos la partición del tiempo en intervalos disjuntos  $t_0, W_1, W_1 + \Delta W_1, W_2, W_2 + \Delta W_2, \dots, W_n + \Delta W_n$ . Claramente,  $\{w_i \in [W_i, W_i + \Delta W_i), i = 1, 2, \dots, n\}$  que se puede expresar igualmente como

$$\{N(t_0, W_1) = 0, N(W_1, W_1 + \Delta W_1) = 1, N(W_1 + \Delta W_1, W_2) = 0, \dots, N(W_n, W_n + \Delta W_n) = 1\}.$$

Debido a la independencia de los incrementos de  $N(t)$ , tenemos

$$\begin{aligned} P\{w_i \in [W_i, W_i + \Delta W_i), i = 1, 2, \dots, n\} \\ &= e^{-[\Lambda(W_1) - \Lambda(t_0)]} \prod_{i=1}^n [\Lambda(W_i + \Delta W_i) - \Lambda(W_i)] e^{-[\Lambda(W_i + \Delta W_i) - \Lambda(W_i)]} \\ &= \prod_{i=2}^n e^{-[\Lambda(W_i) - \Lambda(W_{i-1})]} = \left[ \prod_{i=1}^n \int_{W_i}^{W_i + \Delta W_i} \lambda(\sigma) d\sigma \right] e^{-\int_{t_0}^{W_n + \Delta W_n} \lambda(\sigma) d\sigma}. \end{aligned}$$

Sustituyendo este resultado en la definición de función de densidad conjunta de los  $n$  primeros tiempos de ocurrencia, obtenemos por fin

su expresión:

$$\begin{aligned}
 p_w^{(n)}(W) &\equiv \lim_{\max \Delta W_i \rightarrow 0} \frac{P\{w_i \in [W_i, W_i + \Delta W_i], i = 1, 2, \dots, n\}}{\prod_{i=1}^n \Delta W_i} \\
 &= \begin{cases} [\prod_{i=1}^n \lambda(W_i)] \exp\left(-\int_{t_0}^{W_n} \lambda(\sigma) d\sigma\right); & t_0 \leq W_1 \leq \dots \leq W_n \\ 0, & \text{en otro caso.} \end{cases} \quad (1.19)
 \end{aligned}$$

Hemos calculado esta función de densidad para un  $\lambda(t)$  general. Si el proceso de Poisson es homogéneo con intensidad constante,  $\lambda$ , de (1.19) obtenemos

$$p_w^{(n)}(W) = \begin{cases} \lambda^n e^{-\lambda(W_n - t_0)}; & t_0 \leq W_1 \leq \dots \leq W_n \\ 0, & \text{en otro caso.} \end{cases}$$

Estudiemos algunas propiedades de los tiempos de ocurrencia de los puntos del proceso de Poisson.

**Propiedad 1.3.1** *La secuencia de los tiempos de ocurrencia es una cadena de Markov con probabilidad de transición*

$$p_{w_n/w_{n-1}}(W_n/W_{n-1}) = \lambda(W_n) e^{-[\Lambda(W_n) - \Lambda(W_{n-1})]} \quad (1.20)$$

donde  $\Lambda(t) = \int_{t_0}^t \lambda(\sigma) d\sigma$ .

**Demostración.** La demostración es inmediata. Calculemos la probabilidad condicionada de  $w_n$  dados  $w_i = W_i, i = 1, \dots, n-1$

$$p_{w_n/w_{n-1}, \dots, w_1}(W_n/W_{n-1}, \dots, W_1) = \frac{p_w^{(n)}(W)}{p_w^{(n-1)}(W)} = \lambda(W_n) e^{\int_{W_{n-1}}^{W_n} \lambda(\sigma) d\sigma}.$$

Observamos que la función de densidad no depende de  $W_i, i = 1, \dots, n-1$ , por tanto, trivialmente

$$p_{w_n/w_{n-1}, \dots, w_1}(W_n/W_{n-1}, \dots, W_1) = p_{w_n/w_{n-1}}(W_n/W_{n-1})$$

lo que significa que efectivamente la secuencia de  $w_1, \dots, w_n$  es una cadena de Markov con la función de densidad dada en (1.20). ■

**Propiedad 1.3.2** *La función de densidad de los tiempos de ocurrencia adelantada (función de densidad condicionada del  $n$ -ésimo tiempo entre ocurrencias,  $t_n$ , dados los  $n-1$  tiempos de ocurrencia anteriores,  $w_1, \dots, w_{n-1}$  para  $n = 2, 3, \dots$ ) es*

$$p_{t_n/w_{n-1}, \dots, w_1}(T/W_{n-1}, \dots, W_1) = \lambda(W_{n-1} + T)e^{-[\Lambda(W_{n-1}+T) - \Lambda(W_{n-1})]} \quad (1.21)$$

con  $n = 2, 3, \dots$ , y  $T \geq 0$ . Además

$$p_{t_n/w_{n-1}, \dots, w_1}(T/W_{n-1}, \dots, W_1) = p_{t_n/w_{n-1}}(T/W_{n-1}),$$

con  $n = 2, 3, \dots$ ,  $T \geq 0$ . Definiendo  $w_0 = W_0 = t_0$ , también se cumplen estas expresiones para  $n = 1$ .

**Demostración.** Ya que  $w_n = w_{n-1} + t_n$  y puesto que la secuencia de tiempos de ocurrencia es una cadena de Markov,

$$\begin{aligned} p_{t_n/w_{n-1}, \dots, w_1}(T/W_{n-1}, \dots, W_1) &= p_{w_n/w_{n-1}, \dots, w_1}(W_n + T/W_{n-1}, \dots, W_1) \\ &= p_{w_n/w_{n-1}}(W_{n-1} + T), \quad n = 2, 3, \dots, T \geq 0. \end{aligned}$$

Usando la función de densidad de transición de  $w_{n-1}$  a  $w_n$  dada en (1.20), la función de densidad adelantada para un proceso de Poisson es

$$p_{t_n/w_{n-1}, \dots, w_1}(T/W_{n-1}, \dots, W_1) = \lambda(W_{n-1} + T)e^{-[\Lambda(W_{n-1}+T) - \Lambda(W_{n-1})]}$$

con  $n = 2, 3, \dots$ , y  $T \geq 0$ . Además,

$$p_{t_n/w_{n-1}, \dots, w_1}(T/W_{n-1}, \dots, W_1) = p_{t_n/w_{n-1}}(T/W_{n-1}),$$

con  $n = 2, 3, \dots$ ,  $T \geq 0$ . Definiendo  $w_0 = W_0 = t_0$ , también se cumplen estas expresiones para  $n = 1$ . ■

**Observación.** En el caso de un proceso de Poisson homogéneo con  $\lambda(t) = \lambda$ , la expresión anterior resulta

$$p_{t/w_{n-1}, \dots, w_1}(T/W_{n-1}, \dots, W_1) = \lambda e^{-\lambda T}.$$

Un ejemplo en el que es útil esta propiedad es el caso en el que los puntos son fallos de un sistema. Si éstos siguen un proceso de Poisson y se quiere estudiar la probabilidad de que el sistema continúe operativo durante un periodo de tiempo después de los fallos conocidos, según la propiedad que acabamos de ver, sólo es necesario fijarse en el último de los fallos, además de darnos la función de densidad deseada.

**Propiedad 1.3.3** *La función de densidad condicionada de los  $n$  primeros tiempos de ocurrencia dado  $N(t_0, W_n) = n$  es*

$$p_w(W/N(t_0, W_n) = n) = n! \left( \int_{t_0}^{W_n} \lambda(t) dt \right)^{-n} \left( \prod_{i=1}^n \lambda(W_i) \right) \quad (1.22)$$

**Demostración.** La función de densidad de los  $n$  primeros tiempos de ocurrencia,  $p_w^{(n)}(W)$ , en el intervalo  $t_0 \leq \sigma < W_n$ , tiene implícita la variable aleatoria  $N(W_n)$ . Es obvio que la función de densidad conjunta de los  $n$  primeros tiempos de ocurrencia y la variable  $N(W_n)$ ,  $p_w(W, N(W_n) = n)$ , es  $p_w^{(n)}(W)$ . También es evidente que

$$P[N(W_n) = n/w_1 = W_1, \dots, w_n = W_n] = 1$$

y como conocemos por (1.1)  $P[N(s, t) = n]$ , podemos hallar la función de densidad condicionada de los primeros tiempos de ocurrencia dado que conocemos  $N(t_0, W_n) = n$ . Esta no es más que el cociente entre (1.19) y (1.1). Simplificando, concluimos que

$$p_w(W/N(t_0, W_n) = n) = n! \left( \int_{t_0}^{W_n} \lambda(t) dt \right)^{-n} \left( \prod_{i=1}^n \lambda(W_i) \right),$$

como queríamos demostrar. ■

**Observación.** En el caso particular de un proceso de Poisson homogéneo con  $\lambda(t) = \lambda$ , tenemos que

$$p_w(W/N(t_0, W_n) = n) = n!(W_n - t_0)^{-n}.$$

El desarrollo para la obtención de la  $p_w(W/N(t_0, t) = n)$ , hemos podido realizarlo de forma tan sencilla gracias a los axiomas que escogimos para la definición de proceso de Poisson. Otra obtención de esta misma expresión la encontramos en el artículo de Bar-David (1969), debido a la definición del proceso de la que parte.

### 1.3.2 Distribución de los tiempos entre ocurrencias

Tanto en los estadísticos de recuento como en la distribución de los tiempos entre ocurrencias, el caso de proceso de Poisson homogéneo se obtenía como simple caso particular del no homogéneo, y las propiedades de los tiempos de ocurrencias eran análogas sin más que sustituir la función intensidad  $\lambda(t)$  por la constante  $\lambda$ . Sin embargo, con respecto a los tiempos entre ocurrencias obtendremos distintas distribuciones y propiedades para el caso homogéneo y no homogéneo.

Calculemos la probabilidad de que el  $n$ -ésimo tiempo entre ocurrencias,  $t_n$ , sea mayor que  $T$ , dado que  $t_i = T_i, i = 1, \dots, n-1$ . Debido a la relación entre  $w_i$  y  $t_i$  y por la propiedad de Markov de los tiempos de ocurrencias, se calcula

$$\begin{aligned} & P[t_n > T/t_1 = T_1, t_2 = T_2, \dots, t_{n-1} = T_{n-1}] = \\ & P[w_n > t_0 + T_1 + \dots + T_{n-1} + T/w_1 = t_0 + T_1, \dots, w_{n-1} = t_0 + T_1 + \dots + T_{n-1}] \\ & = P[w_n > t_0 + T_1 + \dots + T_{n-1} + T/w_{n-1} = t_0 + T_1 + \dots + T_{n-1}]. \end{aligned}$$

Diferenciando con respecto a  $T$  y usando (1.21) llegamos a la expresión



para la función de densidad de  $t_n$  condicionada a  $t_1, t_2, \dots, t_{n-1}$

$$\begin{aligned} p_{t_n/t_{n-1}, \dots, t_1}(T/T_{n-1}, \dots, T_1) & \quad (1.23) \\ & = \lambda(t_0 + T_1 + T_2 + \dots + T_{n-1} + T)e^{-[\Lambda(t_0+T_1+\dots+T_{n-1}+T)-\Lambda(t_0+T_1+\dots+T_{n-1})]} \end{aligned}$$

Claramente, esta expresión para un proceso de Poisson no homogéneo, al depender de  $T_1, \dots, T_{n-1}$  nos indica que los tiempos entre ocurrencias no son independientes, lo que derrumba el mito de que un proceso de Poisson tiene tiempos entre ocurrencias independientes e idénticamente distribuidos. Sí es cierto en el caso particular de un proceso de Poisson homogéneo, como veremos en el siguiente teorema.

**Teorema 1.3.4** *Un proceso de Poisson homogéneo con intensidad  $\lambda$  tiene tiempos entre ocurrencias independientes e idénticamente distribuidos, con distribución exponencial de parámetro  $\lambda$ .*

Este teorema es a menudo en la literatura, definición del proceso de Poisson homogéneo, en aquellos campos en los que tienen más interés los tiempos de y entre ocurrencias, que el número de ocurrencias.

**Demostración.** Es obvio que

$$p_t^{(n)} = p_{t_n/t_{n-1}, \dots, t_1}(T_n/T_{n-1}, \dots, T_1) p_{t_{n-1}/t_{n-2}, \dots, t_1}(T_{n-1}/T_{n-2}, \dots, T_1) \cdots p_{t_1}(T_1)$$

y por (1.23) obtenemos

$$\begin{aligned} p_t^{(n)}(T) & = \\ & \lambda(t_0 + T_1 + \dots + T_{n-1} + T_n)e^{-[\Lambda(t_0+T_1+\dots+T_{n-1}+T_n)-\Lambda(t_0+T_1+\dots+T_{n-1})]} \\ & \quad \cdots \lambda(t_0 + T_1)e^{-[\Lambda(t_0+T_1)-\Lambda(t_0)]} = \lambda^n e^{-\lambda \sum_{i=1}^n T_i}, \\ & \quad T_i \geq 0, i = 1, \dots, n. \end{aligned}$$

con lo que el teorema queda demostrado. ■

Una conclusión inmediata del teorema anterior es la *función de densidad del  $n$ -ésimo tiempo de ocurrencia*,  $w_n = t_1 + \dots + t_n$ , para un proceso de Poisson homogéneo con intensidad  $\lambda$ . Esta es

$$p_{w_n}(W) = \begin{cases} \frac{(\lambda W)^n}{(n-1)!} \lambda e^{-\lambda W}; & W \geq 0 \\ 0; & \text{en otro caso.} \end{cases}$$

Así que la media de  $w_n$  es  $n/\lambda$  y su varianza,  $n/\lambda^2$ . Según lo visto,  $w_n$  es una suma de variables aleatorias independientes e idénticamente distribuidas con varianza finita si  $\lambda \neq 0$ . En este caso, por el teorema central del límite la distribución de  $w_n$ , con  $n$  suficientemente grande, se puede aproximar por una Normal de media  $n/\lambda$  y varianza  $n/\lambda^2$ .

El recíproco del anterior teorema también se verifica:

Si  $\{t_n\}$  es una sucesión de variables aleatorias independientes e idénticamente distribuidas con distribución exponencial de parámetro  $\lambda$ , llamando a  $t_n$  tiempo entre ocurrencias, el proceso de esos sucesos es un proceso de Poisson homogéneo con intensidad  $\lambda$ . Puede verse la demostración en Gross-Harris (1974).

### 1.3.3 Función de densidad muestral y de verosimilitud

La función de densidad muestral y su logaritmo son cruciales especialmente a la hora de la estimación del parámetro de un proceso de Poisson. Aunque nuestro objetivo final es el estudio y estimación de un proceso de Poisson doblemente estocástico y no nos vamos a detener en la estimación de un proceso de Poisson no homogéneo, hallaremos la expresión de la función de densidad muestral pues nos permitirá dar una expresión de la distribución condicionada de los tiempos de ocurrencia y del funcional característico.

Sea  $\Omega = \{W_1, \dots, W_n, N(t) = n\}$  una realización de  $w$  en el proceso de Poisson  $\{N(\sigma) : t_0 \leq t\}$  en el intervalo  $[t_0, t)$ . La *función de densidad muestral*,  $p(\Omega)$ , se define como

$$p(\Omega) \equiv \begin{cases} P[N(t) = 0], & N(t) = 0 \\ p_w[W, N(t) = n], & N(t) = n \geq 1 \end{cases}$$

donde

$$p_w[W, N(t) = n] = P[N(t) = n/w_1 = W_1, \dots, w_n = W_n] p_w^{(n)}(W).$$

Observando la definición de la función de densidad muestral podemos decir que define la probabilidad de obtener exactamente  $n$  puntos en el intervalo  $[t_0, t)$  y justo en los instantes  $w_1 = W_1, \dots, w_n = W_n$ . Nos queda dar una expresión para  $p(\Omega)$ .

**Teorema 1.3.5** *Función de densidad muestral*

Sea  $\{N(t) : t_0 \leq \sigma < t\}$  un proceso de Poisson no homogéneo con intensidad  $\lambda(t); t \geq t_0$ . Entonces la función de densidad muestral para  $N$  es

$$p(\Omega) = \begin{cases} e^{-\int_{t_0}^t \lambda(\sigma) d\sigma}, & N(t) = 0 \\ [\prod_{i=1}^n \lambda(W_i)] e^{-\int_{t_0}^t \lambda(\sigma) d\sigma}, & N(t) = n \geq 1. \end{cases} \quad (1.24)$$

**Demostración.** Con respecto a la primera parte, por (1.1),

$$P[N(t) = 0] = e^{-\int_{t_0}^t \lambda(\sigma) d\sigma}.$$

Acerca de la segunda parte podemos decir que como

$$\begin{aligned} P[N(t) = n/w_1 = W_1, \dots, w_n = W_n] &= \\ &= P[N(t) - N(W_n^+)] = e^{-\int_{W_n}^t \lambda(\sigma) d\sigma} \end{aligned}$$

y usando (1.19), el teorema se verifica. ■

Definamos la *integral de recuento* como sigue

$$\int_{t_0}^t \ln[\lambda(\sigma)] N(d\sigma) = \begin{cases} 0, & N(t) = 0 \\ \sum_{i=1}^{N(t)} \ln[\lambda(w_i)], & N(t) \geq 1. \end{cases}$$

La función de densidad muestral vista en (1.24) puede expresarse usando la integral de recuento de la siguiente forma:

$$p(\Omega) = \exp \left[ - \int_{t_0}^t \lambda(\sigma) d\sigma + \int_{t_0}^t \ln[\lambda(\sigma)] N(d\sigma) \right]. \quad (1.25)$$

Con esta forma de escribir la función de densidad queda claro que  $p(\Omega)$  es función del proceso aleatorio  $\{N(\sigma) : t_0 \leq \sigma < t\}$  en el intervalo  $[t_0, t)$ .

Dada una realización arbitraria,  $w$ , del proceso, el logaritmo de  $p(w)$  es el *logaritmo de la función de verosimilitud muestral* y viene dada por

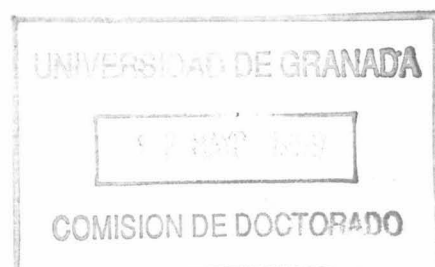
$$\mathcal{L} \equiv \ln[p(w)] = - \int_{t_0}^t \lambda(\sigma) d\sigma + \int_{t_0}^t \ln[\lambda(\sigma)] N(d\sigma) \quad (1.26)$$

Como adelantamos anteriormente, veamos algunas aplicaciones de la función de densidad muestral.

### Distribución condicional de los tiempos de ocurrencia

Para un intervalo  $[t_0, T]$ , usando el resultado de la función de densidad muestral (1.24), tenemos que:

$$p_w(W/N(T) = n) = \frac{p_w(W, N(T) = n)}{P[N(T) = n]} = n! \prod_{i=1}^n \frac{\lambda(W_i)}{\int_{t_0}^T \lambda(\sigma) d\sigma}. \quad (1.27)$$



Sean  $u_1, \dots, u_n$  variables aleatorias independientes e idénticamente distribuidas con densidad de probabilidad común

$$p_{u_i}(U_i) = \begin{cases} \frac{\lambda(U_i)}{\int_{t_0}^T \lambda(\sigma) d\sigma}, & t_0 \leq U_i \leq T \\ 0, & \text{fuera.} \end{cases}$$

Claramente, debido a la independencia de las variables, la probabilidad conjunta para  $(u_1, \dots, u_n) = u$  resulta

$$p_u(U) = \prod_{i=1}^n \frac{\lambda(U_i)}{\int_{t_0}^T \lambda(\sigma) d\sigma}, \quad t_0 \leq U_i \leq T; \quad i = 1, \dots, n.$$

Sean  $o_1, \dots, o_n$  las variables aleatorias ordenadas asociadas a  $u_1, \dots, u_n$ . Hay  $n!$  permutaciones y como son sucesos mutuamente excluyentes, sus probabilidades se suman y la función de densidad conjunta para  $(o_1, \dots, o_n) = o$  es la siguiente

$$p_o(O) = n! \prod_{i=1}^n \frac{\lambda(O_i)}{\int_{t_0}^T \lambda(\sigma) d\sigma}.$$

Podemos observar que es la misma que la función de densidad condicionada de los tiempos de ocurrencia  $w_1, \dots, w_n$  de un proceso de Poisson con la condición de que  $N(T) = n$ , que hallamos en (1.27).

En el caso de que  $\lambda(U_i) = \lambda$ , los  $o_i$  son variables aleatorias independientes e idénticamente distribuidas con distribución uniforme en  $[t_0, T]$  y resulta que su función de densidad conjunta es igual a la de los tiempos de ocurrencia de un proceso de Poisson homogéneo con la condición de que  $N(t_0, T) = n$ . Este resultado se usa para verificar si un proceso puntual empírico es homogéneo o no.

### Funcional característico de un proceso de Poisson

Como decíamos en la página 8, cuando tratábamos de dar una expresión para el funcional característico de un proceso de Poisson (1.6), usando la propiedad de la distribución de la distribución condicional de los tiempos de ocurrencia visto anteriormente, se demuestra que el funcional característico de  $N(T)$  es

$$\Phi_N(iv) = \exp \left[ \int_{t_0}^T \lambda(\sigma) (e^{iv(\sigma)} - 1) d\sigma \right] \quad (1.28)$$

La función característica es el caso particular del funcional con

$$v(\sigma) = \begin{cases} 0, & t_0 \leq \sigma < s \\ \alpha, & s \leq \sigma < t \\ 0, & t \leq \sigma < T, \end{cases}$$

con lo que  $\int_{t_0}^T v(\sigma) N(d\sigma) = \alpha N(s, t)$ . Sustituyendo  $v(\sigma)$  en el funcional, la función característica de  $N(s, t) = N(t) - N(s)$  resulta ser:

$$\Phi_N(iv) = \exp \left[ (e^{i\alpha} - 1) \int_s^t \lambda(\sigma) d\sigma \right] = M_{s,t}(i\alpha).$$

## 1.4 Proceso de Poisson según la Teoría de Martingalas

La teoría de martingalas ha sido desarrollada intensivamente en las últimas dos décadas, tanto por ingenieros como por matemáticos y estadísticos. La razón por la cual ha tenido tanto desarrollo en los últimos años ha sido la necesidad de encontrar una teoría de estimación y predicción para procesos puntuales más potente que la clásica. Para procesos del tipo de una difusión, la teoría de martingalas es una

buena herramienta para los problemas mencionados, además de dar información de la estructura del proceso y de su función de verosimilitud, así como para resultados de convergencia. El inconveniente de los métodos de esta teoría es su restricción a procesos de tipo evolutivo en el tiempo, aunque los casos de este tipo son muy abundantes en la práctica. La idea de que la teoría de martingalas podía ser una herramienta eficaz para los procesos puntuales la dió Watanabe (1964) que fué quien dió una caracterización de martingala para el proceso de Poisson y más tarde comenzó a desarrollar la teoría de martingalas junto a Kunita (1967). Brémaud ha realizado importantes estudios sistemáticos de esta teoría, apareciendo siempre como punto de referencia para libros y artículos acerca de la misma.

Nosotros centraremos el interés en la aplicación de las martingalas al estudio del proceso de Poisson que es el que nos ocupa. En primer lugar definiremos y estudiaremos conceptos que serán el marco en el que nos movamos en lo sucesivo.

### 1.4.1 Conceptos básicos

Haremos en esta subsección, un inciso en nuestro estudio del proceso de Poisson, para ver algunos conceptos básicos de la teoría de martingalas y que así, cuando veamos nuestro proceso desde esta teoría sea mucho más directo y se tenga establecida una notación conocida. Para un estudio más detallado consultar por ejemplo ( Daley y Vere-Jones, 1988) o ( Andersen et al. (1993)).

Notaremos por  $\{X_t(\omega)\} = \{X(t, \omega)\} = \{X(t)\}$  a un proceso estocástico real-valuado en  $t \in (0, \infty) \equiv \mathbb{R}_+$ . Se puede interpretar este proceso como una familia indizada de variables aleatorias con un espacio de probabilidad común  $(\Omega, \mathcal{E}, \mathcal{P})$  con índices en  $\mathbb{R}_+$ , pero será más apropiado ahora interpretarlo como una función en el espacio  $\mathbb{R}_+ \times \Omega$ .

## 1.4 Proceso de Poisson según la Teoría de Martingalas 31

El proceso  $X : \mathbb{R}_+ \times \Omega \longrightarrow \mathbb{R}$  es medible cuando, para todo  $A \in \mathcal{B}(\mathbb{R})$ ,

$$\{(t, \omega : X(t, \omega) \in A) \in \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{E},$$

donde  $\mathcal{B}(\mathbb{R}_+) \otimes \mathcal{E}$  denota el producto de las dos  $\sigma$ -álgebras de  $\mathcal{B}(\mathbb{R}_+)$  y  $\mathcal{E}$ , respectivamente. Visto el proceso como familia indizada de variables aleatorias, es medible si y sólo si sus trayectorias son continuas c.s. o monótonas c.s. y continuas a la derecha.

Una familia de sub- $\sigma$ -álgebras de  $\mathcal{E}$ ,  $\mathcal{F} \equiv \{\mathcal{F}_t : 0 \leq t \leq \infty\}$ , se denomina *filtración* o *historia* del proceso y contiene información del mismo. Si además  $X(t, \omega)$  es  $\mathcal{F}_t$ -medible para todo  $t$ , se dice que  $X$  es  $\mathcal{F}$ -adaptado.

Consideremos la  $\sigma$ -álgebra generada por todas las posibles evoluciones del proceso en el intervalo  $(0, t]$  es decir,  $\{X(s, \omega) : 0 < s \leq t\}$ ; la notaremos por

$$\mathcal{H}_t \equiv \{\{\omega : X(s, \omega) \leq s\} : 0 < s \leq t\}.$$

La  $\sigma$ -álgebra  $\mathcal{H}_0$ , será la generada por la información a priori que tengamos del proceso  $X$ . Claramente,  $\mathcal{H}_s \subseteq \mathcal{H}_t$  para  $0 < s < t < \infty$ . Por tanto, tenemos que  $\mathcal{H} \equiv \{\mathcal{H}_t : 0 \leq t \leq \infty\}$  es una familia expansiva de sub- $\sigma$ -álgebras de  $\mathcal{E}$  por lo que es una filtración o historia del proceso. Como  $X(t, \omega)$  es  $\mathcal{H}_t$ -medible para todo  $t$ ,  $X$  es  $\mathcal{H}$ -adaptado.

Llamaremos *historia natural*, *mínima* o *interna* de  $X$  a la historia  $\mathcal{H}$  tal que  $\mathcal{H}_0 = \liminf_{t>0} \mathcal{H}_t = \{\emptyset, \Omega\}$  y  $\mathcal{H}_\infty = \bigcap_{t>0} \mathcal{H}_t$ . Los dos primeros nombres de esta historia reflejan el hecho que es la mínima familia anidada de  $\sigma$ -álgebras para la que  $X$  es adaptado.

Sea  $X$  un proceso  $\mathcal{F}$ -adaptado; se dice *progresivamente medible* con respecto a  $\mathcal{F}$ , si para todo  $t \in \mathbb{R}_+$  y todo  $A \in \mathcal{B}(\mathbb{R})$ , se verifica que

$$\{(s, \omega) : 0 < s \leq t, X(s, \omega) \in A\} \in \mathcal{B}((0, t]) \otimes \mathcal{F}_t$$

Los conjuntos de la forma  $[s, t] \times U; 0 \leq s < t, U \in \mathcal{F}_t, t > 0$ , generan un  $\sigma$ -álgebra en  $\mathbb{R}_+ \times \Omega$ , que llamaremos  $\sigma$ -álgebra  $\mathcal{F}$ -progresiva.



Así que un proceso  $X$  progresivamente medible con respecto a  $\mathcal{F}$  es equivalente a que  $X(t, \omega)$  sea medible con respecto a la  $\sigma$ -álgebra  $\mathcal{F}$ -progresiva.

Definamos una nueva  $\sigma$ -álgebra necesaria para lo siguiente. Notaremos  $\Psi^{\mathcal{F}}$  a la sub- $\sigma$ -álgebra de  $\mathcal{B}(\mathbb{R}_+) \otimes \mathcal{E}$  generada por conjuntos de la forma  $[s, t] \times U$ , donde  $U \in \mathcal{F}_s$ ,  $t > s$  y  $0 \leq s < \infty$  y la denominaremos  $\sigma$ -álgebra *predecible*. Su nombre indica el hecho intuitivo de que puede ser predecido un tiempo "futuro"  $t$  dada la evolución del proceso hasta el instante "presente"  $s$ .

Se llama a  $X$ ,  $\mathcal{F}$ -predecible, cuando es  $\Psi^{\mathcal{F}}$ -medible; es decir, si para cualquier  $A \in \mathcal{B}(\mathbb{R})$ ,

$$\{(t, \omega) : X(s, \omega) \in A\} \in \Psi^{\mathcal{F}}$$

El típico ejemplo de proceso  $\mathcal{F}$ -predecible es cualquier proceso  $\mathcal{F}$ -adaptado que sea continuo a la izquierda.

La *historia continua a la izquierda*, que notaremos por  $\mathcal{F}_{(-)} \equiv \{\mathcal{F}_{t-}\}$ , asociada a la historia  $\mathcal{F}$  es aquella tal que

$$\mathcal{F}_{0-} = \mathcal{F}_0 \text{ y } \mathcal{F}_{t-} = \limsup_{s < t} \mathcal{F}_s = \bigvee_{s < t} \mathcal{F}_s$$

**Lema 1.4.1** *Un proceso  $\mathcal{F}$ -predecible es  $\mathcal{F}_{(-)}$ -adaptado.*

**Demostración.** Consideremos un proceso de la forma

$$X(t, \omega) = I(a, b] I_U(\omega) \text{ con } 0 < a < b < \infty, U \in \mathcal{F}_a$$

que es  $\mathcal{F}$ -predecible por construcción de  $\Psi^{\mathcal{F}}$ .

Para cualquier  $t$ ,  $\{\omega : X(t, \omega) = 1\} = \emptyset$  si  $t \leq a$  ó  $b < t$  y  $\{\omega : X(t, \omega) = 1\} = U$  si  $a < t \leq b$ , por lo que el proceso es  $\mathcal{F}_{(-)}$ -medible,  $\forall t$ .

Puesto que cualquier función  $\mathcal{F}$ -predecible se puede expresar como combinación lineal de funciones de este tipo y como los procesos  $\mathcal{F}_{(-)}$ -adaptados forman una clase cerrada para combinaciones lineales y límites monótonos, el lema queda probado. ■

Dada una historia  $\mathcal{F}$ , se llama *tiempo de parada respecto de  $\mathcal{F}$*  a la variable aleatoria no negativa  $T : \Omega \rightarrow [0, \infty]$  tal que

$$\{w : T(w) \leq t\} \in \mathcal{F}_t; \quad 0 \leq t < \infty$$

Si  $S$  y  $T$  son tiempos de parada respecto de  $\mathcal{F}$ , también lo son  $S \vee T$  y  $S \wedge T$ .

Notando  $\{T(w) = \infty\} = \bigcap_n \{T(w) > n\}$  y como  $\{T(w) = \infty\} \in \mathcal{F}_\infty$ , podemos considerar *tiempos de parada extendidos* a aquellos tal que  $P[T(w) < \infty] < 1$ .

**Lema 1.4.2** *Sea  $X(t)$  un proceso estocástico monótono creciente, continuo a la derecha y  $\mathcal{F}$ -adaptado y sea  $Y$  una variable aleatoria  $\mathcal{F}_0$ -medible. Entonces  $T(w) \equiv \inf \{t : X(t, w) \geq Y(w)\}$  es un tiempo de parada respecto a  $\mathcal{F}$ , posiblemente extendido, y si  $X$  es  $\mathcal{F}$ -predecible, entonces  $T$  es un tiempo de parada respecto a  $\mathcal{F}_{(-)}$  (extendido).*

**Demostración.** Si  $Y = cte.$ , puesto que  $X$  es creciente,  $X(t) \geq Y$  si y sólo si  $T \leq t$ , y como  $X$  es  $\mathcal{F}$ -adaptado, tenemos que  $\{w : X(t, w) \geq Y\} \in \mathcal{F}_t$ , por tanto  $\{w : T(w) \leq t\} \in \mathcal{F}_t$ .

En el caso más general,  $X(t, w) - Y(w)$  es monótono creciente, continuo a la derecha y  $\mathcal{F}$ -adaptado (pues como  $Y$  es  $\mathcal{F}_0$ -medible, es  $\mathcal{F}_t$ -medible,  $\forall t > 0$ ). Siguiendo el mismo razonamiento anterior para  $cte. = 0$ , tenemos lo que queremos.

Por fin, si  $X$  es  $\mathcal{F}$ -predecible, es  $\mathcal{F}_{(-)}$ -adaptado, como vimos en el lema 1.4.1, así podemos asegurar que  $T$  es un tiempo de parada respecto a  $\mathcal{F}_{(-)}$ . ■

Si tenemos un proceso estocástico  $X(t)$ , usaremos la siguiente notación

$$X(t \wedge T) = \begin{cases} X(t), & t \leq T \\ X(T), & t > T \end{cases}$$

Se demuestra para este tipo de procesos la proposición que sigue (ver (Daley y Vere-Jones, 1988)).

**Proposición 1.4.3** *Sea  $\mathcal{F}$  una historia,  $T$  un tiempo de parada respecto de  $\mathcal{F}$  y  $X(t)$  un proceso. Entonces  $X(t \wedge T)$  es medible,  $\mathcal{F}$ -progresivo o  $\mathcal{F}$ -predecible, según lo sea  $X(t)$ . En cualquier caso, la variable aleatoria  $X(T)$  es  $\mathcal{F}_\infty$ -medible.*

Esta proposición tiene como interesante corolario el siguiente resultado.

**Corolario 1.4.4** *Si un proceso  $X(t)$  es  $\mathcal{F}$ -progresivo e integrable c.s. en intervalos finitos, entonces*

$$Y(t, \omega) = \int_0^t X(s, \omega) ds$$

*es  $\mathcal{F}$ -progresivo,  $Y(T)$  es una variable aleatoria si  $T < \infty$  e  $Y(t \wedge T)$  es también  $\mathcal{F}$ -progresivo.*

Recordemos la definición de martingala pues será el concepto fundamental con el que vamos a tratar de ahora en adelante.

Sea  $(\Omega, \mathcal{E}, \mathcal{P})$  un espacio probabilístico,  $\mathcal{F}$  una historia en  $(\Omega, \mathcal{E})$  y  $\{X(t) : 0 \leq t < \infty\}$  un proceso real-valuado,  $\mathcal{F}$ -adaptado y tal que  $E[|X(t)|] < \infty, \forall t$ . Entonces  $X$  es una  $\mathcal{F}$ -martingala si para  $0 \leq s < t < \infty$ ,

$$E[X(t)/\mathcal{F}_s] = X(s) \quad \text{c.s.} \quad (1.29)$$

Si en lugar de igualdad en (1.29) se tiene la desigualdad  $\leq$  o  $\geq$ , la definición de super y submartingala, respectivamente.

Finalmente, citaremos la forma del teorema de descomposición de Doob-Meyer que usaremos más adelante. La demostración puede consultarse en (Lipster y Shiriyayev, 1977).

**Teorema 1.4.5 (de descomposición de Doob-Meyer)** *Sea  $\mathcal{F}$  una historia y  $X$  un proceso acotado,  $\mathcal{F}$ -submartingala con trayectorias continuas a la izquierda. Entonces existe una única  $\mathcal{F}$ -martingala, uniformemente integrable,  $Y(\bullet)$ , y un único proceso acumulativo y  $\mathcal{F}$ -predecible,  $A(\bullet)$ , tal que*

$$X(t) = Y(t) + A(t)$$

Para  $A(\bullet)$  no decreciente y de trayectorias continuas a la derecha se demuestra además que el hecho de que sea  $\mathcal{F}$ -predecible es equivalente a que

$$E \int_0^u Z(t) A(dt) = E \int_0^u Z(t-) A(dt) \tag{1.30}$$

para cualquier  $\mathcal{F}$ -martingala  $Z(\bullet)$  acotada y  $u \geq 0$ .

Como para cualquier proceso  $\mathcal{F}$ -adaptado y acumulativo,  $\xi$ , y cualquier  $\mathcal{F}$ -martingala,  $Z$ , se verifica que

$$E \left[ Z(u) \int_0^u \xi(dt) \right] = E \left[ \int_0^u Z(t) \xi(dt) \right],$$

la propiedad anterior es equivalente a

$$E [Z(u) A(u)] = E \int_0^u Z(t-) A(dt) \tag{1.31}$$

La condición de que el proceso sea acotado impide la posibilidad de “explosiones” e impone la restricción a la serie de tiempos de parada (ver lema 1.4.2)  $\{T_n\}$ ,  $n = 0, 1, 2, \dots$  definida por

$$T_n \equiv T_n(w) = \sup \{t : X(t, w) < n\} = \begin{cases} \infty, & X(t, w) < n, \text{ para todo } 0 < t < \infty \\ \inf \{t : X(t, w) \geq n\}, & \text{en otro caso} \end{cases} \quad (1.32)$$

de que  $T_n \rightarrow \infty$  c.s. cuando  $n \rightarrow \infty$ .

### 1.4.2 El proceso de Poisson es una semimartingala

El resultado que perseguimos ahora, dicho de forma intuitiva, es que un proceso de un punto tiene un *compensador* que lo “convierte” en martingala. Luego lo particularizaremos al caso del proceso de Poisson. En realidad, esto se puede demostrar para el caso general de cualquier proceso acumulativo, pero excede nuestras pretensiones.

Sea  $\xi$  un proceso acumulativo en  $\mathbb{R}^+$  y  $\mathcal{F}$  una historia del mismo. Se denomina *compensador* a un proceso monótono, no decreciente, continuo a la derecha y predecible, que notaremos  $A(\bullet)$ , tal que para cada  $n$  y tiempo de parada  $T_n$  definido en (1.32) el proceso  $\{\xi(t - T_n) - A(t - T_n) : 0 \leq t < \infty\}$  sea una martingala con respecto a  $\mathcal{F}$ . A un proceso, que admite un compensador, se le denomina *semimartingala*.

En primer lugar, demostraremos que un proceso puntual de un sólo punto en algún instante de tiempo tiene compensador y luego lo generalizaremos al caso de proceso puntual de varios puntos. Recordamos que considerar a un proceso, puntual o de recuento, no es más que dos visiones de una misma cosa.

**Proposición 1.4.6** *Un proceso  $N$  de un sólo punto, generado por la variable aleatoria (instante de ocurrencia)  $X$ , definido por*

$$N(t, w) = I_{(0,t)}(T(w))$$

tiene un  $\mathcal{H}$ -compensador

$$A(t, w) = \int_0^{t \wedge X(w)} (1 - F(x))^{-1} dFx; \quad 0 < t < \infty,$$

donde  $F$  es la función de distribución de  $X$ .

**Demostración.** Llamemos  $H(t \wedge X(w))$  a la integral del enunciado (de integrated hazard function).

$H(t \wedge X(w))$  es monótona creciente y continua a la derecha en  $t$ . Veamos que es predecible:

$t \wedge X(w)$  es predecible pues

$$\{(t, w) : t \wedge X(w) > s\} = \{t > s\} \times \{w : X(w) > s\}$$

y el segundo conjunto claramente pertenece a  $\mathcal{H}_s$  y por tanto el conjunto completo tiene la forma de un conjunto del  $\sigma$ -álgebra predecible.

Además,  $H$  es monótona creciente y continua a la izquierda en  $x$  así que tiene inversa tal que  $H(x) \geq y \iff x \geq H^{-1}(y)$ . Por tanto,  $\{(t, w) : t \wedge X(w) \geq H^{-1}(y)\}$  es un conjunto predecible. Con lo que queda demostrado que  $H(t \wedge X(w))$  es predecible.

Falta comprobar que se verifica

$$E[N(t \wedge X) - H(t \wedge X) / \mathcal{H}_s] = N(s \wedge X) - H(s \wedge X) \quad c.s. \quad (1.33)$$

Por la estructura de la historia interna,  $\mathcal{H}$ , para cualquier función acotada  $g(\bullet)$ ,

$$E[g(X) / \mathcal{H}_s] = \begin{cases} g(X), & \text{en } \{X(w) \leq s\} \\ E[g(X) I_{\{X > s\}}] / E(I_{\{X > s\}}), & \text{en } \{X(w) > s\} \end{cases}$$

que puede escribirse en función de  $F$  de la siguiente manera:

$$E[g(X) / \mathcal{H}_s] = \frac{1}{1 - F(s)} \int_s^\infty g(u) F(du); \text{ en } \{X(w) > s\} \quad (1.34)$$

En  $\{X(w) \leq s\}$ ,  $N(t \wedge X) = N(s \wedge X) = 1$  y  $H(t \wedge X) = H(s \wedge X) = H(X)$ , por tanto,  $\{X(w) \leq s\}$  es  $\mathcal{H}_s$ -medible, y (1.33) se verifica en este caso.

En su complementario,  $\{X(w) > s\}$ , debido a (1.34),

$$E[N(t \wedge X)/\mathcal{H}_s] = \frac{1}{1 - F(s)} \int_s^t F(du) = \frac{F(t) - F(s)}{1 - F(s)}$$

e integrando por partes,

$$(1 - F(s)) [E[H(t \wedge X)/\mathcal{H}_s] - H(s)] = F(t) - F(s)$$

Por fin,  $E[N(t \wedge X) - H(t \wedge X)/\mathcal{H}_s] = -H(s)$ , con lo que se verifica la propiedad de martingala también en el caso de  $\{X(w) > s\}$ . Con lo que se completa la demostración de la proposición. ■

Por fin, para alcanzar nuestra meta, falta ampliar esta proposición al caso de proceso de Poisson. Basta pensar en el proceso de Poisson,  $N(t)$ , como la suma de procesos de un punto. Es decir, si tomamos la serie de tiempos en los que se da una ocurrencia del proceso, estos son una serie de tiempos de parada  $\{T_n\}$  con la propiedad vista en (1.32) y además se puede escribir

$$N(t) = \sum_{n=1}^{\infty} [N(t \wedge T_n) - N(t \wedge T_{n-1})] \equiv \sum_{n=1}^{\infty} N^n(t)$$

donde cada  $N^n(t)$  es un proceso de un punto con incremento en  $T_n$ . La demostración de que  $N(t)$  admite un compensador es casi igual a la anterior (ajustando la propiedad de martingala, debido a las diferencias en los sumandos de  $N(t)$ ) y utilizando la aditividad de los compensadores. Gracias al teorema de Doob-Meyer (teorema 1.4.5) y puesto que el proceso  $N(t)$  es sólo localmente acotado, se consigue demostrar que el proceso de Poisson es un semimartingala que puede expresarse como la suma de una martingala local,  $M(t)$ , y un proceso acumulativo y predecible,  $A(t)$ , que se prueba que es de varianza acotada

, en caso contrario no se podría realizar las integrales que aparecen en (1.30) o (1.31) (ver ( Andersen et al. (1993)) para más detalles). Según esto, se puede escribir

$$N(t) = M(t) + A(t)$$

donde  $M$  y  $A$  son los ya mencionados.

Sin entrar en más detalles de integración estocástica se demuestra que el compensador de una martingala local cuyo cuadrado es localmente integrable, es el límite en probabilidad de aproximaciones de la forma

$$\sum \text{var} [M(t_i) - M(t_{i-1})/\mathcal{F}_{t_{i-1}}]$$

para particiones cada vez más finas  $0 < t_1 < \dots < t_n$ . Intuitivamente, usando el lema 1.4.1 y que  $M = N - A$  podemos comprender que se llegue a que el compensador del proceso de Poisson es  $A(t) = \int_0^t \lambda(s) ds$ , donde  $\lambda(t)$  es la conocida intensidad del proceso de Poisson.

Por fin podemos escribir el siguiente teorema.

**Teorema 1.4.7** *Un proceso de Poisson con intensidad  $\lambda(t)$ , es una semimartingala es decir, tiene una única descomposición de la forma*

$$N(t) = M(t) + A(t)$$

donde  $M$  es una martingala local y  $A(t)$  es  $\mathcal{F}$ -predecible y de varianza acotada tal que  $A(t) = \int_0^t \lambda(s) ds$ .







## Capítulo 2

# Generalizaciones del proceso de Poisson

El proceso de Poisson homogéneo es el caso más simple de proceso de Poisson y resultaba claramente insuficiente a la hora de modelizar casos reales, para esto se ampliaba su concepto al caso no homogéneo que resulta ser una mejora notable según comentamos anteriormente. Aún así, sigue habiendo casos en los que no modeliza suficientemente al proceso puntual observable; para ampliar todavía más las posibilidades del proceso que estamos estudiando relajaremos algunas de sus propiedades. Vamos a estudiar el proceso de Poisson multidimensional (sin la restricción de que los puntos pertenezcan a la recta real), el proceso de Poisson trasladado (un proceso de Poisson con posibles supresiones e inserciones), el proceso de Poisson compuesto (aquél en el que se considera a cada punto con una variable asociada que aporta más información del fenómeno en estudio) y por último, el proceso de Poisson doblemente estocástico (un proceso de Poisson cuya intensidad es a su vez un proceso estocástico). Éste será el proceso que estudiaremos con mayor profundidad puesto que ese es nuestro objetivo principal.

## 2.1 Proceso de Poisson multidimensional

Hasta el momento, hemos tomamos como espacio en el que ocurrían los puntos,  $\mathcal{X}$ , la semirrecta  $[t_0, \infty)$ ; realmente nos hemos restringido al espacio tiempo. Pero fijémonos en algunos de los ejemplos que avanzábamos en la página 1. Tanto en el que se trababa del número de estrellas en una región del espacio, como el del número y localización de los movimientos sísmicos, el espacio  $\mathcal{X}$  no es unidimensional, sino multidimensional. Por lo tanto, necesitaríamos una generalización del proceso de Poisson a varias dimensiones.

Un *proceso de Poisson multidimensional* es la generalización de ese proceso en el que el espacio  $\mathcal{X}$  puede ser el espacio  $n$ -dimensional  $\mathbb{R}^n$ , de forma que la coordenada de cada punto es un vector  $n$ -dimensional  $x$ . Si  $\mathcal{A}$  es un subconjunto de  $\mathcal{X}$ ,  $N(\mathcal{A})$  es el número de puntos con coordenadas en  $\mathcal{A}$ .

Un proceso de Poisson multidimensional  $\{N(\mathcal{A}) : \mathcal{A} \subseteq \mathcal{X}\}$  con intensidad  $\lambda(x)$  verifica:

1. para cada  $\mathcal{A} \subseteq \mathcal{X}$ ,  $N(\mathcal{A})$  tiene distribución de Poisson con parámetro

$$\int_{\mathcal{A}} \lambda(x) dx;$$

2. si  $\mathcal{A}_1, \mathcal{A}_2, \dots, \mathcal{A}_k, \dots$  son subconjuntos disjuntos de  $\mathcal{X}$ , entonces las variables  $N(\mathcal{A}_1), N(\mathcal{A}_2), \dots, N(\mathcal{A}_k), \dots$  son estadísticamente independientes;
3. si  $\mathcal{A}_1, \mathcal{A}_2, \dots, \mathcal{A}_k, \dots$  son subconjuntos disjuntos de  $\mathcal{X}$ , entonces

$$N\left(\bigcup_{k=1}^{\infty} \mathcal{A}_k\right) = \sum_{k=1}^{\infty} N(\mathcal{A}_k).$$

La propiedad (2) es la análoga a la propiedad de incrementos independientes.

A continuación vamos a generalizar la función de densidad muestral vista en (1.25) y su logaritmo en (1.26). Sea  $\Omega$  una realización particular del proceso de Poisson en  $\mathcal{X}$  con  $N(\mathcal{X}) = n$ . Sean los puntos  $\{X_1, \dots, X_n, N(\mathcal{X}) = n\}$ . La esfera centrada en el punto  $X_i$  de radio  $\rho_i$  se notará  $S(X_i) = \{x : \|x - X_i\| \leq \rho_i\}$ . Para un orden dado de puntos  $\{X_1, \dots, X_n\}$  y teniendo en cuenta que  $S(X_1), \dots, S(X_n)$  y  $\mathcal{X} - \bigcup_{i=1}^n S(X_i)$  son disjuntos, por la propiedad de “incrementos” independientes tenemos que:

$$\begin{aligned} P[N(S(X_1)) = 1, \dots, N(S(X_n)) = 1, N(\mathcal{X}) = n] &= \\ \left[ \prod_{i=1}^n \int_{S(X_i)} \lambda(x) dx \exp\left(-\int_{S(X_i)} \lambda(x) dx\right) \right] \exp\left(-\int_{\mathcal{X} - \bigcup_{i=1}^n S(X_i)} \lambda(x) dx\right) &= \\ = \left[ \prod_{i=1}^n \int_{S(X_i)} \lambda(x) dx \right] \exp\left(-\int_{\mathcal{X}} \lambda(x) dx\right). \end{aligned}$$

Dividiendo ambos miembros entre el volumen de las  $n$  esferas, haciendo que el volumen tienda a 0 y como hay  $n!$  formas de ordenar  $X_1, \dots, X_n$ , la función de densidad para una realización del proceso queda de la siguiente forma:

$$p(\Omega) = \frac{1}{n!} \left[ \prod_{i=1}^n \lambda(X_i) \right] e^{-\int_{\mathcal{X}} \lambda(x) dx}. \quad (2.1)$$

También puede escribirse,

$$p(w) = \frac{1}{N(\mathcal{X})!} \exp\left[-\int_{\mathcal{X}} \lambda(x) dx + \int_{\mathcal{X}} \ln[\lambda(x)] N(dx)\right],$$

donde

$$\int_{\mathcal{X}} \ln[\lambda(x)] N(dx) = \begin{cases} 0, & N(\mathcal{X}) = 0 \\ \sum_{i=1}^{N(\mathcal{X})} \ln[\lambda(X_i)], & N(\mathcal{X}) \geq 1. \end{cases}$$

Esta expresión para  $p(w)$  es análoga a (1.25). El factor  $1/N(\mathcal{X})!$  aparece porque en el espacio  $\mathcal{X}$  multidimensional los puntos no están ordenados.

El logaritmo de la función de verosimilitud es

$$\mathcal{L} = - \int_{\mathcal{X}} \lambda(x) dx + \int_{\mathcal{X}} \ln[\lambda(x)] N(dx) - \ln[N(\mathcal{X})!]. \quad (2.2)$$

Análogamente al caso unidimensional, calculamos la densidad de probabilidad conjunta de los puntos aleatorios  $x_1, \dots, x_n$  de  $\mathcal{X}$  dado  $N(\mathcal{X}) = n$ , usando (2.1) y la propiedad (1) de la definición de proceso de Poisson multidimensional:

$$p_{x_1, \dots, x_n / N(\mathcal{X})}(X_1, \dots, X_n / n) = \frac{p(\Omega)}{P[N(\mathcal{X}) = n]} \\ \frac{\frac{1}{n!} [\prod_{i=1}^n \lambda(X_i)] e^{-\int_{\mathcal{X}} \lambda(x) dx}}{\frac{1}{n!} \left[ - \int_{\mathcal{X}} \lambda(x) dx \right]^n e^{-\int_{\mathcal{X}} \lambda(x) dx}} = \prod_{i=1}^n \frac{\lambda(X_i)}{- \int_{\mathcal{X}} \lambda(x) dx}. \quad (2.3a)$$

Esta densidad de probabilidad condicionada nos dice que las localizaciones de los puntos en  $\mathcal{X}$ , dado el número de puntos, son variables aleatorias independientes e idénticamente distribuidas.

También de forma análoga al caso unidimensional, en el que se usaba la densidad condicionada anterior, el funcional característico de un proceso de Poisson multidimensional viene dado por

$$\Phi_N(iv) = \exp \left[ - \int_{\mathcal{X}} \lambda(x) (e^{iv(x)} - 1) dx \right].$$

## 2.2 Proceso de Poisson trasladado

Los datos de un proceso puntual observable pueden provenir de otro proceso puntual oculto; al proceso observable lo llamaremos *trasladado*. El proceso oculto suele ser el realmente interesante; algunos ejemplos son:

- para la reconstrucción de imágenes de órganos humanos, se introduce en el cuerpo del paciente sustancia radiactiva. Se detectan los fotones emitidos, pero estos han sufrido alguna variación en su trayectoria al pasar por los diferentes tejidos e incluso se puede perder la medición de alguno.
- Ocurre algo similar al intentar reconstruir las imágenes de astros; la luz que emiten se distorsiona al pasar por la atmósfera.
- En telecomunicaciones, se trabaja también con procesos trasladados ya sea de imágenes como de sonidos afectados por ruidos. Lo interesante es reconstruir los originales.

Los puntos del proceso oculto, que ocurren en el espacio  $\mathcal{X}$ , se denominan *input del proceso*. Los puntos observados son los puntos *output del proceso* y se recogen en el espacio  $\mathcal{Y}$ .

La localización  $x$  de un punto se traslada a la localización  $y$  del espacio output de acuerdo con la densidad de probabilidad condicional  $p(y/x)$ , denominada *probabilidad de transición*, que se puede entender como la función de diseminación de los puntos. Un conjunto de puntos input pasará a ser un conjunto de outputs dependiendo de una función de densidad conjunta; nosotros nos restringiremos al caso en el que los puntos son trasladados independientemente unos de otros, por lo que la función de densidad conjunta es el producto de las funciones de transición para cada punto.

Sean  $N(\mathcal{X})$  e  $M(\mathcal{Y})$  el número de puntos que ocurren en  $\mathcal{X}$  e  $\mathcal{Y}$ , respectivamente; en general no son iguales debido a que algunos inputs han podido perderse y no aparecer en los outputs y han podido añadirse otros debido a algún ruido. Podemos ver un esquema de la

idea que hemos descrito en la siguiente figura 2.2.1.:

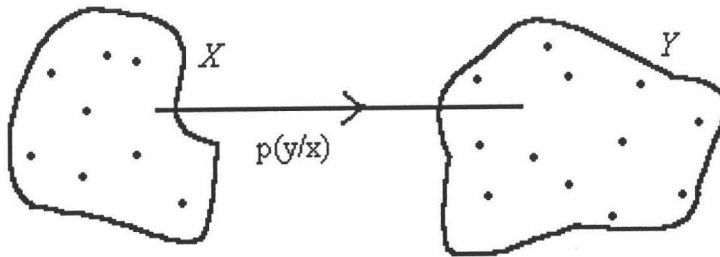


Figura 2.2.1.: Traslado de puntos de un proceso de Poisson

En general, y como ejemplo tal vez más importante por su significado y amplitud de miras, se puede interpretar cualquier proceso puntual como proceso trasladado, puesto que siempre se comete algún error en la observación y posterior medición de los puntos del proceso, por lo que lo que observamos es el proceso “distorsionado”.

En el caso que nos ocupa del proceso de Poisson y con la restricción mencionada del traslado de puntos de forma independiente, veremos en el siguiente teorema que el proceso de Poisson trasladado es un proceso de Poisson.

**Teorema 2.2.1 (Un p. P. trasladado es un proceso de Poisson)**  
 Sea el proceso de Poisson,  $\{N(A) : A \subseteq \mathcal{X}\}$ , con función de intensidad integrable  $\{\lambda(x) : x \in \mathcal{X}\}$ . Este proceso es trasladado al espacio  $\mathcal{Y}$  en forma del proceso de recuento  $\{M(B) : B \subseteq \mathcal{Y}\}$ , donde cada punto se traslada de forma independiente con la densidad de transición  $p(y/x)$ .  
 Entonces:

- a) si no hay supresiones ni inserciones, el proceso output  $\{M(B) : B \subseteq \mathcal{Y}\}$  es un proceso de Poisson con función de intensidad

$$\mu(y) = \int_{\mathcal{X}} p(y/x) \lambda(x) dx;$$

- b) si los puntos inputs son suprimidos independientemente con probabilidad  $d(y/x)$  al ser trasladados al espacio output y si se insertan los puntos de un proceso de Poisson independiente con intensidad  $\mu_0(y)$ , en este caso, el proceso output es un proceso de Poisson con intensidad

$$\mu(y) = \int_{\mathcal{X}} p(y/x) s(y/x) \lambda(x) dx + \mu_0,$$

donde  $s(y/x) = 1 - d(y/x)$  es la probabilidad de que un punto trasladado de  $\mathcal{X}$  a  $\mathcal{Y}$  sobreviva.

#### Demostración.

- (a) El funcional característico del proceso output viene dado por

$$\begin{aligned} E \left[ \exp \left( i \int_{\mathcal{Y}} v(y) M(dy) \right) \right] &= E \left[ \exp \left( i \sum_{j=0}^{M(\mathcal{Y})} v(y_j) \right) \right] \\ &= E \left\{ E \left[ \exp \left( i \sum_{j=0}^{N(\mathcal{X})} v(y_j) \right) / N(\mathcal{X}) \right] \right\}, \end{aligned} \quad (2.4)$$

donde  $N(\mathcal{X}) = M(\mathcal{Y})$  pues no hay inserciones ni supresiones de inputs.

Dadas la localizaciones  $x_1, \dots, x_{N(\mathcal{X})}$  de los  $N(\mathcal{X})$  inputs, debido a (2.3a) son independientes e idénticamente distribuidas y su función de densidad conjunta es

$$\frac{\lambda(x)}{\int_{\mathcal{X}} \lambda(x) dx}.$$



Puesto que cada punto se traslada independientemente con probabilidad de transición  $p(y/x)$ , dado  $N(\mathcal{X})$ , los outputs son independientes con función de densidad

$$\frac{\int_{\mathcal{X}} p(y/x) \lambda(x) dx}{\int_{\mathcal{X}} \lambda(x) dx}.$$

Por tanto

$$\begin{aligned} E \left[ \exp \left( i \sum_{j=0}^{N(\mathcal{X})} v(y_j) \right) / N(\mathcal{X}) \right] &= [E (e^{iv(y)} / N(\mathcal{X}))]^{N(\mathcal{X})} \\ &= \left[ \frac{\int_{\mathcal{Y}} \left( \int_{\mathcal{X}} e^{iv(y)} p(y/x) \lambda(x) dx \right) dy}{\int_{\mathcal{X}} \lambda(x) dx} \right]^{N(\mathcal{X})}. \end{aligned}$$

Sustituyendo esta expresión en (2.4), el funcional característico del proceso output es

$$\begin{aligned} E \left[ \exp \left( i \int_{\mathcal{Y}} v(y) M(dy) \right) \right] &= \\ &= \sum_{n=0}^{\infty} \left[ \frac{\int_{\mathcal{Y}} \int_{\mathcal{X}} e^{iv(y)} p(y/x) \lambda(x) dx dy}{\int_{\mathcal{X}} \lambda(x) dx} \right]^n \frac{(\int_{\mathcal{X}} \lambda(x) dx)^n}{n!} e^{-\int_{\mathcal{X}} \lambda(x) dx} = \\ &= e^{-\int_{\mathcal{X}} \lambda(x) dx} \exp \left[ \int_{\mathcal{Y}} \int_{\mathcal{X}} e^{iv(y)} p(y/x) \lambda(x) dx dy \right] \\ &= e^{-\int_{\mathcal{X}} \lambda(x) dx} \exp \left[ \int_{\mathcal{Y}} e^{iv(y)} \mu(y) dy \right] \end{aligned}$$

donde  $\mu(y) = \int_{\mathcal{X}} p(y/x) \lambda(x) dx$  y como

$$\begin{aligned} \int_{\mathcal{Y}} \mu(y) dy &= \int_{\mathcal{Y}} \int_{\mathcal{X}} p(y/x) \lambda(x) dx dy \\ &= \int_{\mathcal{X}} \lambda(x) \left( \int_{\mathcal{Y}} p(y/x) dy \right) dx = \int_{\mathcal{X}} \lambda(x) dx \end{aligned}$$

tenemos que

$$E \left[ \exp \left( i \int_{\mathcal{Y}} v(y) M(dy) \right) \right] = \exp \left[ \int_{\mathcal{Y}} (e^{iv(y)} - 1) \mu(y) dy \right]$$

y tal y como vimos en (1.28), este es el funcional característico de un proceso de Poisson definido sobre el espacio  $\mathcal{Y}$  con intensidad  $\mu(y) = \int_{\mathcal{X}} p(y/x) \lambda(x) dx$  como se esperaba.

(b) Nos basaremos en dos ideas para demostrar esta parte del teorema:

intuitivamente es claro que al suprimir de forma independiente puntos del proceso de Poisson con intensidad  $\lambda(x)$  y probabilidad de supresión  $d(y/x)$ , el proceso resultante es un proceso de Poisson con intensidad  $(1 - d(y/x))\lambda(x)$ ; además, la suma de dos procesos de Poisson,  $Z$  y  $Z_0$ , con intensidades  $\mu(y)$  y  $\mu_0(y)$  respectivamente, es otro proceso de Poisson,  $Y$ , con intensidad  $\mu(y) + \mu_0(y)$ , lo que se puede probar fácilmente mediante la función característica, tal y como sigue: llamemos  $M_Y(iv)$  a la función característica del proceso suma  $\{Y(\sigma) : t_0 \leq \sigma < T\}$  y  $M_Z(iv)$  y  $M_{Z_0}(iv)$  a las de los procesos sumandos. Por definición de la función característica tenemos que

$$\begin{aligned} M_Y(iv) &= E [\exp (ivY(s, t))] = \\ &= E [\exp (iv[Z(s, t) + Z_0(s, t)])] = E [e^{ivZ(s, t)} e^{ivZ_0(s, t)}]; \end{aligned}$$

debido a la independencia de los procesos  $Z$  y  $Z_0$  y como sabemos que son procesos de Poisson, la expresión queda

$$\begin{aligned} M_Y(iv) &= E [e^{ivZ(s, t)}] E [e^{ivZ_0(s, t)}] = \\ &= \exp \left[ (e^{iv} - 1) \int_s^t \mu(\sigma) d\sigma \right] \exp \left[ (e^{iv} - 1) \int_s^t \mu_0(\sigma) d\sigma \right] \\ &= \exp \left[ (e^{iv} - 1) \int_s^t (\mu + \mu_0)(\sigma) d\sigma \right] \end{aligned}$$

que es la función característica de un proceso de Poisson con intensidad  $\mu(t) + \mu_0(t)$ , como esperábamos.

Según estas dos ideas y usando la parte (a) del teorema, también se verifica lo expuesto en (b). ■

El proceso output, puesto que es un proceso de Poisson como hemos visto en el teorema 2.2.1, sus estadísticos de recuento son los mismos que los calculados en el apartado anterior, con la diferencia de que la función de intensidad de este proceso es la dada por el mismo teorema.

De esta forma el logaritmo de la función de verosimilitud para un proceso de Poisson trasladado es, de acuerdo con (2.1)

$$\mathcal{L} = - \int_{\mathcal{Y}} \mu(y) dy + \int_{\mathcal{Y}} \ln[\mu(y)] M(dy) - \ln[M(\mathcal{Y})!]$$

Para los casos particulares (a) y (b) del teorema, la función de máximo-verosimilitud queda de la siguiente forma:

- sin inserciones ni supresiones,

$$\mathcal{L} = - \int_{\mathcal{X}} \lambda(x) dx + \int_{\mathcal{Y}} \ln \left[ \int_{\mathcal{X}} p(y/x) \lambda(x) dx \right] M(dy) - \ln[M(\mathcal{Y})!]$$

- con inserciones y supresiones,

$$\begin{aligned} \mathcal{L} = & - \int_{\mathcal{X}} \beta(x) \lambda(x) dx + \int_{\mathcal{Y}} \mu_0(y) dy \\ & + \int_{\mathcal{Y}} \ln \left[ \int_{\mathcal{X}} p(y/x) s(y/x) \lambda(x) dx + \mu_0(y) \right] M(dy) - \ln[M(\mathcal{Y})!] \end{aligned}$$

donde  $\beta(x) = \int_{\mathcal{Y}} p(y/x) s(y/x) dy$  es la probabilidad media de supervivencia de un punto trasladado de  $\mathcal{X}$  a  $\mathcal{Y}$ .

## 2.3 Proceso de Poisson compuesto

Un *proceso puntual marcado* es un proceso puntual con una variable aleatoria asociada a cada punto. Cada suceso o punto puede tener asociado una cantidad aleatoria interesante de estudiar, ésta será su *marca*. Un ejemplo ilustrativo puede ser la ocurrencia de descargas eléctricas producidas por una fuente cualquiera; además de estudiarse el número de sucesos que ocurren interesa estudiar la cantidad de energía descargada en ese instante. En este ejemplo, la cantidad de energía en el instante  $i$ -ésimo, sería la marca  $i$ -ésima.

Sea  $\mathcal{U}$  el espacio de posibles marcas. Llamaremos  $u_i$  a la marca del suceso ocurrido en el instante (si el proceso es temporal)  $w_i$ . El espacio  $\mathcal{U}$  puede ser numerable o no numerable, por ejemplo:

- un proceso puntual con varias ocurrencias en un instante, puede pensarse como un proceso puntual marcado en el que la marca  $u_i$  del instante  $w_i$  sea un entero positivo que exprese el número de ocurrencias en  $w_i$ . Entonces  $\mathcal{U}$  es numerable e igual a  $\mathbb{N}$ . Este ejemplo es la generalización del proceso de Poisson a que varios sucesos pueden ocurrir simultáneamente, es decir, un proceso de Poisson que no cumple la propiedad de regularidad.
- La energía descargada en un fenómeno eléctrico, como mencionamos antes, puede ser la marca en un proceso puntual; entonces,  $\mathcal{U}$  es  $[0, \infty)$  y por tanto no numerable.
- Si se estudia un proceso multidimensional en el espacio, la ocurrencia en el instante  $w_i$  estaría marcada por la posición bi-tridimensional del punto.

En la primera aproximación que hacíamos a los procesos que estamos estudiando en la sección 1.1, definíamos un proceso de recuento



como el proceso que contaba las ocurrencias de un proceso puntual. Asimismo, un *proceso acumulador de marcas*,  $x(t)$ , se define como:

$$x(t) = \sum_{i=0}^{N(t)} u_i,$$

donde  $u_0 = 0$ ,  $\{N(t) : t \geq t_0\}$  es un proceso de recuento que indica el número de sucesos en  $[t_0, t)$  y  $u_i$  es la marca del suceso en el instante  $w_i$ . Es decir,  $x(t)$ , el proceso acumulador de marcas, es la suma de las marcas de los puntos que han ocurrido en  $[t_0, t)$ .

Nosotros restringiremos nuestro estudio al caso particular de un proceso puntual marcado donde

1.  $\{N(t) : t \geq t_0\}$  es un proceso de Poisson no homogéneo con función de intensidad  $\{\lambda(t) : t \geq t_0\}$  y
2.  $\{u_i\}$  son variables aleatorias independientes e idénticamente distribuidas e independientes del proceso  $N(t)$ .

Un proceso puntual marcado con estas propiedades se denomina *proceso de Poisson compuesto o marcado*. De la misma forma se denomina su proceso acumulador de marcas, por ser equivalentes.

Puesto que las marcas son independientes entre sí, el proceso acumulador,

$$x(t) = \sum_{i=0}^{N(t)} u_i,$$

posee incrementos independientes.

### 2.3.1 Estadísticos de un proceso de Poisson compuesto

#### Funcional y función característicos

El funcional característico de un proceso de Poisson compuesto  $\{x(t) : t_0 \leq t < T\}$  se define como

$$\Phi_x(iv) = E \left[ \exp \left( i \int_{t_0}^T v'(t) dx(t) \right) \right],$$

donde  $v'(t)$  es el vector traspuesto de  $v(t)$  y  $\{v(t) : t_0 \leq t < T\}$  es una función vectorial arbitraria. La integral se define como la siguiente integral de recuento:

$$\int_{t_0}^T v'(t) dx(t) = \begin{cases} 0, & N(T) = 0 \\ \sum_{i=1}^{N(T)} v'(w_i)u_i, & N(T) \geq 1. \end{cases}$$

Evaluemos  $\Phi_x(iv)$ :

$$\begin{aligned} \Phi_x(iv) &= \sum_{n=0}^{\infty} E \left[ \exp \left( i \int_{t_0}^T v'(\sigma) d\sigma \right) / N(T) = n \right] = \\ &= P[N(T) = 0] + \\ &+ \sum_{n=1}^{\infty} P[N(T) = n] E \left[ \exp \left( i \sum_{i=1}^{N(T)} v'(w_i)u_i \right) / N(T) = n \right] \end{aligned} \quad (2.5)$$

Debido a la independencia entre las marcas y de éstas con  $N(T)$  y que son idénticamente distribuidas tenemos que:

$$\begin{aligned}
 & E \left[ \exp \left( i \sum_{j=1}^{N(T)} v'(w_j) u_j \right) / N(T) = n \right] = \\
 & = E \left\{ E \left[ \exp \left( i \sum_{j=1}^{N(T)} v'(w_j) u_j \right) / N(T) = n; w_1, w_2, \dots, w_{N(T)} \right] / N(T) = n \right\} \\
 & = E \left[ \prod_{j=1}^{N(T)} E \left( e^{v'(w_j) u_j} / N(T) = n; w_1, w_2, \dots, w_{N(T)} \right) / N(T) \right] \\
 & = E \left\{ \prod_{j=1}^{N(T)} M_u(i v(w_j)) / N(T) = n \right\},
 \end{aligned}$$

donde  $M_u(i\alpha)$  es la función característica de las marcas,

$$M_u(i\alpha) \equiv E[e^{i\alpha' u}].$$

Usando un razonamiento análogo al que utilizamos al calcular el funcional característico de un proceso de Poisson, en (1.28), llegamos a que:

$$E \left[ \exp \left( i \sum_{j=1}^{N(T)} v'(w_j) u_j \right) / N(T) = n \right] = \left[ \frac{\int_{t_0}^T \lambda(t) M_u(i v(t)) dt}{\int_{t_0}^T \lambda(t) dt} \right]^n$$

y sustituyendo en (2.5), tenemos

$$\begin{aligned} \Phi_x(iv) &= \\ & e^{-[\Lambda(T)-\Lambda(t_0)]} + \sum_{n=1}^{\infty} \frac{1}{n!} \left( \int_{t_0}^T \lambda(t) [M_u(iv(t)) - 1] dt \right)^n = \\ & \exp \left\{ \int_{t_0}^T \lambda(t) [M_u(iv(t)) - 1] dt \right\}. \end{aligned} \quad (2.6)$$

Tal y como vimos para el proceso de Poisson, la función característica es un caso particular del funcional característico. Para hallarla, tomemos la función  $v(\sigma)$  tal que

$$v(\sigma) = \begin{cases} 0, & t_0 \leq \sigma < s \\ \alpha, & s \leq \sigma < t \\ 0, & t \leq \sigma T. \end{cases}$$

Entonces,

$$\begin{aligned} \Phi_x(iv) &= E \left[ \exp \left( \int_{t_0}^T \right) \right] = E \left[ e^{\alpha'(x(t)-x(s))} \right] = \\ & = \exp \left[ \int_s^t \lambda(\sigma) d\sigma (M_u(i\alpha) - 1) \right] = M_{x(t)-x(s)}(i\alpha) \end{aligned} \quad (2.7)$$

que es la función característica del incremento  $x(t) - x(s)$ .

Si tomamos  $t = t_0$ , (2.7) es la función característica de  $x(t)$ :

$$M_{x(t)}(i\alpha) = \exp \left[ (M_u(i\alpha) - 1) \left( \int_{t_0}^t \lambda(\sigma) d\sigma \right) \right] \quad (2.8)$$

### Esperanza y matriz de covarianzas

Usaremos la función característica que es generatriz de momentos, para hallar tanto la esperanza como la matriz de covarianza de  $x(t)$ .



Sabemos que

$$E[x(t)] = \left. \frac{\partial M_{x(t)}(i\alpha)}{i\partial\alpha} \right|_{\alpha=0}$$

y que

$$E[x(t)x'(t)] = \left. \frac{\partial^2 M_{x(t)}(i\alpha)}{i^2\partial\alpha^2} \right|_{\alpha=0},$$

donde  $\partial M/\partial\alpha$  es el vector gradiente de  $M$  con respecto a  $\alpha$  y  $\partial^2 M/\partial\alpha^2$  es la matriz Hessiana.

Usando (2.8) tenemos

$$\begin{aligned} \left. \frac{\partial M_{x(t)}(i\alpha)}{i\partial\alpha} \right|_{\alpha=0} &= \left. \frac{\partial}{i\partial\alpha} \exp \left[ \int_{t_0}^t \lambda(\sigma) d\sigma (M_u(i\alpha) - 1) \right] \right|_{\alpha=0} \\ &= E[u] \int_{t_0}^t \lambda(\sigma) d\sigma \end{aligned}$$

es decir, el valor esperado de  $x(t)$  es la media de las marcas por el número medio de puntos en  $[t_0, t)$ .

La matriz de covarianza instantánea de  $x(t)$ , que notaremos  $\Sigma(t)$  es

$$\Sigma(t) = E[x(t)x'(t)] - E[x(t)]E[x'(t)] = E[uu'] \int_{t_0}^t \lambda(\sigma) d\sigma,$$

o sea, es el producto del valor cuadrático medio de las marcas y la varianza del número de puntos en  $[t_0, t)$ .

LLamaremos  $K_x(s, t)$  a la matriz de covarianza, que es por definición

$$\begin{aligned} K_x(s, t) &= E[x(s)x'(t)] - E[x(s)]E[x'(t)] = \\ &E[x(s)(x(t) - x(s))'] + E[x(s)x'(s)] - E[x(s)]E[x'(t)] \end{aligned}$$

por la independencia de los incrementos. Calculando la matriz de covarianza se observa que  $K_x(s, t) = K_x(t, s)$ , con

$$K_x(s, t) = \Sigma(t), \quad \text{para } t_0 \leq t < s.$$

Como  $\int_{t_0}^t \lambda(\sigma) d\sigma = \Lambda(t)$  y es no decreciente,

$$K_x(s, t) = \Sigma \min(s, t) = E[uu'] \min[\Lambda(t), \Lambda(s)].$$

A partir de la esperanza y matriz de covarianza de un proceso de Poisson compuesto, tomando las marcas adecuadas, resultan los mismos momentos para un proceso de Poisson homogéneo y no homogéneo, lo que se resume en la tabla siguiente:

P. Poisson \ Momento	Media	M.C. I.	M. C.
Compuesto	$E[u]\Lambda(t)$	$E[uu']\Lambda(t)$	$E[uu'] \min[\Lambda(t), \Lambda(s)]$
No homogéneo	$\Lambda(t)$	$\Lambda(t)$	$\min[\Lambda(t), \Lambda(s)]$
Homogéneo	$\lambda(t - t_0)$	$\lambda(t - t_0)$	$\lambda \min(t - t_0, s - t_0)$

donde M.C. I. es la matriz de covarianzas instantánea y M.C., la matriz de covarianzas del proceso.

Un proceso de Poisson compuesto puede representarse como una superposición de procesos de Poisson independientes; para ello hay que estudiar por separado los casos en los que el espacio de marcas,  $\mathcal{U}$ , sea numerable de los no numerables. No vamos a abordar el estudio de estas representaciones en este trabajo pues excederían los objetivos de esta sección. Según esta misma división de los procesos de Poisson compuestos en aquellos con espacio de marcas numerable o no, se llegaría a distintas expresiones de la función de densidad muestral. Este tema, tratado en la Memoria de Licenciatura (Rodríguez-Bouzas, 1996), tiene como referencia original el libro de Snyder and Miller (1991).

### 2.3.2 Casos particulares notables

Vamos a detenernos en un par de casos particulares del proceso de Poisson compuesto porque es interesante ver desde este nuevo punto de vista que un proceso de Poisson con supresiones aleatorias sigue siendo un proceso de Poisson. Veremos también que el problema que amenudo se plantea en teoría de colas y otros múltiples casos reales, de que los puntos no ocurren de uno en uno, sino que pueden darse varias ocurrencias a la vez, por lo que no se podrían modelizar mediante un proceso de Poisson, sí que se corresponde a un cierto tipo de proceso de Poisson compuesto, el llamado proceso de Poisson generalizado.

Sea un proceso de Poisson compuesto con un espacio de marcas numerable  $\{U_1, \dots, U_k, \dots\}$ . Llamemos  $p_k$  a la probabilidad de que la marca sea  $U_k$ ;  $k = 1, 2, \dots$ . Entonces la función característica de las marcas queda

$$M_u(i\alpha) = \sum_{k=1}^{\infty} p_k e^{i\alpha' U_k}$$

y por tanto, el funcional característico del proceso acumulador en este caso particular resulta

$$\Phi_x(iv) = \exp \left[ \int_{t_0}^T \lambda(\sigma) \left( \sum_{k=1}^{\infty} p_k e^{iv'(\sigma) U_k} - 1 \right) d\sigma \right]. \quad (2.9)$$

Observemos dos casos particulares de entre éstos que tienen el espacio de las marcas,  $U$ , numerable:

- Sea  $U = \{0, 1\}$ , entonces el proceso de Poisson compuesto es un proceso de Poisson en el que se pierde, o no se cuenta, un punto con probabilidad  $p$ , la marca de ese punto será 0. El proceso acumulador,  $x(t)$ , cuenta los puntos con marca 1 y (2.9) queda

$$\Phi_x(iv) = \exp \left[ \int_{t_0}^T (1-p)\lambda(\sigma)(e^{iv(\sigma)} - 1) d\sigma \right].$$

Esta sería la demostración de que un proceso de Poisson con intensidad  $\lambda(t)$  con supresiones aleatorias, es un proceso de Poisson con intensidad  $(1 - p)\lambda(t)$ .

- Un proceso de Poisson en el que se permite que los puntos puedan ocurrir de forma simultánea, que ocurran  $k$  sucesos con probabilidad  $p_k$ , deja de ser un proceso de Poisson según lo hemos definido. Este se puede interpretar como un proceso de Poisson compuesto cuyo espacio de marcas es el espacio numerable  $U = \{1, 2, \dots, k, \dots\}$ . El proceso de Poisson compuesto correspondiente se denomina *proceso de Poisson generalizado*<sup>1</sup>. Es por tanto un proceso de Poisson sin la condición de regularidad. En este caso,

$$M_u(i\alpha) = \sum_{k=1}^{\infty} p_k e^{i\alpha k},$$

y  $\Phi_x$  resulta ser

$$\Phi_x(iv) = \exp \left[ \int_{t_0}^T \lambda(\sigma) \left( \sum_{k=1}^{\infty} p_k e^{iv(\sigma)k} - 1 \right) \right].$$

Suele ser difícil desarrollar esta expresión y obtener la distribución de  $x(t)$ . En el caso de que  $p_k = 0, k \neq 1$  y  $p_1 = 1$ , claramente  $x(t)$  se reduce a un proceso de Poisson.

- Un proceso de Poisson con función de intensidad de la forma

$$\lambda(t) = \sum_{i=1}^K \lambda_i(t) \quad (2.10)$$

<sup>1</sup>En la literatura puede encontrarse que denominan *generalizado* al proceso de Poisson no homogéneo, pues es una generalización del proceso de Poisson (homogéneo), lo que podría llevar a error.

se denomina *proceso de Poisson mixto*. Puesto que la suma de procesos de Poisson independientes es un proceso de Poisson con intensidad la suma de las intensidades, el proceso de Poisson mixto puede interpretarse como la suma de  $K$  procesos de Poisson. Los puntos de los  $K$  procesos se mezclan para formar el proceso de Poisson mixto.

Un suceso que ocurre en el instante  $t$  es originario del proceso  $i$ -ésimo con una probabilidad  $p_i$ . Llamando  $A$  al suceso de que ocurra un punto de la mezcla de puntos en  $[t, t + \Delta t)$  y  $A_i$  al suceso de que ocurra un punto proveniente de la componente  $i$ -ésima en  $[t, t + \Delta t)$ ,

$$p_i = P(A_i/A) = P(A/A_i) \frac{P(A_i)}{P(A)} \approx 1 \times \frac{\lambda_i(t) \Delta t}{\lambda(t) \Delta t}$$

por lo que podemos tomar

$$p_i = \frac{\lambda_i(t)}{\sum_{i=1}^K \lambda_i(t)} = \frac{\lambda_i(t)}{\lambda(t)}.$$

La definición y estudio de un proceso de Poisson mixto no es algo trivial como podría pensarse si sólo se le interpreta como un proceso de Poisson con una intensidad más o menos complicada. Un proceso de Poisson con una intensidad exactamente de la forma (2.10) se puede interpretar como proceso de Poisson compuesto de la siguiente forma; imaginemos que los puntos estén marcados, de forma que cada marca indica la procedencia del punto; se puede tomar como espacio de marcas,  $\mathcal{U}$ , al conjunto  $\{1, 2, \dots, K\}$ . Esta visión del proceso de Poisson mixto como caso especial del compuesto es interesante a la hora de abordar la estimación de los parámetros que influyen en la intensidad pues podemos apoyarnos en la expresión de la función de densidad muestral obtenida a través de la representación de un proceso

de Poisson compuesto con espacio de marcas numerable a la que nos referíamos en la página 57, facilitándose así el trabajo de estimación. Esta interpretación del proceso de Poisson mixto es tanto más ventajosa cuanto mejor sea el conocimiento de  $\lambda$  y su descomposición. Para más detalles damos la misma referencia que entonces.

## 2.4 Proceso de Poisson doblemente estocástico

En este capítulo ya hemos comentado distintas razones para generalizar el proceso de Poisson. Otra razón puede ser el hecho de que la ocurrencia de los sucesos de un proceso puntual dependa de un proceso externo. El proceso puntual que condicionado a ese proceso externo resulte ser un proceso de Poisson se denomina *proceso de Poisson doblemente estocástico*; también se conoce como *proceso de Cox* pues fue Cox (1995), quien comenzó a estudiarlo en profundidad. En ocasiones se encuentra en la literatura que denominan a este proceso, proceso de Poisson condicionado, compuesto o mixto, pero nosotros hemos usado esas dos últimas denominaciones para los procesos estudiados en la sección 2.3 y subsección 2.3.2, respectivamente.

El proceso de Poisson doblemente estocástico (P.P.D.E.) es, por tanto, e intuitivamente, una generalización del proceso de Poisson cuya intensidad en la ocurrencia de los puntos está influenciada por un proceso externo de forma que la intensidad resulta ser un proceso estocástico en sí mismo. A ese proceso externo se le denomina *proceso de información* por ser el que rige la evolución del proceso puntual.

Más concretamente, llamando  $\{x(t) : t \geq t_0\}$ , a un proceso continuo a la izquierda, al proceso externo que influye en la ocurrencia de los puntos del proceso de recuento  $\{N(t) : t \geq t_0\}$ , el proceso de

Poisson doblemente estocástico se define como sigue.

**Definición 2.4.1 (Proceso de Poisson doblemente estocástico)**

Un proceso de recuento,  $\{N(t) : t \geq t_0\}$ , es un proceso de Poisson doblemente estocástico con proceso de intensidad  $\{\lambda(t, x(t)) : t \geq t_0\}$  si para casi toda trayectoria del proceso  $\{x(t) : t \geq t_0\}$ ,  $N(\bullet)$  es un proceso de Poisson con función de intensidad  $\{\lambda(t, x(t)) : t \geq t_0\}$ .

En otras palabras:

El proceso  $\{N(t) : t \geq t_0\}$  es un proceso de Poisson doblemente estocástico con proceso de intensidad  $\{\lambda(t, x(t)) : t \geq t_0\}$  si condicionado a  $\{x(t) : t \geq t_0\}$  es un proceso de Poisson con función de intensidad  $\{\lambda(t, x(t)) : t \geq t_0\}$ .

Aunque no es habitual en la literatura, también encontramos en el artículo de Serfozo (1972) una definición del P.P.D.E. según su media. Como él bien dice, un P.P.D.E. es en esencia un proceso de Poisson cuya intensidad ( $\lambda$ ), o media ( $\Lambda$ ), es un proceso estocástico, pues recordamos que  $\Lambda(t) = \int_{t_0}^t \lambda(\sigma) d\sigma$ . Esta observación tan simple es vista normalmente sólo en su vertiente exclusivamente matemático, sin embargo, para nosotros será relevante por su significado e idea base de nuestra aportación como se verá más adelante. Serfozo define el P.P.D.E. de la siguiente manera.

**Definición 2.4.2** Sea  $\{\Lambda(t) : t \geq t_0\}$  un proceso estocástico no decreciente, real valuado y continuo a la derecha definido en el espacio de probabilidad  $(\Omega, F, P)$  con  $\Lambda(0) = 0$  c.s. Sea  $\mathcal{G}$  el menor  $\sigma$ -álgebra en  $\Omega$  donde cada  $\Lambda(t)$  sea medible. El proceso no negativo que toma valores enteros no negativos  $\{N(t) : t \geq t_0\}$  definido en el mismo espacio se llama P.P.D.E. con media  $\Lambda(t)$  si  $N(t)$  es un proceso de Poisson con función media  $\Lambda(t)$  condicionando a  $\mathcal{G}$ . Es decir, si  $X(0) = 0$  c.s. y dados  $t_0 < t_1 < \dots < t_r$  y  $n_1, \dots, n_r$  se tiene

- $P[N_{t_1} - N_{t_0} \leq n_1, \dots, N_{t_r} - N_{t_{r-1}} \leq n_r / \mathcal{G}] = \prod_{k=1}^r P[N_{t_k} - N_{t_{k-1}} \leq n_k / \mathcal{G}]$  c.s. para todo  $0 \leq t_0 \leq t_1 \leq \dots \leq t_r$  y  $n = 0, 1, 2, \dots$
- $P[N_t - N_s \leq n / \mathcal{G}] = \frac{1}{n!} (\Lambda_t - \Lambda_s)^n \exp[-(\Lambda_t - \Lambda_s)]$  c.s.

En caso de que  $\Lambda(t)$  sea una función determinista del tiempo, el P.P.D.E. se reduce a un proceso de Poisson (no homogéneo). Sin reducirnos a ese caso extremo, el P.P.D.E. más simple es aquel en el que el proceso de información  $\{x(t) : t \geq t_0\}$  es un conjunto finito de variables aleatorias independientes del tiempo.

El P.P.D.E. tiene muchas aplicaciones en múltiples contextos. Un ejemplo fácil, es la representación de procesos de Poisson superpuestos donde el número de procesos superpuestos cambia aleatoriamente, sin ocurrencias simultáneas. Es decir, supongamos  $k$  procesos de Poisson independientes con intensidad  $\{\lambda_1, \dots, \lambda_k\}$  cada uno interviniendo en un cierto instante o no, según el proceso estocástico  $\{\xi(t) = (\xi_1(t), \dots, \xi_k(t)), t \geq 0\}$ , donde  $\xi_i(t) = 1$  ó  $0$  significa que el  $i$ -ésimo proceso de Poisson interviene o no. Entonces el número de sucesos que ocurren en  $(0, t]$  es un P.P.D.E. con intensidad  $\lambda(t) = \sum_{i=1}^k \lambda_i \xi_i(t)$ .

### 2.4.1 Estadísticos de recuento

Los estadísticos de recuento pueden calcularse por el llamado método de condicionamiento, usando la propiedad que define al P.P.D.E. de que resulta ser un proceso de Poisson,  $\{N(t) : t \geq t_0\}$ , al condicionar al proceso información.

#### Función masa de probabilidad

Así pues, la probabilidad de que el número de puntos que ocurren en  $[t_0, t)$  es  $n$  viene dado por



$$P[N(t) = n] = E\{P[N(t) = n/x(\sigma) : t_0 \leq \sigma < t]\} =$$

$$E\left\{\frac{1}{n!} \left(\int_{t_0}^t \lambda(\sigma, x(\sigma))d\sigma\right)^n \exp\left(-\int_{t_0}^t \lambda(\sigma, x(\sigma))d\sigma\right)\right\} \quad (2.11)$$

para  $n = 0, 1, 2, \dots$

Veamos un ejemplo en el que hallemos la probabilidad de que el proceso tome el valor  $n$ .

**Ejemplo 2.4.3** *Proceso de Poisson a escala aleatoria (randomly scaled Poisson process)*

Supongamos que el proceso información es simplemente una variable aleatoria no negativa,  $x$ , y que  $\lambda(t, x) = x v(t)$ ,  $t \geq t_0$ , donde  $v(t)$  es una función determinista del tiempo. Entonces, evaluando (2.11) se tiene

$$P[N(t) = n] = \frac{1}{n!} \left(\int_{t_0}^t v(\sigma)d\sigma\right)^n E\left[x^n \exp\left(-x \int_{t_0}^t v(\sigma)d\sigma\right)\right]$$

y si existe la función generatriz de momentos de la variable aleatoria  $x$ , en cuyo caso notaremos por  $G_x$ , podemos expresar la probabilidad de la siguiente manera

$$P[N(t) = n] = \frac{1}{n!} \left(\int_{t_0}^t v(\sigma)d\sigma\right)^n G_x^{(n)}\left(-\int_{t_0}^t v(\sigma)d\sigma\right) \quad (2.12)$$

donde  $G_x^{(n)}$  significa la derivada  $n$ -ésima de la función generatriz de momentos.  $\square$

**Ejemplo 2.4.4** Sea un P.P.D.E. y supongamos que la intensidad,  $\{\lambda(t); t \in [0, T]\}$ , es un proceso de banda estrecha; es decir  $\lambda(t) = A \cos[2\pi t + \theta]$ ,  $A \geq 0$  y  $\theta$  uniforme en  $[0, 2\pi]$ . Nosotros supondremos que  $A$  tiene distribución Rayleigh y por lo tanto el proceso  $\{\lambda(t); t \in [0, T]\}$  es gaussiano.

En este caso, la función de covarianza es

$$R(t, s) = \frac{1}{2} E[A] \cos[2\pi(t - s)]$$

y resolviendo la ecuación

$$\alpha \varphi(t) = \int_0^T R(t, s) \varphi(s) ds$$

resulta la raíz doble  $\alpha = T \frac{E[A^2]}{2}$ , valores propios de la covarianza y resulta que las funciones propias asociadas son

$$f_1 = \sqrt{\frac{2}{T}} \operatorname{sen}(2\pi t) \quad \text{y} \quad f_2 = \sqrt{\frac{2}{T}} \cos(2\pi t)$$

por tanto el proceso intensidad,  $\{\lambda(t); t \in [0, T]\}$ , puede expresarse según el desarrollo en serie de Karhunen-Loève

$$\lambda(t) = \operatorname{sen}(2\pi t) \xi_1 + \cos(2\pi t) \xi_2$$

donde  $\xi_1$  y  $\xi_2$  son variables aleatorias normales independientes, de media cero y varianza  $\alpha$ .

La probabilidad de que el P.P.D.E., con esta intensidad, tome el valor  $n$  y sustituyendo en (2.11) queda

$$P[N(t) = n] = E \left\{ \frac{1}{n!} \left( \int_{t_0}^t (\operatorname{sen}(2\pi\sigma) \xi_1 + \cos(2\pi\sigma) \xi_2) d\sigma \right)^n \exp \left( - \int_{t_0}^t (\operatorname{sen}(2\pi\sigma) \xi_1 + \cos(2\pi\sigma) \xi_2) d\sigma \right) \right\}$$

y usando la independencia de  $\xi_1$  y  $\xi_2$  y la expresión (2.12) del ejemplo 2.4.3, obtenemos que

$$\begin{aligned}
 P[N(t) = n] &= \frac{1}{n!} \left( \int_{t_0}^t \text{sen}(2\pi t) d\sigma \right)^n G_{\xi_1}^n \left( \int_{t_0}^t \text{sen}(2\pi t) d\sigma \right) \\
 &\quad + \frac{1}{n!} \left( \int_{t_0}^t \text{cos}(2\pi t) d\sigma \right)^n G_{\xi_1}^n \left( \int_{t_0}^t \text{cos}(2\pi t) d\sigma \right) \\
 &= \frac{1}{n!} \left( \frac{1}{n} \sqrt{\frac{1}{2T}} \right)^n \left[ (\text{cos}(2\pi t) - \text{cos}(2\pi t_0))^n G_{\xi_1}^n (\text{cos}(2\pi t) - \text{cos}(2\pi t_0)) + \right. \\
 &\quad \left. (\text{sen}(2\pi t_0) - \text{sen}(2\pi t))^n G_{\xi_1}^n (\text{sen}(2\pi t_0) - \text{sen}(2\pi t)) \right]
 \end{aligned}$$

□

Pensando en que es posible que se tenga conocimiento acerca de la media del P.P.D.E. dado el proceso información y no de su intensidad, sería interesante ver la probabilidad  $P[N(t) = n]$  desde ese otro punto de vista. Sea  $\Lambda(t, x(t))$  la función paramétrica del proceso de Poisson condicionado; en ese caso, como sabemos,  $\Lambda(t, x(t)) = \int_{t_0}^t \lambda(\sigma, x(\sigma)) d\sigma$  y es la media del P.P.D.E. Obviamente, dada esta relación y como ya comentábamos anteriormente, la función paramétrica es también un proceso influenciado por el proceso de información. Por tanto, podemos reescribir (2.11) como sigue

$$P[N(t) = n] = \frac{1}{n!} E \{ \Lambda(t, x(t))^n e^{-\Lambda(t, x(t))} \} = \frac{1}{n!} G_{\Lambda(t, x(t))}^n(-1) \quad (2.13)$$

donde  $G_{\Lambda(t, x(t))}(s)$  es la función generatriz de momentos de  $\Lambda(t, x(t))$ .

A pesar de que hemos obtenido de forma trivial una expresión para  $P[N(t) = n]$ , su evaluación en casos concretos es en general difícil encontrar una expresión explícita, si no imposible. Se pueden encontrar más ejemplos en Snyder and Miller (1991) o en Jiménez (1991). Nosotros calcularemos la probabilidad de que el proceso sea igual a  $n$  en el caso de que la función paramétrica sea una Normal truncada para cada instante de tiempo,  $t$ , pero lo haremos en la sección 2.5 reservada para este caso de P.P.D.E. pues será estudiado más en profundidad.

### Función característica

Usando nuevamente el método de condicionamiento, la función característica de un P.P.D.E. con función de información  $\{x(t) : t \geq t_0\}$  es la siguiente,

$$M_{N(t)}(iu) = E \left\{ \exp \left[ (e^{iu} - 1) \int_{t_0}^t \lambda(\sigma, x(\sigma)) d\sigma \right] \right\} \quad (2.14)$$

**Ejemplo 2.4.5** *Supongamos nuevamente un P.P.D.E. del tipo del ejemplo 2.4.4. En este caso,*

$$\begin{aligned} M_{N(t)}(iu) &= E \left\{ \exp \left[ (e^{iu} - 1) \left( \int_{t_0}^t f_1(\sigma) \xi_1 d\sigma + \int_{t_0}^t f_2(\sigma) \xi_2 d\sigma \right) \right] \right\} \\ &= M_{\xi_1} \left( i \left( -i \int_{t_0}^t f_1(\sigma) d\sigma \right) \right) M_{\xi_2} \left( i \left( -i \int_{t_0}^t f_2(\sigma) d\sigma \right) \right) \end{aligned}$$

donde se ha usado la independencia de las variables aleatorias y puesto que son Normales de media cero y varianza  $\alpha$ , quedaría

$$\begin{aligned} M_{N(t)}(iu) &= \exp \left\{ \frac{\alpha^2}{2\pi^2 T^2} (e^{iu} - 1)^2 [(\cos(2\pi t) - \cos(2\pi t_0))^2 \right. \\ &\quad \left. + (\sin(2\pi t_0) - \sin(2\pi t))^2] \right\} \end{aligned}$$

□

La función característica de un P.P.D.E. se puede escribir en función de la media del proceso o función paramétrica,  $\Lambda(t, x(t))$ , como sigue

$$M_{N(t)}(iu) = E \left\{ \exp \left[ (e^{iu} - 1) \Lambda(t, x(t)) \right] \right\} = M_{\Lambda(t)}(e^{iu} - 1) \quad (2.15)$$

Nuevamente hay que decir que aunque se tenga esta expresión de la función característica, su evaluación no es siempre posible. Sí que a partir de ella resulta fácil observar que

$$E[N(t)] = E[\Lambda(t)] \text{ y que } Var[N(t)] = Var[\Lambda(t)] + E[\Lambda(t)] \quad (2.16)$$

En caso de que la intensidad del P.P.D.E.,  $\lambda(t, x(t))$ , sea gaussiana se puede encontrar en Valderrama et al.(1995) una expresión fácilmente evaluable para la función característica que damos en el siguiente teorema.

**Teorema 2.4.6** *La función característica de un P.P.D.E.  $\{N(t) : t \geq t_0\}$  con intensidad aleatoria gaussiana  $\{\lambda(t) : t \geq t_0\}$  viene dada por*

$$M_{N(t)}(iu) = \exp \left\{ (e^{iu} - 1)^2 \int_{t_0}^t \int_{t_0}^s R_\lambda(v, s) dv ds + (e^{iu} - 1) \int_{t_0}^t m_\lambda(s) ds \right\} \quad (2.17)$$

donde  $m_\lambda(t)$  y  $R_\lambda(t, s)$  son la media y función de covarianza, respectivamente, de la intensidad.

La demostración, que se puede ver en el artículo citado, tiene como base fundamental el desarrollo en serie de Karhunen-Loève que se hace de la intensidad  $\lambda(t)$ .

Nosotros hemos calculado también la función característica para un P.P.D.E. con media Normal truncada que se analizará en la sección 2.5.

### 2.4.2 Teoremas límite para los estadísticos de recuento

Debido a la dificultad de encontrar una expresión explícita para la probabilidad  $P[N(t) = n]$  vista en (2.11) o (2.13) para un  $t$  finito, es interesante caracterizarla asintóticamente para un  $t$  grande. Para esto veremos varios teoremas acerca de la distribución límite del P.P.D.E. cuando  $t$  tiende a infinito.

Notemos  $N_{cn}(t)$  y  $\Lambda_{cn}(t)$  a  $N(t)$  y  $\Lambda(t)$  centrados y normalizados, respectivamente, es decir:

$$N_{cn}(t) = \frac{N(t) - m_N(t)}{\sigma_N(t)} \quad \text{y} \quad \Lambda_{cn}(t) = \frac{\Lambda(t) - m_\Lambda(t)}{\sigma_\Lambda(t)}$$

donde  $m_N(t)$  y  $m_\Lambda(t)$  son las funciones del tiempo medias de  $N$  y  $\Lambda$ , así como  $\sigma_N(t)$  y  $\sigma_\Lambda(t)$  son las desviaciones típicas de los mismos.

**Teorema 2.4.7** *Supongamos  $m_\Lambda(t) \rightarrow \infty$  y  $\sigma_\Lambda^2(t)/m_\Lambda(t) \rightarrow c$  cuando  $t \rightarrow \infty$ , siendo  $c$  una constante real positiva. Si  $\Lambda_{cn}(t)$  converge en distribución hacia  $\xi$  cuando  $t \rightarrow \infty$  entonces,  $N_{cn}(t)$  converge en distribución a  $g(1+c)^{-1/2} + \xi(1+c^{-1})^{-1/2}$  cuando  $t \rightarrow \infty$ , siendo  $g$  una Normal standard e independiente de  $\xi$ .*

**Demostración.** Estudiemos el límite de la función característica

de  $N_{cn}(t)$ . En primer lugar veamos que

$$\begin{aligned}
 E[e^{iuN_{cn}(t)}] &= \exp\left[-iu\frac{m_N(t)}{\sigma_N(t)}\right] E[e^{iN(t)/\sigma_N(t)}] \\
 &= \exp\left[-iu\frac{m_N(t)}{\sigma_N(t)}\right] E[\exp((e^{iu/\sigma_N(t)} - 1)\Lambda(t))] \\
 &= \exp\left[-iu\frac{m_N(t)}{\sigma_N(t)} + m_\Lambda(t) (e^{iu/\sigma_N(t)} - 1)\right] \\
 &= E[\exp((e^{iu/\sigma_N(t)} - 1)\Lambda_{cn}(t)\sigma_N(t))]
 \end{aligned}$$

Usando (2.16) y el desarrollo en serie de Taylor en torno al cero de  $e^{iu/\sigma_N(t)}$ , se demuestra fácilmente que

$$\lim_{t \rightarrow \infty} \exp\left[-iu\frac{m_N(t)}{\sigma_N(t)} + m_\Lambda(t) (e^{iu/\sigma_N(t)} - 1)\right] = \exp\left[-\frac{u^2}{2(c+1)}\right] \quad (2.18)$$

Para poder introducir el límite dentro de la esperanza que conforma el segundo factor en el que hemos expresado la función característica

veamos que

$$\begin{aligned}
 & \left| \exp \left( (e^{iu/\sigma_N(t)} - 1) \Lambda_{cn}(t) \sigma_N(t) \right) \right| = \\
 & = \left| \exp \left( (e^{iu/\sigma_N(t)} - 1) \Lambda_{cn}(t) \sigma_N(t) \right) \right| \\
 & = \left| \exp \left\{ \left[ \cos \left( \frac{u}{\sigma_N(t)} \right) + i \operatorname{sen} \left( \frac{u}{\sigma_N(t)} \right) - 1 \right] (\Lambda(t) - m_\Lambda(t)) \right\} \right| \\
 & = \left| \exp \left[ \Lambda(t) \left( \cos \left( \frac{u}{\sigma_N(t)} \right) - 1 \right) \right] \cdot \right. \\
 & \cdot \left. \exp \left[ -m_\Lambda(t) \left( \cos \left( \frac{u}{\sigma_N(t)} \right) - 1 \right) \right] e^{i(\Lambda(t) - m_\Lambda(t)) \operatorname{sen} \left( \frac{u}{\sigma_N(t)} \right)} \right| \\
 & \leq \exp \left[ -m_\Lambda(t) \left( \cos \left( \frac{u}{\sigma_N(t)} \right) - 1 \right) \right] \leq \\
 & \leq \exp \left[ \frac{m_\Lambda(t) u^2}{2 \sigma_N^2(t)} \right] \leq e^{\frac{u^2}{2}} < \infty
 \end{aligned}$$

para lo que hemos utilizado que  $\Lambda(t) \geq 0$ ,  $\cos \left( \frac{u}{\sigma_N(t)} \right) - 1 \leq 0$ , la exponencial de un complejo puro está acotada por 1, que  $1 - \cos x \leq \frac{x^2}{2}$  y que de (2.16),  $\frac{m_\Lambda(t)}{\sigma_N^2(t)} \leq 1$ . Por tanto, y gracias a esta acotación independiente de  $t$ , podemos aplicar el teorema de convergencia acotada generalizado, por lo que

$$\begin{aligned}
 & \lim_{t \rightarrow \infty} E \left\{ \exp \left( (e^{iu/\sigma_N(t)} - 1) \Lambda_{cn}(t) \sigma_N(t) \right) \right\} = \\
 & E \left\{ \lim_{t \rightarrow \infty} \left[ \exp \left( (e^{iu/\sigma_N(t)} - 1) \Lambda_{cn}(t) \sigma_N(t) \right) \right] \right\} = \\
 & E \left\{ \exp \left[ iu \frac{\xi}{\sqrt{1+c^{-1}}} \right] \right\} \quad (2.19)
 \end{aligned}$$

y junto con el límite calculado en (2.18) se prueba el teorema. ■

La demostración de este teorema se basa en la de Grandell (1971)



hecha para el caso en que la intensidad es un proceso estacionario y la encontrada en Snyder and Miller (1991).

**Teorema 2.4.8** *Supongamos las convergencias,  $m_\Lambda(t) \rightarrow \infty$  y que  $\sigma(t)_\Lambda^2(t)/m_\Lambda(t) \rightarrow 0$  cuando  $t \rightarrow \infty$ . Entonces,  $N_{cn}(t)$  tiende en distribución a  $g$  cuando  $t \rightarrow \infty$ , siendo  $g$  una Normal standard.*

**Demostración.** La demostración de este teorema es igual al anterior excepto que el límite en (2.18) es en este caso  $\exp[-\frac{1}{2}u^2]$  y el límite (2.19) resulta ser 1. ■

**Teorema 2.4.9** *Supongamos que se cumple que  $m_\Lambda(t) \rightarrow \infty$  y que  $\sigma(t)_\Lambda^2(t)/m_\Lambda(t) \rightarrow \infty$  cuando  $t \rightarrow \infty$ . Si  $\Lambda_{cn}(t)$  converge en distribución hacia  $\xi$  cuando  $t \rightarrow \infty$  entonces,  $N_{cn}(t)$  también converge en distribución a  $\xi$  cuando  $t \rightarrow \infty$ .*

**Demostración.** De nuevo, la demostración de este teorema es igual al anterior excepto que el límite en (2.18) es ahora 1 y el límite (2.19) resulta ser  $E[e^{iu\xi}]$ . ■

### 2.4.3 Estadísticos de tiempo

El método de condicionamiento, es decir, la propiedad de que un P.P.D.E.  $\{N(t) : t \geq t_0\}$  es un P.P. condicionando al proceso información  $\{x(t) : t \geq t_0\}$ , que usamos ya para calcular la función masa de probabilidad y la función característica del P.P.D.E., también nos servirá para obtener una expresión de los estadísticos de tiempo.

Llamando  $p_W^{(n)}(w)$  a la función de densidad conjunta de los  $n$  primeros tiempos de ocurrencia  $w_1, \dots, w_n$  y por el método de condicionamiento y (1.19); ésta tiene la forma

$$p_W^{(n)}(w) = E \left\{ \prod_{i=1}^n \lambda(w_i, x(w_i)) \exp \left[ - \int_{t_0}^{w_n} \lambda(\sigma, x(\sigma)) d\sigma \right] \right\}$$

con  $t_0 \leq w_1 \leq w_2 \leq \dots \leq w_n$ .

De la misma manera, con el método de condicionamiento y (1.25), la función de densidad muestral es la siguiente

$$p[\{N(t) : t_0 \leq \sigma < T\}] = E \left\{ \exp \left[ - \int_{t_0}^T \lambda(\sigma, x(\sigma)) d\sigma + \int_{t_0}^T L_n \lambda(\sigma, x(\sigma)) N(d\sigma) \right] \right\} \quad (2.20)$$

donde la segunda integral es una integral de recuento.

#### 2.4.4 El P.P.D.E. es una semimartingala

Abordamos en la sección 1.4, la visión del proceso de Poisson según la teoría de martingalas. Pretendemos ahora hacer lo mismo pero para el P.P.D.E.; no sólo porque es la interpretación más actual de los procesos de recuento si no porque permite tratarlos con relativa facilidad. Podemos ver que la extensión de procesos puntuales a procesos puntuales marcados se hace con no demasiado esfuerzo extra con la herramienta de la teoría de martingalas (Last y Brandt, 1995) o a procesos de recuento con intensidad aleatoria (Daley y Vere-Jones (1988) y Andersen et al. (1993)). Incluso se estudia con éxito el problema de filtrado no lineal de un proceso de recuento gracias a esta teoría, aún en casos complicados como aquel en el que la intensidad de un proceso de Poisson sea una difusión cualquiera (Boel y Beneš, 1980), campo de estudio que excede a esta memoria.

De hecho, la demostración de que el P.P.D.E. es una semimartingala se resuelve como extensión inmediata de la demostración de que lo sea un proceso de Poisson no homogéneo, (ver proposición 1.4.6, su discusión y el enunciado del teorema 1.4.7). Si  $\{X(t) : t \geq t_0\}$  es el proceso información del P.P.D.E., bastaría haber tomado las  $\sigma$ -álgebras  $\mathcal{F}_t$  de la historia  $\mathcal{F}$  de la forma

$$\mathcal{F}_t = \mathcal{H}_t^N \vee \mathcal{H}_t^X$$

donde en el segundo miembro hemos denotado de esa manera a las  $\sigma$ -álgebras de las historias internas de los procesos  $\{N(t) : t \geq t_0\}$  y  $\{X(t) : t \geq t_0\}$ .

## 2.5 P.P.D.E. con media aleatoria Normal truncada

Como sabemos,  $E[N(t)] = \Lambda(t, x(t))$ . Nosotros hemos considerado que hay muchos procesos de recuento de Poisson cuya media podría modelizarse mediante una Normal truncada en cada instante de tiempo. Por lo que en esta sección nos vamos a dedicar a estudiar en más profundidad el caso particular de P.P.D.E. con función paramétrica o media aleatoria que se distribuye como una Normal truncada para cada  $t, t \geq t_0$ .

### 2.5.1 Función masa de probabilidad

Recordando la expresión de que el P.P.D.E. tome un cierto valor que vimos en (2.13)

$$P[N(t) = n] = \frac{1}{n!} G_{\Lambda(t, x(t))}^n(-1) \equiv \frac{1}{n!} G_{\Lambda}^n(-1) \quad (2.21)$$

donde  $G_{\Lambda(t, x(t))}(s)$  es la función generatriz de momentos de  $\Lambda(t, x(t))$  y que notaremos  $G_{\Lambda}(s)$  a partir de ahora, pues ya se sabe que la función paramétrica depende del tiempo y el proceso información y resulta una notación más clara.

Si  $\Lambda(t, x(t))$  es una Normal truncada para cada  $t$ , su función de densidad dada en Johnson y Kotz (1970) es:



$$\begin{aligned}
 f(\Lambda(t, x(t))) &\equiv f(\Lambda) \\
 &= \left( \int_A^B e^{-\frac{1}{2} \left( \frac{y-\mu}{\sigma} \right)^2} dy \right)^{-1} \exp \left[ -\frac{1}{2} \left( \frac{\Lambda - \mu}{\sigma} \right)^2 \right], A < \Lambda < B
 \end{aligned}
 \tag{2.22}$$

Puede también expresarse

$$\begin{aligned}
 f(\Lambda(t, x(t))) &\equiv f(\Lambda) \\
 &= \frac{1}{\sigma\sqrt{2\pi}} \left( P \left[ \frac{A - \mu}{\sigma} < Z < \frac{B - \mu}{\sigma} \right] \right)^{-1} \exp \left[ -\frac{1}{2} \left( \frac{\Lambda - \mu}{\sigma} \right)^2 \right]
 \end{aligned}
 \tag{2.23}$$

donde  $A$  y  $B$  son los puntos inferior y superior, respectivamente, de truncamiento, con  $A \geq 0$  para que la función paramétrica sea no negativa, y  $\mu \in \mathbb{R}$  y  $\sigma \geq 0$ . Observemos que también  $A, B, \mu$  y  $\sigma$  son funciones de  $t$ ,  $A(t), B(t), \mu(t)$  y  $\sigma(t)$ , lo que hemos omitido en la expresión de la función de densidad para no complicar excesivamente la notación. De ahora en adelante, notaremos la Normal truncada en el instante  $t$ , por  $N_T(A(t), B(t), \mu(t), \sigma(t)) \equiv \aleph_T(A, B, \mu, \sigma)$ .

Usando que  $G_\Lambda^n(s) = E[\Lambda^n e^{s\Lambda}]$  y (2.23), fácilmente deducimos que la derivada  $n$ -ésima de la función generatriz de momentos de  $\Lambda$  tiene la forma

$$G_\Lambda^n(s) = \frac{1}{\sigma\sqrt{2\pi}} \left( P \left[ \frac{A - \mu}{\sigma} < Z < \frac{B - \mu}{\sigma} \right] \right)^{-1} \int_A^B \Lambda^n e^{s\Lambda} e^{-\frac{1}{2} \left( \frac{\Lambda - \mu}{\sigma} \right)^2} d\Lambda
 \tag{2.24}$$

donde  $Z$  es la Normal estándar  $N(0, 1)$ .

Evaluando la expresión (2.24) en  $-1$  y teniendo en cuenta (2.21), el estadístico de recuento para un P.P.D.E. con media aleatoria Normal truncada resulta ser

$$P[N(t) = n] = \frac{1}{n! \sigma \sqrt{2\pi}} \left( P \left[ \frac{A - \mu}{\sigma} < Z < \frac{B - \mu}{\sigma} \right] \right)^{-1} \int_A^B \Lambda^n e^{-\Lambda} e^{-\frac{1}{2} \left( \frac{\Lambda - \mu}{\sigma} \right)^2} d\Lambda \quad (2.25)$$

para  $n = 0, 1, 2, \dots$  donde  $0 < A < \mu < B$ , todas ellas funciones del tiempo como sabemos.

### 2.5.2 Función característica

Para calcular la función característica del P.P.D.E. con media una Normal truncada también necesitamos calcular, tal como vimos en (2.15), la función característica de  $\Lambda$ , que en este caso es una Normal truncada  $N_T(A, B, \mu, \sigma)$ .

$$\begin{aligned} M_\Lambda(iu) &= \int_A^B e^{iu\Lambda} f(\Lambda) d\Lambda \\ &= P \left[ \frac{A - \mu - iu\sigma^2}{\sigma} < Z < \frac{B - \mu - iu\sigma^2}{\sigma} \right] \cdot \\ &\quad \cdot \left( P \left[ \frac{A - \mu}{\sigma} < Z < \frac{B - \mu}{\sigma} \right] \right)^{-1} \exp \left[ iu\mu - \frac{1}{2} u^2 \sigma^2 \right] \quad (2.26) \end{aligned}$$

donde  $Z$  es una variable aleatoria que se distribuye mediante una Normal standard.

Volviendo al P.P.D.E. del que nos ocupamos, y usando (2.26), su

función característica es por tanto de la forma

$$M_{N(t)}(iu) = M_{\Lambda(t)}(e^{iu} - 1) = \frac{P[A - i(e^{iu} - 1)\sigma^2 < \xi < B - i(e^{iu} - 1)\sigma^2]}{P[A < \xi < B]} \cdot \exp\left[\mu(e^{iu} - 1) - \frac{1}{2}(e^{iu} - 1)^2\sigma^2\right]$$

donde  $\xi$  es una variable con distribución  $N(\mu(t), \sigma(t))$ . Por facilitar la notación, notaremos la fracción que aparece en el segundo miembro de la ecuación anterior como  $C(iu, t)$ . Observando además la exponencial que aparece en ese mismo miembro de esa expresión de la función característica, se puede escribir abreviadamente

$$M_{N(t)}(iu) = C(iu, t) M_{\xi}(e^{iu} - 1) \quad (2.27)$$

También se puede interpretar la función característica del proceso vista en (2.17) como el producto de  $C(iu, t)$  por la función característica de un proceso de Poisson  $\{N^*(t); t \geq t_0\}$  con intensidad gaussiana  $\{\lambda^*(t); t \geq t_0\}$  tal que

$$E[\lambda^*(t)] = \mu'(t) \text{ y } \int_{t_0}^t \int_{t_0}^t R(u, v) du dv = \sigma^2(t)$$

donde  $R(u, v)$  es la función de covarianza de  $\lambda^*(t)$ , ver el artículo de Valderrama et al. (1995), por lo que tendríamos

$$M_{N(t)}(iu) = C(iu, t) M_{N^*(t)}(iu) \quad (2.28)$$

Ambas formas de escribir la función característica del P.P.D.E. con media Normal truncada, nos ayudan notablemente a la hora de hacer los cálculos para hallar la media y varianza del proceso, que serían imposibles de realizar usando sus correspondientes definiciones,

puesto que los momentos de orden 1 y 2 de la variable  $\xi$  son ya por todos conocidos, así como los del proceso  $N^*(t)$  calculados en el citado artículo.

Usando (2.27), tenemos que la esperanza se calcularía

$$\begin{aligned} E[N(t)] &= E[\Lambda(t)] = \left. \frac{\partial M_{N(t)}(iu)}{\partial iu} \right|_{i0} = \left. \frac{\partial M_{\Lambda(t)}(e^{iu} - 1)}{\partial iu} \right|_{i0} \\ &= \left. \frac{\partial C(iu, t)}{\partial iu} \right|_{i0} M_{\xi}(e^{i0} - 1) + C(i0, t) \left. \frac{\partial M_{\xi}(e^{iu} - 1)}{\partial iu} \right|_{i0} \end{aligned} \quad (2.29)$$

y para la varianza del proceso necesitamos el momento de segundo orden

$$\begin{aligned} E[N^2(t)] &= \left. \frac{\partial^2 M_{N(t)}(iu)}{\partial iu^2} \right|_{i0} = \\ &= \left. \frac{\partial^2 C(iu, t)}{\partial iu^2} \right|_{i0} M_{\xi}(e^{i0} - 1) + 2 \left. \frac{\partial C(iu, t)}{\partial iu} \right|_{i0} \left. \frac{\partial M_{\xi}(e^{iu} - 1)}{\partial iu} \right|_{i0} + \\ &+ C(i0, t) \left. \frac{\partial^2 M_{\xi}(e^{iu} - 1)}{\partial iu^2} \right|_{i0} \end{aligned}$$

Puesto que  $\xi \rightsquigarrow N(\mu, \sigma)$ , es fácil observar que

$$M_{\xi}(e^{i0} - 1) = 1; \quad \left. \frac{\partial M_{\xi}(e^{iu} - 1)}{\partial iu} \right|_{i0} = E[\xi] \quad \text{y} \quad \left. \frac{\partial^2 M_{\xi}(e^{iu} - 1)}{\partial iu^2} \right|_{i0} = E[\xi^2] \quad (2.30)$$

se comprueba también que

$$\begin{aligned} C(i0, t) &= 1; \quad \left. \frac{\partial C(iu, t)}{\partial iu} \right|_{i0} = \sigma \frac{e^{-\frac{1}{2}(\frac{B-\mu}{\sigma})^2} - e^{-\frac{1}{2}(\frac{A-\mu}{\sigma})^2}}{P[A < \xi < B]} \quad \text{y} \\ \left. \frac{\partial^2 C(iu, t)}{\partial iu^2} \right|_{i0} &= \frac{\frac{B-\mu}{2\sigma} e^{-\frac{1}{2}(\frac{B-\mu}{\sigma})^2} - \frac{A-\mu}{2\sigma} e^{-\frac{1}{2}(\frac{A-\mu}{\sigma})^2}}{P[A < \xi < B]} + \sigma \frac{e^{-\frac{1}{2}(\frac{B-\mu}{\sigma})^2} - e^{-\frac{1}{2}(\frac{A-\mu}{\sigma})^2}}{P[A < \xi < B]} \end{aligned} \quad (2.31)$$

Tenemos ya los cálculos necesarios para obtener la expresión de la esperanza y la varianza. Que aunque los hemos hecho para la descomposición de la función característica del proceso como la función  $C(iu, t)$  por la función característica de  $\xi$  en  $e^{iu} - 1$ , resultaría igual haber tomado la interpretación que también hacíamos anteriormente de la misma función característica del proceso como  $C(iu, t)$  por la función característica del P.P.D.E.  $N^*(t)$  con intensidad gaussiana  $\lambda^*(t)$ . Queda, por tanto, que la esperanza y la varianza de un P.P.D.E. con media aleatoria Normal truncada son los siguientes:

$$E[N(t)] = \sigma \frac{f(B) - f(A)}{P[A < \xi < B]} + \mu \quad (2.32)$$

y

$$\begin{aligned} \text{Var}[N(t)] = & \quad (2.33) \\ & \frac{1}{2\sigma} \frac{(B - \mu) f(B) - (B - \mu) f(A)}{P[A < \xi < B]} + \frac{f(B) - f(A)}{P[A < \xi < B]} [\sigma - 2\mu\sigma + 2\mu] + \sigma^2 \end{aligned}$$

donde  $\xi$  se distribuye como una  $N(\mu, \sigma)$  y  $f(x) = \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]$ . Recordamos de nuevo que  $A, B, \mu$  y  $\sigma$  son todas ellas funciones del tiempo.

### 2.5.3 Estimación de la moda

Para cualquier instante de tiempo dado,  $t = t_1$ , los parámetros de la Normal truncada serían constantes y así, la probabilidad  $P[N(t_1) = n]$  (ver 2.25) sería únicamente una función de  $n$ ; llamémosla  $\varphi(n)$ .

Intentaremos encontrar el entero con máxima probabilidad en el instante  $t_1$  para el P.P.D.E. con media Normal truncada. El problema que surge es que no podemos encontrar los extremos de  $\varphi(n)$  derivando e igualando a cero como si la función de  $n$  fuese continua, pues no lo es. En la siguiente proposición (Rodríguez-Bouzas et al., 1998a) intentamos encontrar una solución a este problema.



**Proposición 2.5.1** *El entero para el cual la probabilidad  $P[N(t_1) = n]$ ,  $n = 0, 1, 2, \dots$  es máxima se encuentra en el intervalo acotado  $[A + 1, B - 1]$ .*

**Demostración.** *Dado un momento de tiempo fijo  $t = t_1$ , la ecuación (2.25) se puede escribir*

$$\varphi(n) = P[N(t_1) = n] = \frac{c}{n!} \int_A^B \Lambda^n e^{-\Lambda} e^{-\frac{1}{2} \left(\frac{\Lambda-\mu}{\sigma}\right)^2} d\Lambda,$$

$$A < \Lambda < B \text{ y } c = \text{cte.}$$

Sea la función  $f(n) = \frac{\Lambda^n}{n!} e^{-\Lambda} e^{-\frac{1}{2} \left(\frac{\Lambda-\mu}{\sigma}\right)^2}$ ,  $A < \Lambda < B$ . Maximizar  $\varphi(n)$  es lo mismo que maximizar  $f(n)$  puesto que difieren en una constante y la integral conserva la monotonicidad.

Como  $e^{-\Lambda} e^{-\frac{1}{2} \left(\frac{\Lambda-\mu}{\sigma}\right)^2}$  no depende de  $n$ , estudiemos el comportamiento de  $\frac{\Lambda^n}{n!}$ .

Observando que  $e^\Lambda = \sum_{n=0}^{\infty} \frac{\Lambda^n}{n!}$ ,  $\forall \Lambda \in \mathbb{R}$ , sabemos entonces que  $\frac{\Lambda^n}{n!} \rightarrow 0$  cuando  $n \rightarrow \infty$ . Por lo que existe un  $n_0$  tal que  $\forall n \geq n_0$ ,  $\frac{\Lambda^n}{n!}$  es decreciente en  $n$ ,  $\forall \Lambda \in \mathbb{R}$ . Tomando un  $\Lambda \in [A, B]$  fijo, se verifica que

$$n \geq n_0(\Lambda) : \frac{\Lambda^n}{n!} \geq \frac{\Lambda^{n+1}}{(n+1)!} \text{ si y sólo si } n \geq \Lambda - 1,$$

por lo que el entero en el que  $f(n)$  alcanza su máximo para todo  $\Lambda \in [A, B]$  es menor o igual que  $B - 1$ .

También para un  $\Lambda$  fijo,  $A < \Lambda < B$  se verifica que

$$\frac{\Lambda^n}{n!} \leq \frac{\Lambda^{n+1}}{(n+1)!} \text{ si y sólo si } n \leq \Lambda + 1$$

así,  $\frac{\Lambda^n}{n!}$  es creciente para  $n \leq \Lambda - 1$  y en consecuencia, para todo  $\Lambda \in [A, B]$ , el entero en el que  $f(n)$  alcanza su máximo es mayor o igual que  $A + 1$ .

Así pues, queda probado que el entero en el que  $\varphi(n)$  es máxima se encuentra en el intervalo acotado  $[A + 1, B - 1]$ . ■

La importancia de esta proposición es que permite buscar el valor más probable del proceso de Poisson en el instante  $t_1$  evaluando solamente  $P[N(t_1) = n]$  en los enteros del intervalo  $[A + 1, B - 1]$  aunque esté definida para todo  $n = 0, 1, 2, \dots$

Teniendo un proceso de recuento real, conocido en  $[t_0, T]$ , podemos intentar modelizarlo mediante un P.P.D.E. y si creemos que su media es aleatoria y se distribuye mediante una Normal truncada para cada instante de tiempo, podemos encontrar su valor más probable en un instante de tiempo futuro  $t_1$  usando la proposición anterior. Los pasos que deberíamos seguir son:

- 1.- estimar los parámetros de la Normal truncada con la que se distribuye la media del P.P.D.E., para cada  $t$ ,  $t_0 \leq t \leq T$ ;
- 2.- comprobar que la modelización de la media de los datos por esas Normales truncadas es adecuada;
- 3.- estudiar la evolución de los parámetros en el tiempo,  $t$ , y así encontrar la Normal truncada adecuada para el instante  $t_1$ ;
- 4.- el último paso consistiría en encontrar el entero para el cual  $P[N(t_1) = n]$  es máxima.

Seguir estos pasos no es fácil en la práctica; analicemos uno por uno.

- 1.- Si disponemos de varias trayectorias del proceso y los datos de las medias de ocurrencias para cada  $t$ ,  $t_0 \leq t \leq T$ ; entonces tendremos una muestra  $\Lambda_1, \dots, \Lambda_r$  para cada  $t$  con lo que estimaríamos la Normal truncada,  $N_T(A(t), B(t), \mu(t), \sigma(t))$ , para cada  $t$ .

En el artículo de Mittal y Dahiya (1987), se demuestra que los estimadores máximo verosímiles (EMV) para una Normal truncada son infinitos con probabilidad positiva. En este artículo se propone otro método de estimación, la estimación modal bayesiana que proporciona los *estimadores máximo verosímiles modificados (EMVM)* demostrándose que existen con probabilidad uno y son mejores que los EMV.

Tomando una chi-cuadrado como densidad a priori para  $\frac{1}{\sigma^2}$ ,

$$f(\sigma^2) = f(\theta) = c(\nu) \theta^{-\frac{\nu-2}{2}} \exp\left(\frac{-1}{2\theta}\right), \theta > 0$$

donde  $c(\nu) = 2^{-\frac{\nu}{2}} [\Gamma(\frac{\nu}{2})]^{-1}$  y sin información a priori para  $\mu$ , la función de verosimilitud modificada es ahora:

$$L_m(\mu, \sigma / \Lambda_1, \dots, \Lambda_r) = \frac{c_2 \sigma^{2-\nu} \exp[-(\sum_{i=1}^r (\Lambda_i - \mu) + 1) / 2\sigma^2]}{\left[ \int_A^B \exp(- (t - \mu)^2 / 2\sigma^2) \right]^r}$$

donde  $c_2$  es una función de  $\nu$ .

Así, los EMVM de la Normal truncada para un instante  $t$  verifican el siguiente sistema de ecuaciones:

$$\begin{aligned} \bar{\Lambda} &= \frac{\int_A^B y \exp[-\frac{1}{2} (\frac{y-\mu}{\sigma})^2] dy}{\int_A^B \exp[-\frac{1}{2} (\frac{y-\mu}{\sigma})^2] dy} \\ S^2 + \frac{1}{r} &= \frac{2\sigma^2}{r} + \frac{\int_A^B y^2 \exp[-\frac{1}{2} (\frac{y-\mu}{\sigma})^2] dy}{\int_A^B \exp[-\frac{1}{2} (\frac{y-\mu}{\sigma})^2] dy} - \\ &\quad - \left( \frac{\int_A^B y \exp[-\frac{1}{2} (\frac{y-\mu}{\sigma})^2] dy}{\int_A^B \exp[-\frac{1}{2} (\frac{y-\mu}{\sigma})^2] dy} \right)^2 \end{aligned}$$

donde  $\bar{\Lambda}$  y  $S$  son la media y desviación standard muestrales, respectivamente, y  $r$  el tamaño muestral; para  $\nu$ , los autores

proponen como valor óptimo  $\nu = 4$  por lo que nosotros también lo usamos.

La solución de este sistema de ecuaciones no lineales lo hemos implementado usando el método de Newton-Raphson. Aplicándolo a los datos para cada instante de tiempo del que tenemos información  $t, t_0 \leq t \leq T$ , podrán estimarse las Normales truncadas para cada uno de ellos.

- 2.- Ha sido también implementado un test de bondad de ajuste de Kolmogorov-Smirnov para la distribución Normal truncada, que aplicado a los datos para todo  $t, t_0 \leq t \leq T$ , nos da información de la bondad de los ajustes de las Normales truncadas estimadas en el paso anterior.
- 3.- Una vez aceptado, en el paso 2, que  $\Lambda(t)$  es una Normal truncada particular para cada  $t, t_0 \leq t \leq T$ , lo que tenemos es un valor de  $A, B, \mu$  y  $\sigma$  para cada uno de los instantes en los que teníamos información. Debemos establecer ahora las funciones  $A(t), B(t), \mu(t)$  y  $\sigma(t)$ . Después de esto, podrían calcularse los parámetros de  $\Lambda(t_1)$  y por tanto su distribución.
- 4.- Habiendo completado los pasos 1, 2 y 3, y usando la proposición 2.5.1, el número de ocurrencias del P.P.D.E. con mayor probabilidad en el instante  $t_1$  es un entero perteneciente al intervalo  $[A(t_1) + 1, B(t_1) - 1]$ . Hemos implementado también la búsqueda de ese valor del punto de máxima probabilidad del proceso de Poisson que estamos estudiando.

Los cuatro pasos han sido implementados en notebooks de *Mathematica 3.0* haciendo uso de Ramírez et al. (1997) debido a que usa el método adaptativo gaussiano para calcular numéricamente las integrales que aparecen en la función de densidad de la media del P.P.D.E.,  $\Lambda(t)$ , y por tanto en la función masa de probabilidad, el

sistema de ecuaciones no lineales, etc. Los programas se han incluido en el Apéndice B de esta memoria.

Estos cuatro pasos conforman un método de estimación del valor del P.P.D.E., es decir de su moda, en el caso de que la media esté influenciada por el proceso información de tal forma que su distribución sea una Normal truncada en cada instante de tiempo. En caso de que la función paramétrica tuviese otra distribución, los pasos generales podrían ser los mismos aunque el test fuese para esa otra distribución, las ecuaciones de los estimadores fuesen las adecuadas a esa distribución, etc.; siempre y cuando se tenga información de dónde buscar el valor de la moda dentro de los infinitos puntos posibles que puede tomar el P.P.D.E., como nosotros conseguimos para nuestro caso en la proposición 2.5.1.

#### **2.5.4 Aplicación: número de efectos de comercio devueltos o impagados**

Pretendemos estimar el número de efectos bancarios devueltos o impagados por cada 10000 habitantes con mayor probabilidad en 1997. Poseemos los datos de este tipo de efectos bancarios desde 1989 a 1996 en diferentes provincias españolas. La media del número de tales efectos fue calculada en quince provincias escogidas aleatoriamente, para cada uno de los años de los que se disponía de información.

Las correspondientes Normales truncadas para esos años se estimaron usando la implementación de los EMVM vistos en el paso en el primer paso del método de estimación propuesto ( Rodríguez-Bouzas et al., 1998b) y se aceptó su bondad de ajuste usando el programa implementado para el test de Kolmogorov-Smirnov del que hablábamos en el paso 2.

Las medias por año por provincia y las distribuciones estimadas son las siguientes:

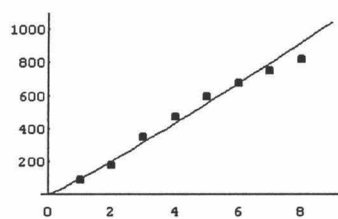
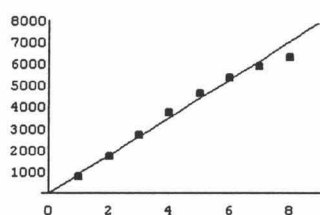
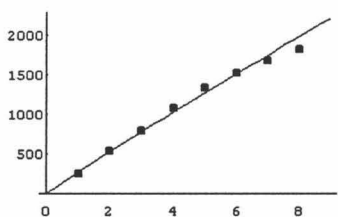
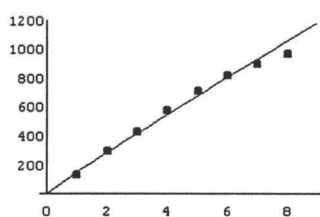
Provin.	1989	1990	1991	1992	1993	1994	1995	1996
Barce.	827.4	1763.6	2742.9	3747.0	4673.7	5364.0	5872.6	6316.5
Cant.	216.4	421.2	657.5	901.24	1137.8	1315.5	1462.4	1585.4
Cast.	366.2	799.4	1222.3	1728.5	2144.1	2440.7	2698.7	2962.2
Coru.	150.8	307.1	472.2	652.79	827.8	960.3	1066.0	1155.3
Grana.	208.5	413.5	722.2	924.6	1127.6	1289.6	1416.8	1537.7
Guada.	90.6	181.9	356.7	473.4	599.4	678.9	753.7	825.9
Huel.	130.3	281.1	413.0	534.7	639.9	714.8	770.1	820.0
León	178.5	343.7	487.7	640.2	784.6	891.4	977.0	1061.3
Lugo	108.4	225.8	352.5	487.8	623.9	710.1	773.0	827.6
Murcia	356.1	738.5	1110.2	1478.3	1803.8	2020.0	2180.6	2327.3
Nava.	405.2	974.4	1458.3	1718.4	2054.2	2310.6	2517.6	2704.1
Rioja	477.9	957.0	1418.2	1882.8	2318.1	2654.0	2942.8	3208.5
Toledo	201.7	412.9	621.6	835.1	1058.4	1228.2	1373.6	1500.1
Valla.	181.7	365.6	545.3	736.8	916.8	1056.8	1180.6	1291.7
Zara.	446.5	932.4	1414.3	1929.4	2373.5	2733.1	3032.9	3290.4

Medias por año para cada provincia

Normales truncadas estimadas:

- 1989 :  $N_T(90.6522, 827.44077, 259.927, 139.368)$
- 1990 :  $N_T(181.996, 1763.62, 542.605, 299.65)$
- 1991 :  $N_T(352.54631, 2742.9004, 812.924, 431.244)$
- 1992 :  $N_T(473.43229, 3747.06821, 1089.49, 582.483)$
- 1993 :  $N_T(599.45003, 4673.71963, 1343.29, 715.946)$
- 1994 :  $N_T(678.984577, 5364.02494, 1534.57, 820.848)$
- 1995 :  $N_T(753.741017, 5872.61065, 1691.77, 901.551)$
- 1996 :  $N_T(820.090958, 6316.50665, 1828.14, 971.78)$

Debemos ahora establecer una expresión para la Normal truncada general  $N_T(A(t), B(t), \mu(t), \sigma(t))$ ; es decir, la distribución de la media,  $\Lambda(t)$ , del P.P.D.E., como decíamos en el paso 3, con objeto de poder calcular después la Normal truncada correspondiente a  $\Lambda(t)$  con  $t$  un instante de tiempo futuro. Para estimar los parámetros de dicha Normal truncada se ha hecho regresión potencial mínimo-cuadrática para cada uno de los parámetros funciones del tiempo, cuyos gráficos de los valores "observados" y predichos son:

parâmetro  $A(t)$ parâmetro  $B(t)$ parâmetro  $\mu(t)$ parâmetro  $\sigma(t)$ 

La expresión funcional de la distribución ,  $N_T(A(t), B(t), \mu(t), \sigma(t))$ , estimada resultó ser

$$\begin{aligned} A(t) &= 94.049383 z^{1.09733}, \\ B(t) &= 879.58482 z^{0.996695}, \\ \mu(t) &= 275.98596 z^{0.949735} \text{ y} \\ \sigma(t) &= 148.18182 z^{0.943997} \end{aligned}$$

donde  $z = t - 1988$ .



Teniendo ya la expresión de la distribución general de  $\Lambda(t)$ , la media del proceso de Poisson doblemente estocástico que estamos estudiando, se distribuye para el año 1997, sin más que sustituir, mediante la Normal truncada  $N_T(1048.269, 7858.952, 2224.1487, 1179.2298)$ .

Por fin, queda solamente calcular el número más probable de efectos bancarios devueltos o impagados en 1997 que según la proposición 2.5.1 debe encontrarse en el intervalo  $[1049, 7857]$ . Usando el programa implementado que mencionábamos en el paso 4 de nuestro método, resultó ser 2223 con probabilidad 0.00040211147, con lo que alcanzamos la meta de nuestro estudio en esta sección.



## Capítulo 3

# Predicción de procesos de Poisson mediante ACP

En este capítulo se aplicará la técnica de reducción de dimensión Análisis en Componentes principales (ACP), en sus versiones multivariante y funcional, para predecir de forma continua procesos estocásticos con funciones muestrales escalonadas, como los de Poisson, a partir de su evolución en el pasado. Los modelos que se proponen son una adaptación para este tipo de procesos de los modelos de Regresión en Componentes Principales (ver, por ejemplo, Jackson (1991)) y de los modelos funcionales de Predicción en Componentes Principales (PCP) desarrollados por Aguilera et al. (1997).

### 3.1 Regresión multivariante en componentes principales

El análisis en componentes principales es una técnica de análisis multivariante que tiene como finalidad reducir la dimensión del problema mediante la construcción de transformaciones lineales de las variables

que son incorreladas y de varianza máxima.

La técnica de Regresión Multivariante en Componentes Principales (RMCP) resuelve el problema de estimar un conjunto de variables de respuesta representadas por un vector aleatorio  $\underline{Y} = (Y_1, \dots, Y_s)'$ , a partir de un conjunto de variables predictoras, denotadas por  $\underline{X} = (X_1, \dots, X_m)'$  ( $m, s \in \mathbb{Z}^+$ ). Gracias a la incorrelación de las componentes principales, esta técnica permite evitar el problema de multicolinealidad de la regresión lineal múltiple. Además, la reducción de dimensión que proporciona el ACP lleva a modelos extremadamente sencillos, de modo que si las componentes principales fuesen fácilmente interpretables las ecuaciones de regresión serían más significativas y fácilmente estimables.

La característica principal de esta técnica es que es necesario realizar un doble ACP, es decir, un ACP clásico para cada conjunto de variables.

Sean  $\{\xi_1, \dots, \xi_m\}$  las componentes principales asociadas al vector aleatorio  $\underline{X}$ , dadas por

$$\xi_i = \sum_{j=1}^m \phi_{ji}(X_j - \mu_{X_j}) = \underline{\phi}_i' (\underline{X} - \underline{\mu}_X), \quad (3.1)$$

donde  $\underline{\phi}_i$  es el  $i$ -ésimo vector propio de la matriz de covarianzas de  $\underline{X}$ , y  $\underline{\mu}_X$  es su vector de medias.

Análogamente, denotemos por  $\{\eta_1, \dots, \eta_s\}$  a las componentes principales (cc.pp.) asociadas al vector aleatorio  $\underline{Y}$ , y definidas por

$$\eta_k = \sum_{j=1}^s \delta_{jk}(Y_j - \mu_{Y_j}) = \underline{\delta}_k' (\underline{Y} - \underline{\mu}_Y), \quad (3.2)$$

donde  $\underline{\delta}_k$  es el  $k$ -ésimo vector propio de la matriz de covarianzas de  $\underline{Y}$  y  $\underline{\mu}_Y$  es el correspondiente vector de medias.

Entonces, un modelo RCPM para predecir cada variable respuesta  $Y_j$  ( $j = 1, \dots, s$ ) viene dado por

$$\tilde{Y}_j = \mu_{Y_j} + \sum_{k=1}^q \hat{\eta}_k \delta_{jk}, \quad q \leq s, \tag{3.3}$$

donde  $\hat{\eta}_k$  ( $k = 1, \dots, s$ ) es el estimador lineal de mínimos cuadrados para la c.p.  $\eta_k$  sobre un subconjunto de cc.pp.  $\{\xi_i\}_{i=1}^m$ , y  $\mu_{Y_j}$  es la media de la variable  $Y_j$ .

Si en la ecuación (3.3)  $q$  iguala a  $s$ , y todas las cc.pp.  $\{\xi_i\}_{i=1}^m$  son usadas para estimar  $\eta_k$ , entonces RMCP proporciona los mismos resultados que la regression lineal múltiple mínimo-cuadrática pero posiblemente con más precisión si la matriz de covarianzas de  $\underline{X}$  tiene problemas de invertibilidad.

### 3.2 Regresión funcional en componentes principales

Como extensión natural de los modelos de regresión multivariante en componentes principales, Aguilera *et al.* (1997) introdujeron los modelos de regresión funcional en componentes principales, utilizándolos para predecir series de tiempo generadas por procesos estocásticos en tiempo continuo (modelos de Predicción en Componentes Principales). Una de las principales ventajas de estos modelos es que proporcionan una predicción continua de un proceso estocástico en un intervalo de tiempo futuro a partir de observaciones discretas del proceso no necesariamente igualmente espaciadas en el pasado.

Conocida la evolución de un proceso estocástico  $\{X(t) : t \in [T_1, T_2]\}$ , los modelos de regresión funcional en componentes principales (RFCP) tienen como objetivo estimar otro proceso estocástico  $\{Y(s) : s \in$

$[T_3, T_4]$ }, haciendo uso del Análisis en Componentes Principales Funcional (ACPF) de cada proceso.

Consideraremos como hipótesis de partida que ambos procesos son de segundo orden, continuos en media cuadrática, y con funciones muestrales de cuadrado integrable en los intervalos de tiempo correspondientes.

Como extensión natural de la definición de Hötelling del ACP, Deville (1974) definió la *i*-ésima componente principal asociada al proceso  $\{X(t) : t \in [T_1, T_2]\}$  como la v.a.

$$\xi_i = \int_{T_1}^{T_2} (X(t) - \mu_X(t)) f_i(t) dt, \quad (3.4)$$

donde  $f_i$ , llamado *i*-ésimo factor principal, es la autofunción normalizada correspondiente al *i*-ésimo mayor autovalor  $\lambda_i$  de la función de covarianza  $C_X(t, s)$ , y  $\mu_X$  es la función media del proceso. Es decir,  $f_i$  y  $\lambda_i$  son las soluciones de la siguiente ecuación integral:

$$\int_{T_1}^{T_2} C_X(t, s) f_i(s) ds = \lambda_i f_i(t) \quad t \in [T_1, T_2]. \quad (3.5)$$

Las componentes principales así definidas tienen las mismas propiedades de optimalidad que en el caso finito. De hecho, la *i*-ésima c.p.,  $\xi_i$ , es una combinación lineal generalizada de las variables del proceso que es centrada y tiene varianza máxima,  $\lambda_i$ , de entre todas aquellas que son incorreladas con  $\{\xi_j\}_{j=1}^{i-1}$ . A la varianza  $\lambda_i$  se le llama *i*-ésimo valor principal y a  $f_i$  *i*-ésimo factor principal.

Un estudio más detallado sobre el ACPF en particular y sobre el contexto más general del Análisis de Datos Funcionales se puede ver en el libro de Ramsay y Silverman (1997).

Si denotamos por  $V^X$  a la varianza total del proceso  $X$  definida por

$$V^X = E \left[ \int_{T_1}^{T_2} (X(t) - \mu_X(t))^2 dt \right] = \sum_{i=1}^{\infty} \lambda_i < \infty, \quad (3.6)$$

la cantidad  $V_i^X = \lambda_i/V^X$  es llamada varianza explicada por la  $i$ -ésima componente principal.

Entonces, el proceso admite la siguiente descomposición en componentes principales, que es conocida en el contexto probabilístico como desarrollo de Karhunen-Loève:

$$X(t) - \mu_X(t) = \sum_i \xi_i f_i(t), \quad t \in [T_1, T_2], \quad (3.7)$$

donde la serie del segundo miembro converge uniformemente en media cuadrática en el intervalo  $[T_1, T_2]$ .

Además, la representación ortogonal del proceso en términos de sus componentes principales es óptima debido a que la serie (3.7) truncada en el  $q$ -ésimo término es el mejor modelo lineal de dimensión  $q$  para  $\{X(t)\}$  en el sentido de mínimos cuadrados (ver, por ejemplo, Fukunaga (1990)), de modo que  $\sum_{i=1}^q \lambda_i$  es la varianza explicada por dicho modelo, y  $\sum_{i=q+1}^\infty \lambda_i$  es el error cuadrático medio mínimo.

Análogamente, denotemos, respectivamente, por  $g_j$  y  $\eta_j$  a los factores y componentes principales asociados al proceso  $\{Y(s) : s \in [T_3, T_4]\}$ . La varianza total de  $Y$  viene dada por  $V^Y = \sum_j \alpha_j$ , donde  $\alpha_j$  es la varianza de la c.p.  $\eta_j$ , de modo que la varianza explicada por la  $j$ -ésima c.p.  $\eta_j$  viene dada por el cociente  $V_j^Y = \alpha_j/V^Y$ .

Entonces un modelo RFCP para estimar el proceso  $Y$  en  $[T_3, T_4]$  es de la forma

$$\tilde{Y}^q(s) = \mu_Y(s) + \sum_{j=1}^q \tilde{\eta}_j^{p_j} g_j(s), \quad s \in [T_3, T_4], \quad (3.8)$$

donde  $\tilde{\eta}_j^{p_j}$  es el estimador lineal de mínimos cuadrados para la c.p.  $\eta_j$  en términos de las  $p_j$  primeras cc.pp. del proceso  $X$  en  $[T_1, T_2]$ . Esto es,

$$\tilde{\eta}_j^{p_j} = \sum_{i=1}^{p_j} \frac{E[\eta_j \xi_i]}{\lambda_i} \xi_i = \sum_{i=1}^{p_j} \beta_i^j \xi_i. \quad (3.9)$$

Aguilera *et al.* (1997) han demostrado que la predicción lineal dada por el modelo (3.8) converge en media cuadrática al estimador lineal mínimo-cuadrático de  $Y(s)$  dadas las variables del proceso  $\{X(t) : t \in [T_1, T_2]\}$ , para cada  $s \in [T_3, T_4]$ . Finalmente, el modelo RFCP definido por (3.8) se suele denotar por RFCP  $(q; p_1, \dots, p_q)$ .

Cuando el proceso  $Y$  es el mismo proceso  $X$  en un intervalo  $[T_3, T_4]$  con  $T_3 > T_2$ , los modelos RFCP se adaptan para predecir el proceso en el intervalo futuro  $[T_3, T_4]$  a partir de su evolución en el pasado  $[T_1, T_2]$  y reciben el nombre de modelos de Predicción en Componentes Principales (PCP).

### 3.2.1 Estimación

Una vez formulados los modelos PCP nos planteamos su estimación a partir de la información proporcionada por  $N$  funciones muestrales independientes del proceso aleatorio  $\{X(t)\}$  en los intervalos pasado y futuro, a las que denotaremos, respectivamente, como sigue:

$$\{X_w(t) : t \in [T_1, T_2]\} \text{ y } \{X_w(s) : s \in [T_3, T_4]\} \quad w = 1, 2, \dots, N.$$

Esquemáticamente, los pasos a seguir para estimar los parámetros del modelo PCP presentado anteriormente son los siguientes:

1. Estimación de los factores y componentes principales muestrales en cada uno de los intervalos de tiempo considerados.

El estimador natural de  $C(t,s)$  es la función de covarianza muestral

$$\hat{C}(t, s) = \frac{1}{N-1} \sum_{w=1}^N (X_w(t) - \bar{X}(t)) (X_w(s) - \bar{X}(s)), \quad (3.10)$$

siendo  $\bar{X}$  el estimador insesgado de la media  $\mu$  definido por

$$\bar{X}(t) = \frac{1}{N} \sum_{w=1}^N X_w(t). \quad (3.11)$$

Como consecuencia de las propiedades de estos estimadores, que han sido estudiadas detalladamente por Deville (1973), los autovalores y autofunciones,  $(\lambda_i, f_i)$ , del núcleo  $C(t,s)$  se estiman mediante los correspondientes autovalores y autofunciones,  $(\hat{\lambda}_i, \hat{f}_i)$ , del núcleo  $\hat{C}(t,s)$ . Por lo tanto, los factores principales muestrales,  $\hat{f}_i$ , en el intervalo  $[T_1, T_2]$  son las soluciones de la siguiente ecuación integral de segundo orden:

$$\int_{T_1}^{T_2} \hat{C}(t,s) \hat{f}_i(s) ds = \hat{\lambda}_i \hat{f}_i(t), \quad t \in [T_1, T_2]. \quad (3.12)$$

Desafortunadamente, resolver la ecuación integral (3.12) es una tarea muy difícil que sólo tiene solución exacta para núcleos muy especiales reduciendo la ecuación integral a una ecuación diferencial. Teniendo en cuenta la naturaleza de las funciones muestrales del proceso, para las que en la práctica es usual disponer de observaciones discretas, en los últimos años han aparecido en la literatura estadística muchos trabajos que resuelven esta ecuación integral basándose en distintas aproximaciones numéricas. El método numérico más simple consiste en aproximar dicha ecuación integral mediante una fórmula de cuadratura compuesta. Aguilera *et al.* (1992) han aplicado la fórmula de cuadratura del trapecio obteniendo muy buenas aproximaciones de los factores principales en los nodos de la partición elegida. Un método más sofisticado es el de proyección ortogonal que consiste en aproximar los factores principales en un subespacio de dimensión finita y resulta particularmente adecuado cuando se



dispone de información a priori sobre la naturaleza de la solución exacta. Para el caso de procesos con trayectorias regulares, Aguilera *et al.* (1995) han resuelto el problema proyectando el proceso original sobre un subespacio finito dimensional de funciones trigonométricas. Además, ha sido contrastado mediante simulación (Aguilera *et al.* (1996a)) que una interpolación spline cúbica de las funciones muestrales proporciona aproximaciones óptimas de los factores principales en este caso. En el siguiente epígrafe se presentará un método de solución aproximada de esta ecuación para el caso de procesos cuyas funciones muestrales son constantes en intervalos aleatorios. Además, para estimar el ACP de un proceso mediante estos métodos de aproximación, se ha desarrollado el programa computacional PCAP, codificado en Turbo Pascal usando programación orientada a objeto y matrices dinámicas (Aguilera *et al.* (1994)).

El estimador natural de la varianza explicada por la  $i$ -ésima componente principal es el cociente  $\hat{\lambda}_i/\hat{V}_P$ , siendo  $\hat{V}_P = \sum_i \hat{\lambda}_i$  un estimador insesgado y consistente para la varianza total del proceso en el pasado  $V_P$ .

Una vez estimados los factores principales muestrales, la  $i$ -ésima c.p. muestral asociada al proceso en el intervalo  $[T_1, T_2]$  viene definida por

$$\hat{\xi}_i = \int_{T_1}^{T_2} (X(t) - \bar{X}(t)) \hat{f}_i(t) dt. \quad (3.13)$$

Una vez realizada la estimación del doble ACPF,  $\hat{f}_i$  y  $\hat{g}_i$  denotarán a los factores principales muestrales en los intervalos  $[T_1, T_2]$  y  $[T_3, T_4]$ , respectivamente. Análogamente,  $\hat{\xi}_{wi}$  y  $\hat{\eta}_{wi}$  denotan a las componentes principales muestrales en los intervalos  $[T_1, T_2]$  y  $[T_3, T_4]$ , respectivamente, para cada individuo  $w = 1, \dots, N$ .

2. Elección del número óptimo  $q$  de cc.pp.  $\hat{\eta}_j$  que serán introducidas en el modelo PCP como variables respuesta. Para ello el criterio más usual consiste en elegir un punto de truncamiento próximo a uno y retener las  $q$  primeras componentes principales tales que la proporción de varianza acumulada por ellas sea al menos el punto de truncamiento fijado.
3. Selección, para cada una de las cc.pp.  $\hat{\eta}_j$  ( $j = 1, \dots, q$ ) elegidas en el paso anterior, de las  $p_j$  componentes principales  $\hat{\xi}_i$  óptimas que serán introducidas como predictores en el modelo PCP. Aunque la práctica usual es borrar automáticamente como predictores aquellas cc.pp. asociadas a valores propios pequeños, está demostrado que no hay razón para que sean las cc.pp. con mayor varianza los mejores predictores. De hecho, puede ocurrir que alguna de las cc.pp. principales menos explicativas en el pasado estén áltamente correladas con las cc.pp. del futuro.

Por ello, nosotros elegiremos como los mejores predictores de cada una de las cc.pp.  $\hat{\eta}_j$  ( $j = 1, \dots, q$ ) aquellas  $p_j$  cc.pp.  $\hat{\xi}_i$  cuya correlación lineal sea significativa alta siguiendo un método de selección stepwise.

4. Estimación en la forma usual del modelo de regresión lineal de cada c.p.  $\hat{\eta}_j$  ( $j = 1, \dots, q$ ) sobre sus predictores  $\hat{\xi}_i$  ( $i = 1, \dots, p_j$ ),

$$\tilde{\eta}_j^{p_j} = \sum_{i=1}^{p_j} \hat{b}_i^j \hat{\xi}_i.$$

5. Estimación del modelo PCP seleccionado

$$\tilde{X}^q(s) = \bar{X}(s) + \sum_{j=1}^q \tilde{\eta}_j^{p_j} \hat{g}_j(s), \quad s \in [T_3, T_4]. \quad (3.14)$$

6. Para cada nueva función muestral  $X_w$  observada sólo en el intervalo  $[T_1, T_2]$ , se predice su evolución en el intervalo  $[T_3, T_4]$  sin más que estimar sus componentes principales en el pasado y sustituir en la ecuación (3.14).

### 3.2.2 Aproximación de estimadores en el ACP de procesos con funciones muestrales escalonadas

El punto de partida de este trabajo es el de los procesos de Poisson cuyas trayectorias permanecen constantes en intervalos aleatorios. En este caso, un método numérico eficiente para estimar de forma aproximada el ACPF consiste en aproximar los factores principales mediante los de la proyección de las funciones muestrales originales sobre el subespacio de las funciones constantes en los intervalos fijos de una partición previamente elegida en  $[T_1, T_2]$ . A continuación se presenta un breve resumen de este procedimiento de aproximación de los factores principales muestrales que ha sido desarrollado en el trabajo de Aguilera *et al.* (1996b).

Fijemos en el intervalo  $[T_1, T_2]$  una partición  $\pi_n$  definida por los nodos

$$T_1 = a_0 < a_1 < \dots < a_n = T_2$$

verificando

$$\Delta_n = \max_{j=1, \dots, n} \{(a_j - a_{j-1})\} \rightarrow 0 \text{ cuando } n \rightarrow \infty.$$

Sea  $E_n$  el subespacio de las funciones constantes sobre cada uno de los intervalos  $(a_{j-1}, a_j]$  ( $j = 1, \dots, n$ ). Una base ortonormal de  $E_n$  viene dada por las funciones

$$\delta_j(t) = (a_j - a_{j-1})^{-1/2} I_j(t),$$

siendo  $I_j$  la función indicadora en el intervalo  $(a_{j-1}, a_j]$ .

Para cada realización particular  $X_w(t)$  del proceso en la muestra, su proyección ortogonal sobre este subespacio de funciones constantes es de la forma

$$X_w^{(n)}(t) = P_n[X_w(t)] = \sum_{j=1}^n Y_{wj} \delta_j(t) = \sum_{j=1}^n M_{wj} I_j(t), \quad (3.15)$$

definiendo, para cada  $w = 1, \dots, N$  y  $j = 1, \dots, n$

$$Y_{wj} = \langle X_w | \delta_j \rangle = \int_{T_1}^{T_2} X_w(t) \delta_j(t) dt = (a_j - a_{j-1})^{1/2} M_{wj}. \quad (3.16)$$

siendo  $M_{wj}$  el valor medio de la trayectoria muestral  $w$  sobre el intervalo  $(a_{j-1}, a_j]$  definido por

$$M_{wj} = (a_j - a_{j-1})^{-1} \int_{a_{j-1}}^{a_j} X_w(t) dt.$$

Entonces, se demuestra que los valores principales muestrales del proceso proyectado  $X^{(n)}(t)$  son los valores propios de la matriz  $\mathbf{R}$  de dimensión  $n \times n$  con elementos

$$R_{ij} = \int_{T_1}^{T_2} \int_{T_1}^{T_2} \hat{C}(t, s) \delta_i(t) \delta_j(s) dt ds = \frac{1}{N-1} \sum_{w=1}^N (Y_{wi} - \bar{Y}_i)(Y_{wj} - \bar{Y}_j) \quad (3.17)$$

definiendo,

$$\bar{Y}_j = \frac{1}{N} \sum_{w=1}^N Y_{wj} = \int_{T_1}^{T_2} \bar{X}(t) \delta_j(t) ds. \quad (3.18)$$

Además, los factores principales muestrales son funciones escalonadas dadas por

$$\hat{f}_i^{(n)} = \sum_{j=1}^n z_{ji} \delta_j,$$

donde el vector columna  $\underline{z}_i$  es el  $i$ -ésimo vector propio de la matriz  $\mathbf{R}$  asociado a su  $i$ -ésimo valor propio  $\hat{\lambda}_i^{(n)}$ . Una vez obtenidos los factores principales aproximados  $\hat{f}_i^{(n)}$ , bajo la condición de normalización  $\sum_{j=1}^n (z_{ji})^2 = 1$ , las correspondientes componentes principales muestrales, a las que denotaremos  $\hat{\xi}_i^{(n)}$ , vienen dadas por

$$\hat{\xi}_{wi}^{(n)} = \sum_{j=1}^n z_{ji} (Y_{wj} - \bar{Y}_j), \quad w = 1, \dots, N. \quad (3.19)$$

Es decir, las componentes principales aproximadas son claramente las componentes principales asociadas a la matriz de datos  $\mathbf{Y}$  de dimensión  $N \times n$  con elementos  $Y_{wj}$  definidos en (3.16).

Observemos finalmente que la elección de la partición estará en función de la densidad de cambios de valor del proceso en el tiempo observado. De este modo los subintervalos no tienen que tener la misma amplitud y serán más finos en aquellos periodos en los que se producen más cambios.

### 3.3 Modelos RMCP y PCP para procesos de Poisson

En esta sección vamos a adaptar los modelos predictivos basados en ACP para predecir un proceso de Poisson en un periodo de tiempo de amplitud  $T$  a partir de su evolución anterior.

Sea  $N(t)$  un proceso de Poisson del que disponemos de los datos de una trayectoria  $n(t)$  observada en un periodo de tiempo  $(0, N \times T]$  con  $n(0) = 0$ . Para poder utilizar un modelo PCP para predecir  $n(t)$  en un intervalo de amplitud  $T$  de la forma  $(N \times T, (N + 1)T]$ , proponemos cortar la trayectoria  $n(t)$  en  $N - 1$  periodos de amplitud  $2T$  obteniendo así  $N - 1$  trayectorias del siguiente proceso

$$X_w(t) = \begin{cases} n((w - 1)T + t) - n((w - 1)T) & t \in (0, T] \\ n((w - 1)T + t) - n(wT) & t \in (T, 2T] \end{cases}$$

donde  $w = 1, \dots, N - 1$ .

Entonces, construimos un modelo PCP para estimar el proceso  $X$  en el intervalo futuro  $(T, 2T]$  a partir de su evolución en el intervalo pasado  $(0, T]$ , que será utilizado para estimar  $X_N(s)$  para todo  $s$  en  $(T, 2T]$  y como consecuencia, predecir  $n(t)$  en el intervalo de amplitud  $T$  de la forma  $(N \times T, (N + 1)T]$ .

En el caso en que no se disponga de los tiempos de ocurrencias de la trayectoria  $n(t)$  sino de un resumen del número acumulado de ocurrencias en los mismos instantes de tiempo  $0 \leq t_0 < t_1 < \dots < t_k \leq T$  para cada periodo de amplitud  $T$ , proponemos cortar la serie de la misma forma y utilizar los modelos RMCP para predecir la trayectoria  $X_N(s_j)$  ( $j = 0, \dots, k$ ) con  $s_j = t_j + T$ . De este modo se obtiene una predicción de la trayectoria  $n(t)$  en los nodos  $t_j + NT$  del intervalo futuro  $(N \times T, (N + 1)T]$ .

Con el objetivo de ilustrar la capacidad para predecir de estos métodos basados en ACP, se presentan a continuación varias aplicaciones.

#### Aplicación con datos reales

Consideraremos de nuevo el proceso de Poisson que representa en cada instante  $t$  el número de efectos de comercio impagados, que fue analizado en el capítulo dos con el objetivo de predecir su distribución en

un instante futuro a partir de la distribución normal truncada de su media que fue predicha basándose en datos mensuales de una muestra de provincias españolas en el periodo 1989-1996.

### 3.3.1 Descripción de los datos

Recordemos que se tomó aleatoriamente una muestra de 15 provincias españolas para las que se tienen datos mensuales del número mensual de efectos bancarios impagados por cada 10.000 habitantes durante los siete años completos correspondientes al periodo 1990-1996. Observemos que si consideramos el proceso de Poisson  $N(t)$  que representa el número acumulado de impagos hasta el instante  $t$ , sólo disponemos de observaciones discretas a fin de mes para cada provincia, a las que denotaremos por

$$\{n_j : j = 1, \dots, 12, 13, \dots, 24, \dots, 12 \times 7 = 84\}.$$

Dado que no disponemos de la trayectoria completa que sería constante por intervalos aleatorios, distintos para cada provincia, utilizaremos modelos RMCP para predecir el número acumulado de impagos por cada 10.000 habitantes en los doce meses del año 1997 para cada provincia.

Cortando la serie de datos para cada provincia en periodos de 24 meses ( $T=12$ ) se obtienen seis realizaciones de un vector aleatorio de dimensión 24 de la forma

$$X_{wj} = \begin{cases} n_{(w-1)12+j} - n_{(w-1)12} & j = 1, \dots, 12 \\ n_{(w-1)12+j} - n_{w12} & j = 13, \dots, 24 \end{cases} \quad w = 1, \dots, 6,$$

donde cada variable  $X_j$  representa el número acumulado de impagos dentro del año correspondiente. Entonces, un modelo RMCP para estimar el vector aleatorio  $(X_{13}, \dots, X_{24})'$  en términos de  $(X_1, \dots, X_{12})'$  proporcionará predicciones del número mensual acumulado de impagos en 1997.

Construcción de los modelos RMCP

Tomemos la provincia de Granada como ejemplo para ilustrar el proceso de identificación y estimación de los modelos RMCP.

En primer lugar se han estimado dos ACP clásicos, uno para el vector  $(X_1, \dots, X_{12})'$  y otro para  $(X_{13}, \dots, X_{24})'$ . Los porcentajes de varizanzas explicadas y acuuladas figuran en la siguiente tabla:

	[1,12]		[13,14]	
	Var. Explicada	Var. Acumulada	Var. Explicada	Var. Acumulada
1	99.7889	99.7889	99.8506	99.8506
2	0.1458	99.9348	0.1003	99.9509
3	0.0489	99.9836	0.0324	99.9833
4	0.0088	99.9924	0.0159	99.9992
5	0.0076	100.0000	0.0008	100.0000
6	0.0000	100.0000	0.0000	100.0000
Varianza Total				
662314.3634		962505.7494		

Observemos que sólo con la primera componente principal se explica más del 99% de la variabilidad del proceso y que con las dos primeras se supera el 99.9%. Esto quiere decir que vamos a construir los modelos RMCP con las dos primeras cc.pp. del intervalo futuro [13,24] como variables de respuesta. Para decidir cuales son las cc.pp. del intervalo pasado [1,12] que serán utilizadas como predictoras de cada de las dos primeras del futuro, se han calculado las correlaciones lineales entre las cc.pp. de los dos periodos que figuran en la siguiente tabla:



	$\hat{\xi}_1$	$\hat{\xi}_2$	$\hat{\xi}_3$	$\hat{\xi}_4$	$\hat{\xi}_5$
$\hat{\eta}_1$	0.8617	-0.4820	0.1226	-0.0911	-0.0424
$\hat{\eta}_2$	-0.2538	-0.1430	0.6415	-0.6175	-0.3498
$\hat{\eta}_3$	-0.1491	-0.3401	-0.5725	-0.0412	-0.7299
$\hat{\eta}_4$	-0.1492	-0.2671	0.4826	0.7759	-0.2673
$\hat{\eta}_5$	-0.3855	-0.7485	-0.1136	-0.08155	0.5212

Como consecuencia de la magnitud de las correlaciones lineales entre cc.pp. y de la aplicación de los métodos usuales de selección *stepwise*, el siguiente modelo RMCP ha sido identificado y estimado:

$$\text{RMCP}(2;2,2): \tilde{X}_j^2 = \bar{X}_j + \tilde{\eta}_1^2 \hat{\delta}_{j1} + \tilde{\eta}_2^2 \hat{\delta}_{j2} \quad j = 13, \dots, 24$$

donde las dos primeras cc.pp.  $\hat{\eta}_1$  y  $\hat{\eta}_2$  son estimadas mediante los siguientes modelos lineales:

$$\tilde{\eta}_1^2 = 1.0390\hat{\xi}_1 - 15.2068\hat{\xi}_2 \quad \tilde{\eta}_2^2 = 1.1077\hat{\xi}_3 - 2.5107\hat{\xi}_4.$$

De este modo sustituyendo en el modelo RMCP por el valor de las componentes principales en el periodo [1,12] para 1996 se obtienen las predicciones mensuales del número acumulado de impagos para el año 1997 que figuran en la Tabla 3.1, junto a la raíz cuadrada del error cuadrático (ECM) en cada mes dado por

$$ECM(j) = \frac{1}{6} \sum_{w=1}^6 (\tilde{X}_{wj}^2 - X_{wj}) \quad j = 13, \dots, 24.$$

Observemos que no es posible medir exactamente la precisión de las predicciones obtenidas porque al no conocerse la población real de 1997, el número de impagos mensual por cada 10.000 habitantes en este año se ha calculado con la población de 1996 que podría haber cambiado sustancialmente en algunas provincia.

**Tabla3.1** (3ªparte)

	Enero	Febre.	Marzo	Abril	Mayo	Junio	Julio	Agosto	Sept.	Oct.	Nov.	Dic.
<b>Nava.</b>												
v.real	197.1	355.6	498.4	659.0	811.3	958.6	1111.9	1240.3	1378.5	1515.1	1625.5	1770.1
pred.	149.3	292.3	462.5	560.4	720.9	870.0	1033.5	1178.0	1329.4	1481.0	1615.0	1773.1
ECM	34	55	69	93	120	130	170	180	200	220	240	260
<b>Rioja</b>												
v.real	309.0	527.2	717.7	934.5	1159.2	1413.0	1665.8	1863.3	2046.4	2242.9	2413.7	2672.7
pred.	282.8	513.3	744.3	949.7	1163.9	1417.5	1670.9	1918.1	2115.7	2322.8	2526.2	2801.2
ECM	65	100	120	160	240	280	320	360	390	430	460	500
<b>Tole.</b>												
v.real	144.5	246.5	339.3	441.8	547.6	666.1	773.3	852.9	930.2	1016.6	1105.9	1230.4
pred.	131.5	230.7	327.6	413.649	509.6	607.7	682.7	750.6	811.8	814.1	957.0	1066.2
ECM	17	27	36	43	59	78	100	120	130	150	170	180
<b>Valla.</b>												
v.real	111.7	199.2	285.4	381.0	475.2	572.2	661.9	737.5	819.2	898.1	965.3	1059.3
pred.	101.3	185.0	278.8	352.3	467.6	577.2	689.2	798.8	898.6	1017.5	1132.4	1270.2
ECM	23	42	49	64	79	89	110	120	150	170	190	200
<b>Zara.</b>												
v.real	292.0	505.4	687.0	891.4	1099.6	1311.2	1519.1	1679.2	1855.1	2043.3	2208.8	2425.3
pred.	542.9	731.0	922.3	1098.4	1309.7	1490.9	1579.0	1718.4	1817.5	1867.6	1969.4	2127.9
ECM	52	92	130	190	240	290	350	400	450	510	550	570



### Predicción de la moda de los impagos acumulados en 1997

A continuación vamos a usar los modelos RMCP para estimar la distribución de probabilidad normal truncada de la media del proceso de Poisson doblemente estocástico que representa al número de impagos para 1997.

Recordemos que en el capítulo anterior se presentó una metodología para estimar los parámetros de la Normal truncada en 1997 a partir de los parámetros estimados para los años anteriores. En este epígrafe presentamos un procedimiento alternativo basado en la estimación de la media para cada provincia a partir de las predicciones mensuales que se obtienen ajustando un modelo RMCP adecuado, análogamente a como se ha hecho en el apartado anterior para la provincia de Granada. Las predicciones y errores cuadráticos medios para el resto de las provincias figuran en la misma tabla que para Granada.

Las medias predichas de este modo para cada provincia en 1997 son las siguientes:

Prov.	Media	Prov.	Media	Prov.	Media
Barce.	6631.2	Guada.	823.7	Nava.	2691.3
Cant.	1701.3	Huel.	857.4	Rioja	3456
Cast.	3250.5	León	1144.3	Toledo	1588.9
Coru.	1199.8	Lugo	871.9	Valla.	1397.0
Grana.	1554.7	Murcia	2434.3	Zara.	3533.62

De la misma forma que estimábamos las Normales truncadas para cada año en la aplicación del capítulo segundo, se obtiene que en este caso, para el año 1997, la media del proceso de Poisson doblemente estocástico se distribuye mediante la Normal truncada,

$$N_T(823.7, 6631.25, 1946.67, 1029.16).$$

Una vez comprobada la bondad de ajuste mediante el test de bondad de ajuste de Kolmogorov-Smirnov para la Normal truncada, debido

Tabla 3.1 (1ª parte)

	Enero	Febre.	Marzo	Abril	Mayo	Junio	Julio	Agosto	Sept.	Oct.	Nov.	Dic.
<b>Barce.</b>												
v.real	495.2	869.4	1217.7	1590.9	1975.4	2360.0	2756.3	3057.2	3411.3	3766.1	4058.1	4452.9
pred.	400.9	745.2	1098.7	1391.9	1745.0	2076.2	2393.6	2678.7	2961.7	3225.8	3469.7	3777.4
ECM	130	230	310	420	530	610	740	820	930	1000	1100	1200
<b>Canta.</b>												
v.real	124.6	227.1	323.3	429.8	527.0	624.8	722.5	801.2	894.4	987.2	1068.0	1165.1
pred.	106.4	205.7	308.9	393.4	500.4	593.0	699.9	799.8	895.3	1004.5	1099.6	1204.9
ECM	34	55	74	100	130	150	180	200	230	260	290	310
<b>Caste.</b>												
v.real	329.1	556.2	762.9	983.1	1186.2	1390.1	1604.9	1770.2	1958.3	2162.5	2318.5	2525.4
pred.	363.3	647.9	938.3	1208.1	1559.7	1890.6	2200.8	2480.1	2756.5	3050.6	3325.9	3657.4
ECM	42	66	83	130	170	220	250	280	300	330	360	390
<b>Coru.</b>												
v.real	88.3	158.8	228.4	308.0	383.9	465.5	546.0	609.1	671.8	738.6	799.8	880.9
pred.	66.8	122.7	178.3	227.65	284.4	340.1	389.5	444.2	483.5	527.3	573.9	635.1
ECM	9.5	12	17	22	26	32	35	42	48	56	63	71
<b>Grana.</b>												
v.real	117.7	205.3	298.7	392.3	481.0	576.2	671.9	752.3	830.3	920.1	993.6	1088.6
pred.	151.5	266.7	388.1	497.5	613.0	724.8	869.3	968.4	1067.4	1204.8	1330.5	1458.0
ECM	11	11	21	20	26	32	37	43	50	57	64	70



**Tabla 3.1** (2ª parte)

	Enero	Febre.	Marzo	Abril	Mayo	Junio	Julio	Agosto	Sept.	Oct.	Nov.	Dic.
<b>Guada.</b>												
v.real	74.0	132.0	180.4	229.2	293.6	360.8	428.4	480.5	532.4	585.6	628.9	684.6
pred.	80.1	144.5	207.6	262.1	331.4	401.1	480.6	564.3	639.9	713.1	785.7	862.0
ECM	11	19	24	30	37	44	56	63	75	87	94	100
<b>Huel.</b>												
v.real	55.9	99.8	145.3	193.0	235.7	272.0	295.4	327.9	362.8	400.9	432.9	471.1
pred.	42.2	85.2	128.7	162.0	195.4	230.8	257.8	289.7	316.5	347.6	382.1	415.9
ECM	12	20	29	38	49	56	71	79	89	100	110	120
<b>León</b>												
v.real	93.8	183.1	267.0	363.9	453.0	544.2	641.5	723.2	804.6	893.5	971.6	1062.1
pred.	79.4	158.2	236.8	299.3	370.8	438.7	515.8	579.5	650.3	723.7	790.0	869.6
ECM	20	27	38	53	66	79	94	100	120	130	140	150
<b>Lugo</b>												
v.real	63.2	119.9	171.3	222.3	271.2	322.8	371.5	410.2	453.5	495.9	535.8	588.4
pred.	62.4	93.7	132.9	174.7	222.5	264.0	308.9	352.8	391.5	435.1	477.8	531.6
ECM	9.3	25	41	46	58	73	87	1.00	1.10	1.30	1.40	1.50
<b>Murc.</b>												
v.real	172.8	320.0	469.4	612.9	773.8	919.7	1085.5	1219.8	1353.9	1494.6	1618.1	1761.2
pred.	126.8	224.2	346.1	445.5	559.3	660.6	761.9	864.0	963.2	1058.2	1149.0	1257.8
ECM	41	85	110	150	190	230	280	310	340	380	410	440



### 3.3 Modelos RMCP y PCP para procesos de Poisson 109

---

a la proposición 2.5.1, la moda del proceso de Poisson debe encontrarse en el intervalo  $[824, 6630]$ . Usando nuevamente el programa de búsqueda del máximo de la función masa de probabilidad de nuestro proceso de Poisson con la media estimada, la moda resultó ser 1946 con probabilidad 0.000449078829729.

Se observa que debido al amplio rango en el que nos movemos para encontrar la moda, ésta se puede considerar muy similar siguiendo el método propuesto en la sección 2.5 y la calculada usando las estimaciones de las medias mediante los modelos RMCP para el año 1997. El método que aplicamos en el segundo capítulo, no requiere los conocimientos teóricos de éste tercer capítulo ni tantos cálculos, como la estimación de la moda hecha en esta sección, por lo que podría pensarse que no merece la pena esta nueva aproximación al problema. Sin embargo, la naturaleza de los datos podría provocar que las regresiones hechas para estimar los parámetros de la media futura del proceso no diesen buenos resultados si éstos fuesen muy cambiantes en el tiempo. Por esta razón debería observarse el comportamiento de las medias en el tiempo antes de decidirse a estimar la moda del proceso de Poisson doblemente estocástico de una u otra forma.



# Apéndice

Este apéndice contiene los programas implementados a los que hacemos referencia en el texto. Todos ellos están realizados en *Mathematica*; hemos escogido este programa porque aproxima las integrales sin primitiva que tenemos que resolver varias veces con un método aproximativo gaussiano, lo que es más adecuado que otros métodos más groseros de aproximación que utilizan otros programas comerciales.

## Programa Est-MVM.nb

Este programa calcula los estimadores máximo verosímiles modificados (de ahí el nombre de Est-MVM.nb) para una variable aleatoria Normal truncada en  $A(t)$  y  $B(t)$ , con media  $\mu(t)$  y desviación típica  $\sigma(t)$ .

```
Clear[A,B,M,S,n,G,F1,F2,F,F11,F12,F21,F22,DF, $\mu$ , $\sigma$ ,a0,a1,aux];
{A, B} = Input[‘‘Los puntos de truncamiento, A y B, de la
Normal truncada son (escribir entre llaves y separando por
comas):  ’’];
{n, M, S} = Input[‘‘El tamaño de la muestra, su media y
su desviación típica son (escribir separando por comas y entre
llaves):  ’’];
G[k_,  $\mu$ _, $\sigma$ _] := NIntegrate[ $x^k \text{Exp}[-\frac{(x-\mu)^2}{2\sigma^2}]$ , {x,A,B}];
F1[ $\mu$ _, $\sigma$ _] :=  $\frac{G[1,\mu,\sigma]}{G[0,\mu,\sigma]} - M$ ;
```



```

F2[μ-,σ-] :=  $\frac{G[2,\mu,\sigma]}{G[0,\mu,\sigma]} - \left(\frac{G[1,\mu,\sigma]}{G[0,\mu,\sigma]}\right)^2 + 2\frac{\sigma^2}{n} + S^2 - \frac{1}{n}$ ;
F[μ-,σ-] := {F1[μ,σ],F2[μ,σ]};
F11[μ-,σ-] :=
 $\frac{G[2,\mu,\sigma]}{\sigma^2 G[0,\mu,\sigma]} - \frac{\mu G[1,\mu,\sigma]}{\sigma^2 G[0,\mu,\sigma]} - \frac{G^2[1,\mu,\sigma]}{\sigma^2 G^2[0,\mu,\sigma]} + \frac{\mu G[1,\mu,\sigma]}{\sigma^2 G[0,\mu,\sigma]}$ ;
F12[μ-,σ-] :=
 $\frac{G[2,\mu,\sigma]}{\sigma^3 G[0,\mu,\sigma]} - \frac{G[2,\mu,\sigma]}{\sigma^3 G[0,\mu,\sigma]} - \frac{G[1,\mu,\sigma]G[2,\mu,\sigma]}{\sigma^3 G^2[0,\mu,\sigma]} + 2\frac{G^2[1,\mu,\sigma]}{\sigma^3 G^2[0,\mu,\sigma]} - \frac{\mu^2 G[1,\mu,\sigma]}{\sigma^3 G[0,\mu,\sigma]}$ ;
F21[μ-,σ-] :=
 $\frac{G[3,\mu,\sigma]}{\sigma^2 G[0,\mu,\sigma]} - \frac{\mu\sigma^2}{\sigma^2 G[0,\mu,\sigma]} - \frac{G[1,\mu,\sigma]G[2,\mu,\sigma]}{\sigma^2 G[0,\mu,\sigma]} + \frac{\mu G[2,\mu,\sigma]}{\sigma^2 G[0,\mu,\sigma]} - 2 F11[\mu, \sigma]$ ;
F22[μ-,σ-] :=
 $\frac{G[4,\mu,\sigma]}{\sigma^3 G[0,\mu,\sigma]} - 2\frac{\mu G[3,\mu,\sigma]}{\sigma^3 G[0,\mu,\sigma]} + \sigma\frac{\mu^2 G[2,\mu,\sigma]}{\sigma^3 G[0,\mu,\sigma]} - \frac{G^2[2,\mu,\sigma]}{\sigma^3 G^2[0,\mu,\sigma]} + 2\frac{\mu G[1,\mu,\sigma]G[2,\mu,\sigma]}{\sigma^3 G^2[0,\mu,\sigma]}$ 
 $-\frac{\mu^2 G[2,\mu,\sigma]}{\sigma^3 G[0,\mu,\sigma]} - 2 F12[\mu, \sigma]$ ;
DF[μ-,σ-] := {{F11[μ,σ],F12[μ,σ]},{F21[μ,σ],F22[μ,σ]}};

tol=Input[‘Error máximo tolerado: ’];

{μ0,σ0} := {M, S};

m = 40; aux = 0;
For[i=1, i≤m, i++, sol = Solve[DF[μ0,σ0].
({μ,σ} - {μ0,σ0}) == F[μ0,σ0],{μ,σ}]; {μ1,σ1}={μ,σ}
/.sol[[1]]; If[F[μ1,σ1] == {0, 0}, aux = 1; Break[]];
If[ $\sqrt{(\mu_1-\mu_0)^2} + \sqrt{(\sigma_1-\sigma_0)^2} \leq \text{tol}$ , aux = 2; Break[]];
{μ0,σ0} = {μ1,σ1}
If[aux == 1, Print[’Solución exacta: ’, {μ0,σ0}],
If[aux == 2, Print[’Solución aproximada: ’, {μ0,σ0}];
Print[’Número de iteraciones: ’, i-1];
Print[’Error máximo cometido ’, tol], Print[’Se ha alcanzado el número máximo de iteraciones y la solución obtenida es: ’, {μ0,σ0}]]]

```

Programa KS.nb

Este programa es un test de bondad de ajuste de Kolmogorov-Smirnov para una Normal truncada  $\mathcal{N}_T(A(t), B(t), \mu(t), \sigma(t))$ .

```

Clear[A,B,Λ,μ,σ,n,Muestra,Y,X];
n=Input['El tamaño de la muestra es '];
Muestra=Input['Introduzca los valores de la muestra se-
parados por comas y entre llaves'];

While[Length[Muestra]!=n, n=Input['El tamaño de la muestra
no concuerda con el número de datos. Vuelva a introducir el
tamaño muestral'];
Muestra=Input['Introduzca los valores de la muestra se-
parados por comas y entre llaves'];
Y = Sort[Muestra];A= Y[[1]];B=Y[[n]];
(*X será la lista de elementos distintos de la muestra*)
X={Y[[1]]};
For[i = 2,i<=n, i++, If[Y[[i]]!= Y[[i-1]],
AppendTo[X,Y[[i]]]];
(*Ahora guardo el número de veces que aparece cada valor
muestral en XX*)
XX = Table[Length[Position[Y,X[[i]]]], {i, 1, Length[X],
1}]
{μ,σ} = Input['A continuación se hará un Test de Kolmo-
gorov-Smirnov estudiando el ajuste de los datos muestrales
a una Normal truncada. Se tomará como punto inferior de trun-
camiento el menor valor muestral y como superior al mayor valor.
Introduzca la media y la desviación típica de la distribu-
ción entre llaves y separadas por comas.'];

```

$$F[\Lambda] := \left( \int_{\frac{A-\mu}{\sigma}}^{\frac{B-\mu}{\sigma}} E^{-\frac{s^2}{2}} ds \right)^{-1} \int_{\frac{A-\mu}{\sigma}}^{\frac{\Lambda-\mu}{\sigma}} E^{-\frac{s^2}{2}} ds ;$$

(\*Guardo en S la evaluación de la función de distribución muestral en cada punto\*)

```
S = Table[Sum[XX[[j]], {j,i}] /n,{i,1,Length[XX],1]
Dn = Max[Table[Abs[S[[i]]-N[F[X[[i]]]]], {i,1,Length[X],
1}]]
```

(\*Comprobar, usando la tabla de Kolmogorov-Smirnov, si se acepta o no la hipótesis de que los datos se distribuyen mediante una Normal truncada con los parámetros propuestos\*)

## Programa normalt.nb

Este programa calcula el valor del proceso de Poisson doblemente estocástico cuya media es la Normal truncada  $\mathcal{N}_T(A(t), B(t), \mu(t), \sigma(t))$ , con máxima probabilidad.

```
Clear[A,B,Λ,μ,σ,k1,k2,P,f,F1,F2,n,m,p];
A=Input[‘El punto inferior de truncamiento es’];
B=Input[‘El punto superior de truncamiento es’];
μ=Input[‘La media de la variable es’];
σ=Input[‘La desviación típica de la variable es’];
K1= (σ ∫A-μB-μ E-s2/2 ds)-1 ;
(*La función de densidad de la Normal Truncada es:*)
f[Λ_] := K1 E-1/2 (Λ-μ)2/σ2;
N[f[Λ]]
Print[‘La gráfica de la función de densidad es’];
Plot[f[Λ], {Λ,A,B}, PlotRange ->All];
Λ = Input[‘¿En qué valor quiere calcular la función de
distribución?’];
F1 := ∫A-μΛ-μ E-u2/2 du ;
F2 := ∫A-μB-μ E-u2/2 du ;
```

```

N[ $\frac{F1}{F2}$ ];
Print["La función de distribución en ",  $\Lambda$ , "es: ", -
N[ $\frac{F1}{F2}$ ]]
k2 =  $\frac{1}{\sqrt{2\pi}}$  F2;
(*La probabilidad de que un proceso de Poisson doblemente
estocástico con media  $\Lambda(t)$  tome el valor n es:*)
P[n_] :=  $\frac{1}{n!}$  k2-1  $\frac{1}{\sigma\sqrt{2\pi}}$  NIntegrate[ $\Lambda^n E^{-\Lambda} E^{-\frac{1}{2}(\frac{\Lambda-\mu}{\sigma})^2}$ , { $\Lambda$ , A, B}]
(* Buscamos el natural, n, que maximiza la probabilidad
anterior.*)
lista=Table[P[n],{n,Floor[A],Ceiling[B]};
ListPlot[lista];
m=Max[lista];
p=Position[lista,m][[1,1]];
Print["El entero en el que se da una mayor probabilidad
es ", Floor[A]+p, " con probabilidad ",m]

```



## Bibliografía

ANDERSEN, P. K.; BORGAN, Ø.; GILL, R.D. AND KEIDING, N. (1993): *Statistical Models Based on Counting Processes*. Springer-Verlag, New York.

AGUILERA DEL PINO, A. M., VALDERRAMA BONNET, M. J., Y DEL MORAL, M. J. (1992): "Un Método para la Aproximación de Estimadores en ACP. Aplicación al Proceso de Ornstein-Uhlenbeck" *Revista de la SOCHE*. Vol. 6, 57-77.

AGUILERA DEL PINO, OCAÑA, F.A. Y A. M., VALDERRAMA BONNET, M. J. (1994): "A Computational Algorithm for PCA of Random Processes" *Proceedings in COMPSTAT. Software Descriptions*. Ed. Dutter , R. y Grossmann, W. 39-40. Berlin, Physica-Verlag.

AGUILERA DEL PINO, GUTIÉRREZ, R., OCAÑA, F.A. Y A. M., VALDERRAMA BONNET, M. J. (1995): "Computational Approaches to Estimation in the Principal Component Analysis of a Stochastic Process" *Applied Stochastic Models and Data Analysis*. Vol. 11, No. 4, 279-299.

AGUILERA DEL PINO, GUTIÉRREZ, R. Y A. M., VALDERRAMA BONNET, M. J. (1996a): "Approximation of estimators in the

PCA of a Stochastic Process using B-splines" *Communications in Statistics*. Vol. 25, No. 3, 671-690.

AGUILERA DEL PINO, OCAÑA, F.A. Y A. M., VALDERRAMA BONNET, M. J. (1996b): "Análisis en Componentes Principales de un Proceso Estocástico con Funciones Muestrales Escalonadas" *Qüestio*. Vol.20, 7-28.

AGUILERA DEL PINO, OCAÑA, F.A. Y A. M., VALDERRAMA BONNET, M. J. (1997): "An Approximated Principal Component Prediction Model for Continuous-Time Stochastic Processes" *Applied Stochastic Models and Data Analysis*. Vol. 13, 61-72.

ASCHER, H. (1989): "Letter to the editor." *Microelectronics & Reliability*. Vol.30, No.4, 807-808.

BAR-DAVID, I. (1969): "Communication under the Poisson Regime." *IEEE Transactions on Information Theory*. Vol.15, No.1, 31-37.

BOEL, R. AND BENEŠ, V. (1980): "Recursive Nonlinear Estimation of a Diffusion Acting as the Rate of an Observed Poisson Process" *IEEE Transactions on Information Theory*. Vol. 26, No. 5, 561-575.

COX, D. R. (1955): "some Statistical Methods Connected with Series of Events" *J. Royal Statistical Society B*. Vol.17, 129-164.

DALEY, D. J. AND VERE-JONES, D. (1988): *An Introduction to the Theory of Point Processes*. Springer-Verlag, New York.

DEVILLE, J. C. (1973): "Estimation of the Eigenvalues and of the Eigenvectors of a Covariance Operator". *Note Interne de l'INSEE*

- DEVILLE, J. C. (1974): "Méthodes Statistiques et Numériques de Analyse Harmonique". *Annles de l'INSEE*. Vol. 15, 3-101.
- FUKUNAGA, K. (1990): *Introduction to Statistical Pattern Recognition*. 2<sup>a</sup> Ed. Academic Press, San Diego.
- GRANDELL, J. (1971): "On Stochastic Processes Generated by a Stochastic Intensity Function" *Skndinavisk Aktuarietidskrift*. 204-240.
- GRIGORIU, M. (1995): *Applied Non-Gaussian Processes, Examples Theory, Simulation, Linear Random Vibration and MATLAB Solutions*. PTR Prentice Hall, N. J.
- GROSS-HARRIS (1974): *Fundamentals of Queueing Theory*. Wiley, New York.
- HÄRTLER, G. (1989): "The nonhomogeneous Poisson process - A model for the reliability of complex repairable systems. -" *Microelectronics & Reliability*. Vol.29, No.3, 381-386.
- JACKSON, J.E. (1991): *A User's Guide to Principal Components*. Wiley, N. J.
- JIMÉNEZ, F. (1991): *Contribución al Estudio de Procesos Poissonianos de Recuento con Intensidad Aleatoria*. Tesis, Universidad de Granada.
- JOHNSON, N. L. AND KOTZ, S. (1970): *Continous Univariate Distributions-I, Distributions in Statistics*. Wiley, New York.
- LAST, G. AND BRANDT, A. (1995) *Marked Point processes on the real line (The dynamic approach)*. Springer-Verlag, New York.



LIPSTER, R. AND SHIRYAYEV, A. N. (1997): *Statistics of Random Processes, I: General Theory*. Springer-Verlag, New York.

MITTAL, M. M. AND DAHIYA, R. C. (1987): "Estimating the Parameters of a Doubly Truncated Normal Distribution." *Communications in Statistics - Simula*. Vol.16, 141-159.

RAMÍREZ, V., GONZÁLEZ, P., PASADAS, M. AND BARRERA, D. (1997): *Matemáticas con Mathematica*. Vols.I y III. Proyecto Sur Ediciones, Granada.

RAMSAY, J. O. Y SILVERMAN, B. M. (1997): *Functional Data Analysis*. Springer-Verlag, N. J.

RIESZ, F. Y SZ-NAGY, B. (1990): *Lecons d'Analyse Fonctionnelle*. Reimpresión de la 3ª edición publicada por Gauthier-Villars y Akademiai Kiad en 1995. Ed. Jacques Gabay. Sceaux.

RODRÍGUEZ-BOUZAS, P. (1996): *Estudio de Algunos Procesos de tipo Poissoniano*. Memoria de Licenciatura, Universidad de Granada.

RODRÍGUEZ-BOUZAS, P., AGUILERA DEL PINO, A. M. Y VALDERRAMA BONNET, M. J. (1998a): "Máxima probabilidad en un Proceso de Poisson Doblemente Estocástico con Esperanza Aleatoria Normal Truncada" XXIV Congreso Nacional de Estadística e Investigación Operativa. Almería.

RODRÍGUEZ-BOUZAS, P., AGUILERA DEL PINO, A. M. Y VALDERRAMA BONNET, M. J. (1998b): "Doubly Stochastic Poisson Process with Random Mean a Truncated Normal Variable" First Euro-Japanese Workshop on Stochastic Modelling in Finance, Insurance, Production and Reliability. Brussels.

- SERFOZO, R. F. (1972): "Conditional Poisson Processes." *J. Appl. Prob.* Vol.9, 288-302.
- SNYDER, D. L. AND MILLER, M. I. (1991): *Random Point Processes in Time and Space*. 2nd edition. Springer Verlag, New York.
- TODOROVIC, P. (1992): *An Introduction to Stochastic Processes and Their Applications*. Springer-Verlag, N. J.
- VALDERRAMA, M. J., JIMÉNEZ, F., GUTIÉRREZ, M. AND MARTÍNEZ-ALMÉCIJA, A. (1995): "Estimation and Filtering on a Doubly Stochastic Poisson Process." *Applied Stochastic Models and Data Analysis*. Vol.11, 13-24.
- ZHAO, M. AND XIE, M. (1996): "On maximum likelihood estimation for a general non-homogeneous Poisson process." *Scandinavian Journal of Statistics*. Vol.23, 597-607.