
Cadenas de Markov de tiempo continuo y aplicaciones

María Valentina Vega

Orientador: Ricardo Fraiman

26 de abril de 2004

Trabajo Monográfico
Licenciatura en Matemática
Universidad de la República
Montevideo - Uruguay.

Introducción

Los procesos estocásticos son sucesiones de eventos gobernados por leyes probabilísticas. Muchas aplicaciones de los procesos estocásticos aparecen en física, ingeniería, biología, medicina y otras disciplinas así como también en otras ramas de la matemática. El objetivo de estas notas es presentar un conjunto de nociones y resultados útiles para trabajar con procesos estocásticos. La presentación que se hace no utiliza teoría de la medida sino que adopta un punto de vista “constructivo”.

Como es conocido, las cadenas de Markov son procesos de corta memoria en el sentido de que solo “recuerdan” el último estado visitado para decidir cual será el próximo. En términos formales, el proceso $\{X_n\}_{n \in \mathbb{N}}$ con espacios de estados E es una cadena de Markov si

$$P(X_{n+1} = y | X_n = x_n, \dots, X_0 = x_0) = P(X_{n+1} = y | X_n = x_n)$$

donde $y, x_0, \dots, x_n \in E$. En procesos con “larga memoria” el valor que toma el proceso en cada paso depende de todo el pasado. Este tipo de procesos fue introducido por Onicescu y Mihoc bajo el nombre de cadenas con conexiones completas en 1935. El primer capítulo de este trabajo es una introducción a este tipo de procesos en donde se encuentra una construcción posible de los mismos y algunos resultados.

En el tercer capítulo se introducen las cadenas de Markov de tiempo continuo y se demuestran resultados importantes de este tipo de procesos. Para el desarrollo de este capítulo se necesita un previo estudio de los procesos de Poisson, el cual se hace en el segundo capítulo de estas notas. En dicho capítulo no solo se definen los procesos de Poisson y se estudian sus propiedades básicas, sino que también se introduce el concepto de proceso de Poisson bi-dimensional. Este tipo de procesos es usado en el desarrollo de las restantes secciones para construir procesos homogéneos uno-dimensionales. La construcción de procesos de Poisson realizada en el segundo capítulo es un caso particular de la construcción propuesta por Neveu (1977). La construcción de procesos de dimensión menor como proyección de procesos de dimensión mayor fué introducida por Kurtz (1989). La mayor diferencia que hay entre cadenas de Markov de tiempo discreto y cadenas de Markov de tiempo continuo es que las primeras saltan de un estado a otro en tiempos enteros $1, 2, \dots$ y las de tiempo continuo lo hacen en tiempos aleatorios τ_1, τ_2, \dots .

El cuarto capítulo tiene como finalidad mostrar algunas de las aplicaciones que existen para los procesos descritos en los capítulos anteriores. Se mostrarán ejemplos relacionados con la biología pero dichos procesos tienen muchas aplicaciones en otras áreas como ser economía, sociología, física, etc.

Para terminar se han anexado al final de este trabajo dos apéndices. El primero describe cierto tipo de proceso utilizado en el desarrollo de estas notas y algunas propiedades del mismo. En el segundo apéndice se definen las funciones generatrices y se presentan algunas de las propiedades de éstas que son utilizadas en el capítulo 4.

Índice general

1. Cadenas con conexiones completas	5
1.1. Especificaciones	5
1.2. Una construcción	7
1.3. Resultados	11
2. Procesos de Poisson	17
2.1. Procesos Puntuales	17
2.2. Procesos de Poisson	19
2.3. Definición formal de un proceso puntual	21
2.4. Propiedades	22
2.5. Resultados	30
2.6. Propiedad de los estadísticos de orden	32
2.7. Definiciones alternativas de procesos de Poisson	34
2.8. Procesos de Poisson en dos o mas dimensiones	39
2.9. Proyecciones	40
2.10. Superposición de procesos de Poisson	42
2.11. Procesos no homogéneos	45
3. Cadenas de Markov de tiempo continuo	48
3.1. Procesos de Markov de salto puro	48
3.2. Propiedades	52
3.3. Explosiones	55
3.4. Ecuaciones de Kolmogorov	58
3.5. Clasificación de estados y medidas invariantes	62
3.6. Esqueletos	63
3.7. Procesos de nacimiento y muerte	65
4. Aplicaciones en Biología	67
4.1. Modelos estocásticos simples para el crecimiento de poblaciones.	67
4.1.1. Procesos de Nacimiento y Muerte.	67
4.1.2. Procesos de Nacimiento y Muerte no homogéneos.	70
4.1.3. Procesos simples de Nacimiento, Muerte e Inmigración.	73
4.2. Modelos estocásticos para el crecimiento de población por sexos.	73

Índice general	4
-----------------------	----------

A. Proceso Casa de Cartas	76
----------------------------------	-----------

B. Funciones Generatrices	78
----------------------------------	-----------

B.1. Algunos teoremas y propiedades.	78
--	----

Capítulo 1

Cadenas con conexiones completas

En este capítulo comenzaremos introduciendo la noción de especificación y en la segunda sección veremos como construir un proceso estocástico que sea “compatible” con una especificación. En la última sección veremos algunos resultados para este tipo de procesos. El concepto de especificación es utilizado en Mecánica estadística.

1.1. Especificaciones

Consideremos E un espacio de estados finito. Sean $\mathbb{N}^* = \mathbb{N} - \{0\}$ y $-\mathbb{N}^* = \{-i : i \in \mathbb{N}^*\}$ los conjuntos de los números enteros positivos y negativos, respectivamente.

Definición 1.1.1. Decimos que una función $\mathcal{P} : E \times E^{-\mathbb{N}^*} \rightarrow [0, 1]$ es una especificación si satisface que para cada $\underline{w} = (w_{-1}, w_{-2}, \dots) \in E^{-\mathbb{N}^*}$:

1. $\mathcal{P}(y; \underline{w}) \geq 0$ para todo $y \in E$
2. $\sum_{y \in E} \mathcal{P}(y; \underline{w}) = 1$

Proposición 1.1.2. Dada $\underline{w} = (w_{-1}, w_{-2}, \dots) \in E^{-\mathbb{N}^*}$ una configuración arbitraria del “pasado”, es posible construir un proceso estocástico $\{X_t^{\underline{w}} : t \in \mathbb{Z}\}$ en E con la propiedad de que para todo $n \geq 0$ y valores arbitrarios $x_0, x_1, \dots, x_n \in E$ se cumple que:

$$P(X_t^{\underline{w}} = x_t : t = 0, \dots, n \mid X_i^{\underline{w}} = w_i : i \in -\mathbb{N}^*) = \prod_{t=0}^n \mathcal{P}(x_t; \widehat{w}_t) \quad (1.1)$$

donde $\widehat{w}_t = (x_{t-1}, \dots, x_0, \underline{w}) = (x_{t-1}, \dots, x_0, w_{-1}, w_{-2}, \dots)$

Demostración. En primer lugar observemos que es posible construir una familia de particiones del intervalo $[0, 1]$

$$\{\{B(y; \underline{w}) : y \in E\} : \underline{w} \in E^{-\mathbb{N}^*}\} \quad (1.2)$$

que satisfaga las condiciones:

1. $|B(y; \underline{w})| = \mathcal{P}(y; \underline{w})$
2. $\cup_{y \in E} B(y; \underline{w}) = [0, 1]$

para todo $\underline{w} \in E^{-\mathbb{N}^*}$. Para ello alcanza con “ordenar” los elementos de E y luego concatenar intervalos de longitud $\mathcal{P}(y; \underline{w})$. La propiedad (2) en la definición de especificación nos permite hacer esta construcción.

Una vez construída la familia de particiones (1.2) definimos la función $F : E^{-\mathbb{N}^*} \times [0, 1] \rightarrow E$ tal que:

$$F(\underline{w}; u) := \sum_{y \in E} y \mathbf{1}_{\{u \in B(y; \underline{w})\}} \quad (1.3)$$

y definimos $X_0^{\underline{w}} = F(\underline{w}; U_0)$ y para $t \geq 1$

$$X_t^{\underline{w}} = F(X_{t-1}^{\underline{w}}, \dots, X_0^{\underline{w}}, \underline{w}; U_t) \quad (1.4)$$

donde $\{U_n : n \in \mathbb{Z}\}$ es una familia de v.a.i.i.d. distribuídas uniformemente en el intervalo $[0, 1]$. Queda ver que el proceso $\{X_t^{\underline{w}} : t \geq 0\}$ verifica la condición (1.1). Sean $n \geq 0$ y $x_0, \dots, x_n \in E$.

$$\begin{aligned} & P(X_t^{\underline{w}} = x_t : t = 0, \dots, n \mid X_i^{\underline{w}} = w_i : i \in -\mathbb{N}^*) = \\ & P(F(X_{t-1}^{\underline{w}}, \dots, X_0^{\underline{w}}, \underline{w}; U_t) = x_t : t = 0, \dots, n \mid X_i^{\underline{w}} = w_i : i \in -\mathbb{N}^*) = \\ & P\left(\sum_{y \in E} y \mathbf{1}_{\{U_t \in B(y; X_{t-1}^{\underline{w}}, \dots, X_0^{\underline{w}}, \underline{w})\}} = x_t : t = 0, \dots, n \mid X_i^{\underline{w}} = w_i : i \in -\mathbb{N}^*\right) = \\ & \prod_{t=0}^n P\left(\sum_{y \in E} y \mathbf{1}_{\{U_t \in B(y; X_{t-1}^{\underline{w}}, \dots, X_0^{\underline{w}}, \underline{w})\}} = x_t \mid X_i^{\underline{w}} = w_i : i \in -\mathbb{N}^*\right) = \\ & \prod_{t=0}^n P\left(\sum_{y \in E} y \mathbf{1}_{\{U_t \in B(y; X_{t-1}^{\underline{w}}, \dots, X_0^{\underline{w}}, X_{-1}^{\underline{w}}, \dots)\}} = x_t\right) = \\ & \prod_{t=0}^n P(\mathbf{1}_{\{U_t \in B(x_t; X_{t-1}^{\underline{w}}, \dots, X_0^{\underline{w}}, X_{-1}^{\underline{w}}, \dots)\}} = 1) = \\ & \prod_{t=0}^n |B(x_t; x_{t-1}, \dots, x_0, w_{-1}, \dots)| = \\ & \prod_{t=0}^n \mathcal{P}(x_t; x_{t-1}, \dots, x_0, w_{-1}, \dots) = \\ & \prod_{t=0}^n \mathcal{P}(x_t; \widehat{w}_t) \end{aligned}$$

donde $\widehat{w}_t = (x_{t-1}, \dots, x_0, w_{-1}, w_{-2}, \dots) \in E^{-\mathbb{N}^*}$, $t = 0, \dots, n$. □

Introduciremos a continuación algunas notaciones y veremos una manera particular de construir las particiones $B(\cdot; \cdot)$ para la especificación \mathcal{P} .

Dados $k \in \mathbb{N}$, $y \in E$ y $\underline{w} \in E^{-\mathbb{N}^*}$ definimos:

$$a_k(y; \underline{w}) := \inf \{ \mathcal{P}(y; w_{-1}, \dots, w_{-k}, \underline{z}) : \underline{z} \in E^{-\mathbb{N}^*} \} \quad (1.5)$$

donde $(w_{-1}, \dots, w_{-k}, \underline{z}) = (w_{-1}, \dots, w_{-k}, z_{-1}, z_{-2}, \dots)$. Observar que $a_0(y; \underline{w}) := \inf \{ \mathcal{P}(y; \underline{z}) : \underline{z} \in E^{-\mathbb{N}^*} \}$ y por lo tanto no depende de \underline{w} . Dado $k \in \mathbb{N}$ definimos:

$$a_k := \inf_{\underline{w} \in E^{-\mathbb{N}^*}} \left(\sum_{y \in E} a_k(y; \underline{w}) \right) \quad (1.6)$$

Para $m \in \mathbb{N}$, sea:

$$\beta_m := \prod_{k=0}^m a_k \quad (1.7)$$

y

$$\beta := \lim_{m \rightarrow +\infty} \beta_m \quad (1.8)$$

Definición 1.1.3. Una especificación \mathcal{P} se dice de conexiones completas si $a_0 > 0$.

Definición 1.1.4. Decimos que una medida ν en $E^{\mathbb{Z}}$ es compatible con la especificación \mathcal{P} si:

$$\nu(\underline{X} \in E^{\mathbb{Z}} : X_n = y \mid X_{n+j} = w_j : j \in -\mathbb{N}^*) = \mathcal{P}(y; \underline{w}) \quad (1.9)$$

para todo $n \in \mathbb{Z}$, $y \in E$ y $\underline{w} = (w_{-1}, w_{-2}, \dots) \in E^{-\mathbb{N}^*}$.

1.2. Una construcción

En esta sección veremos una construcción del proceso establecido en la proposición 1.1.2. Asumamos que el espacio de estados $E = \{1, 2, \dots, q\}$ para algún q entero positivo. Dado $\underline{w} \in E^{-\mathbb{N}^*}$, definimos para $y \in E$ $b_0(y; \underline{w}) := a_0(y; \underline{w})$ y para $k \geq 1$ definimos:

$$b_k(y; \underline{w}) := a_k(y; \underline{w}) - a_{k-1}(y; \underline{w})$$

Para cada $\underline{w} \in E^{-\mathbb{N}^*}$ sea $\{B_k(y; \underline{w}) : y \in E, k \in \mathbb{N}\}$ una partición del intervalo $[0, 1]$ con las siguientes propiedades:

1. Para $y \in E$ y $k \geq 0$, $B_k(y; \underline{w})$ es un intervalo cerrado por izquierda y abierto por derecha con medida de Lebesgue $b_k(y; \underline{w})$.
2. Los intervalos están dispuestos en orden lexicográfico crecientes con respecto a y y a k de tal manera que el extremo izquierdo de un intervalo coincide con el extremo derecho del anterior:

$$B_0(1; \underline{w}), B_0(2; \underline{w}), \dots, B_0(q; \underline{w}), B_1(1; \underline{w}), \dots, B_1(q; \underline{w}), \dots$$

sin intersecciones.

Mas precisamente, llamando $L(A) := \inf\{x : x \in A\}$ y $R(A) := \sup\{x : x \in A\}$, $A \in \mathbb{R}$ la construcción anterior requiere que:

1. $L(B_0(1, \underline{w})) = 0$.
2. $R(B_k(y; \underline{w})) = L(B_k(y + 1; \underline{w}))$ si $y = 1, \dots, q - 1$.
3. $R(B_k(q; \underline{w})) = L(B_{k+1}(1; \underline{w}))$.

Definimos :

$$B(y; \underline{w}) := \bigcup_{k \geq 0} B_k(y; \underline{w}) \quad (1.10)$$

Las propiedades anteriores implican que:

1. $R(B_k(q; \underline{w})) = \sum_{y \in E} a_k(y; \underline{w})$.

Demostración. Debido a la construcción de la partición del intervalo $[0, 1]$ se tiene que:

$$\begin{aligned} R(B_k(q; \underline{w})) &= \sum_{i=0}^k \sum_{j=1}^q |B_i(j; \underline{w})| \\ &= \sum_{i=0}^k \sum_{j=1}^q b_i(j; \underline{w}) \\ &= \sum_{i=1}^k \sum_{j=1}^q (a_i(j; \underline{w}) - a_{i-1}(j; \underline{w})) + \sum_{j=1}^q a_0(j; \underline{w}) \\ &= \sum_{j=1}^q \sum_{i=1}^k (a_i(j; \underline{w}) - a_{i-1}(j; \underline{w})) + \sum_{j=1}^q a_0(j; \underline{w}) \\ &= \sum_{j=1}^q (a_k(j; \underline{w}) - a_0(j; \underline{w})) + \sum_{j=1}^q a_0(j; \underline{w}) \\ &= \sum_{j=1}^q a_k(j; \underline{w}) \end{aligned}$$

□

2. $\lim_{k \rightarrow +\infty} R(B_k(q; \underline{w})) = 1$.

Demostración. Utilizando el resultado de (1) se tiene que:

$$\begin{aligned}
 \lim_{k \rightarrow +\infty} R(B_k(q; \underline{w})) &= \lim_{k \rightarrow +\infty} \sum_{y \in E} a_k(y; \underline{w}) \\
 &= \sum_{y \in E} \lim_{k \rightarrow +\infty} a_k(y; \underline{w}) \\
 &= \sum_{y \in E} \mathcal{P}(y; \underline{w}) \\
 &= 1
 \end{aligned}$$

donde la última igualdad es porque \mathcal{P} es una especificación. \square

3. $|B(y; \underline{w})| = \mathcal{P}(y; \underline{w})$

Demostración. Se deduce de las propiedades de la partición hecha que:

$$\begin{aligned}
 |B(y; \underline{w})| &= \left| \bigcup_{k \geq 0} B_k(y; \underline{w}) \right| \\
 &= \sum_{k \geq 0} |B_k(y; \underline{w})| \\
 &= \sum_{k \geq 0} b_k(y; \underline{w}) \\
 &= \sum_{k \geq 1} (a_k(y; \underline{w}) - a_{k-1}(y; \underline{w})) + a_0(y; \underline{w}) \\
 &= \lim_{k \rightarrow +\infty} a_k(y; \underline{w}) \\
 &\leq \mathcal{P}(y; \underline{w})
 \end{aligned}$$

Por otro lado sabemos que:

$$1 = \sum_{y \in E} |B(y; \underline{w})| \leq \sum_{y \in E} \mathcal{P}(y; \underline{w}) = 1$$

Entonces

$$|B(y; \underline{w})| = \mathcal{P}(y; \underline{w}) \tag{1.11}$$

\square

4. $\left| \bigcup_{y \in E} B(y; \underline{w}) \right| = 1.$

Demostración. Utilizando que los intervalos son disjuntos y la propiedad (3) se tiene que:

$$\begin{aligned}
 \left| \bigcup_{y \in E} B(y; \underline{w}) \right| &= \sum_{y \in E} |B(y; \underline{w})| \\
 &= \sum_{y \in E} \mathcal{P}(y; \underline{w}) \\
 &= 1
 \end{aligned}$$

donde la última igualdad es porque \mathcal{P} es una especificación. \square

Definimos para $l \geq 0$:

$$B_l(\underline{w}) := \bigcup_{y \in E} B_l(y; \underline{w})$$

Observación 1.2.1. 1. Observar que $B_0(y; \underline{w})$ no depende de \underline{w} para todo $y \in E$, ya que es un intervalo de longitud $b_0(y; \underline{w}) = a_0(y; \underline{w})$ lo cual no depende de \underline{w} . Esto implica que $B_0(\underline{w})$ tampoco depende de \underline{w} .

2. Se deduce de la definición de $a_k(y; \underline{w})$ que tanto $B_k(y; \underline{w})$ como $B_k(\underline{w})$ solo dependen de las primeras k coordenadas de \underline{w} ya que son intervalos de longitud $b_k(y; \underline{w}) = a_k(y; \underline{w}) - a_{k-1}(y; \underline{w})$.

Hasta aquí hemos construido una partición con todas las propiedades necesarias. Veamos ahora como construir el proceso buscado. Sea $\underline{w} \in E^{-\mathbb{N}^*}$ una condición de borde arbitraria. Para dicha condición definimos la función $F : E^{-\mathbb{N}^*} \times [0, 1] \rightarrow E$ tal que

$$F(\underline{w}, u) := \sum_{y \in E} y \mathbf{1}_{\{u \in B(y; \underline{w})\}} \quad (1.12)$$

Sea $\underline{U} := \{U_i\}_{i \in \mathbb{Z}}$ una sucesión de v.a.i. uniformemente distribuidas en el intervalo $[0, 1]$. Definimos finalmente el proceso como $X_0^{\underline{w}} := F(\underline{w}, U_0)$ y para $t \geq 0$

$$X_t^{\underline{w}} := F(X_{t-1}^{\underline{w}}, \dots, X_0^{\underline{w}}, \underline{w}; U_t) \quad (1.13)$$

Lema 1.2.2. El proceso $X_t^{\underline{w}}$ definido por 1.13 tiene distribución 1.1, es decir, la distribución es compatible con la especificación \mathcal{P} y con la condición de borde \underline{w} .

Demostración. Por la proposición 1.1.2 solo hay que verificar que los intervalos tienen las longitudes correctas, lo cual queda probado por la propiedad (3) de la partición,

$$|B(y; \underline{w})| = \mathcal{P}(y; \underline{w})$$

\square

Veamos ahora algunas propiedades que nos serán útiles mas adelante. Observemos primero que

$$[0, a_k] \subset \bigcup_{l=0}^k B_l(\underline{w}) \quad \forall \underline{w} \in E^{-\mathbb{N}^*} \quad (1.14)$$

Debido a la construcción de la partición, para probar 1.14 alcanza con ver que:

1. $L(\bigcup_{l=0}^k B_l(\underline{w})) \leq 0$
2. $a_k \leq R(\bigcup_{l=0}^k B_l(\underline{w}))$

Ahora, por un lado tenemos que:

$$\begin{aligned} L\left(\bigcup_{l=0}^k B_l(\underline{w})\right) &= L(B_0(\underline{w})) \\ &= L(B_0(1; \underline{w})) \\ &= 0 \end{aligned}$$

Por otro lado:

$$\begin{aligned} R\left(\bigcup_{l=0}^k B_l(\underline{w})\right) &= R\left(\bigcup_{l=0}^k \bigcup_{y \in E} B_l(y; \underline{w})\right) \\ &= R(B_k(q; \underline{w})) \\ &= \sum_{y \in E} a_k(y; \underline{w}) \end{aligned}$$

donde la última igualdad se debe a la propiedad (1) de la partición. Ahora utilizando la definición de a_k se tiene que:

$$\begin{aligned} a_k &= \inf_{\underline{w} \in E^{-\mathbb{N}^*}} \left(\sum_{y \in E} a_k(y; \underline{w}) \right) \\ &\leq \sum_{y \in E} a_k(y; \underline{w}) \\ &= R\left(\bigcup_{l=0}^k B_l(\underline{w})\right) \end{aligned}$$

lo cual termina la prueba. Observando 1.14 obtenemos que si $u \leq a_k$ entonces para saber el valor de $F(\underline{w}; u)$ solo necesitamos tener en cuenta $w_{-1}, w_{-2}, \dots, w_{-k}$ ya que $u \in [0, a_k] \subset \bigcup_{l=0}^k B_l(\underline{w})$. Mas precisamente, de 1.14 podemos deducir que si $\underline{w}, \underline{v} \in E^{\mathbb{Z}}$ son tales que $w_j = v_j \forall j \in [-k, -1]$ entonces se cumple que:

$$[0, a_k] \cap B_l(\underline{w}) = [0, a_k] \cap B_l(\underline{v}) \tag{1.15}$$

De esta manera tenemos que :

$$u \leq a_k \text{ implica que } F(\underline{w}; u) = F(\underline{v}; u) \tag{1.16}$$

para todo $\underline{w}, \underline{v} \in E^{-\mathbb{N}^*}$ que verifiquen que $w_j = v_j \forall j \in [-k, -1]$, y $u \in [0, 1]$.

Esto implica que si $U_t \leq a_k$ entonces es suficiente mirar k coordenadas en el pasado para calcular $X_t^{\underline{w}}$.

1.3. Resultados

Veremos en esta sección algunos resultados válidos para la construcción hecha en las secciones anteriores.

Para cada $\underline{w}, \underline{v} \in E^{-\mathbb{N}^*}$ sea:

$$\tau^{\underline{w}, \underline{v}} := \inf\{n \geq 0 : X_k^{\underline{w}} = X_k^{\underline{v}}, \forall k \geq n\} \quad (1.17)$$

Es claro que en principio $\tau^{\underline{w}, \underline{v}}$ puede ser infinito. Sin embargo existen ciertas condiciones bajo las cuales ese tiempo es casi seguramente finito. La siguiente proposición establece dicho resultado.

Proposición 1.3.1. *Si $\prod_{k \geq 0} a_k > 0$ entonces para todo \underline{w} y \underline{v} en $E^{-\mathbb{N}^*}$ se cumple que*

$$\sum_{n \geq 0} P(\tau^{\underline{w}, \underline{v}} = n) = 1 \quad (1.18)$$

Demostración. Observemos que la propiedad 1.16 implica que

$$\{n \geq 0 : U_{n+k} \leq a_k \forall k \geq 0\} \subseteq \{n \geq 0 : X_k^{\underline{w}} = X_k^{\underline{v}} \forall k \geq n\}$$

y por lo tanto tenemos que

$$\tau^{\underline{w}, \underline{v}} \leq \min\{n \geq 0 : U_{n+k} \leq a_k \forall k \geq 0\} \quad (1.19)$$

En otras palabras $\tau^{\underline{w}, \underline{v}}$ está dominado estocásticamente por el último tiempo de retorno al origen en el proceso “casa de cartas” con transiciones

$$Q(k, k+1) = a_k \quad ; \quad Q(k, 0) = 1 - a_k \quad (1.20)$$

y estado inicial 0, $k \in \mathbb{N}$ (ver Apéndice A). Ahora por el lema A.0.3 la condición de la hipótesis $\prod_{k \geq 0} a_k > 0$ es equivalente a la transitoriedad del proceso “casa de cartas”, lo cual significa que el proceso puede visitar el origen solo una cantidad finita de veces. Esto implica que el último retorno debe ser en tiempo finito con probabilidad 1, lo cual prueba la proposición. \square

Para $-\infty < s < t \leq +\infty$ definimos:

$$\tau[s, t] := \max\{m \leq s : U_k < a_{k-m}, \forall k \in [m, t]\} \quad (1.21)$$

Observar que $\tau[s, t]$ puede ser $-\infty$. Notaremos $\tau[n] = \tau[n, n] := \max\{m \leq n : U_k < a_{k-m}, \forall k \in [m, n]\}$.

Observación 1.3.2. 1. *Fijado $-\infty < s < +\infty$ se cumple que $\tau[s, t]$ es no creciente con t , ya que si $t \leq t'$ se cumple que :*

$$\begin{aligned} \{m \leq s : U_k < a_{k-m}, \forall k \in [m, +\infty]\} &\subseteq \{m \leq s : U_k < a_{k-m}, \forall k \in [m, t']\} \\ &\subseteq \{m \leq s : U_k < a_{k-m}, \forall k \in [m, t]\} \end{aligned}$$

Luego tomando el máximo de cada conjunto tenemos que :

$$\tau[s, +\infty] \leq \tau[s, t'] \leq \tau[s, t] \quad (1.22)$$

como queríamos ver.

2. Fijado $-\infty < s < +\infty$ se cumple que $\tau[s, t]$ es no decreciente con s , en el sentido que

$$[s', t'] \subseteq [\tau[s, t], t] \text{ implica que } \tau[s', t'] \geq \tau[s, t] \quad (1.23)$$

Esto se debe a que como $\tau[s, t] \leq s' \leq t' \leq t$ entonces

$$\max\{m \leq s' : U_k < a_{k-m}, \forall k \in [m, t']\} \geq \max\{m \leq s : U_k < a_{k-m}, \forall k \in [m, t]\}$$

porque $\max\{m \leq s : U_k < a_{k-m}, \forall k \in [m, t]\} \leq s'$ y si m es tal que $U_k < a_{k-m}$ para todo $k \in [m, t]$ entonces también se cumple que $U_k < a_{k-m}$ para todo $k \in [m, t']$ porque $t' \leq t$.

3. Observar también que $\tau[s, t]$ es un tiempo de parada a izquierda para \underline{U} con respecto a $[s, t]$ en el sentido que

$$\{\tau[s, t] \leq j\} \text{ solo depende de los valores } \{U_i : i \in [j, t]\} \quad (1.24)$$

para $j \leq s$.

Lema 1.3.3. Si $\sum_{m \geq 0} \beta_m = \infty$ entonces para cada $-\infty < s \leq t < +\infty$ la variable aleatoria $\tau[s, t]$ es una v.a. "honesta", es decir que:

$$\sum_{i \in \mathbb{Z}} P(\tau[s, t] = i) = 1 \quad (1.25)$$

Si $\prod_{k=0}^{\infty} a_k > 0$ entonces para cada $-\infty < s$ la variable aleatoria $\tau[s, +\infty]$ es honesta, y entonces:

$$\sum_{i \in \mathbb{Z}} P(\tau[s, +\infty] = i) = 1 \quad (1.26)$$

Demostración. Por la definición 1.21 de τ , tenemos que para $j \leq s$

$$\begin{aligned} \tau[s, t] < j &\iff \forall m \in [j, s], \exists n \in [m, t] : W_n^{m-1} = 0 \\ &\iff \forall m \in [j, s], \exists n \in [s, t] : W_n^{m-1} = 0 \\ &\iff \max\{m < s : \forall n \in [s, t], W_n^m > 0\} < j - 1 \\ &\iff \exists n \in [s, t] : W_n^{j-1} = 0 \end{aligned}$$

donde $W_n^{0;m}$ es el estado n -ésimo del proceso "casa de cartas" con matriz de transición $Q := (Q(x, y))_{x, y \in E}$ tal que

$$Q(x, y) = \begin{cases} a_x & \text{si } y = x + 1 \\ 1 - a_x & \text{si } y = 0 \\ 0 & \text{en otro caso} \end{cases}$$

que empieza en tiempo m en el estado 0. Se cumple entonces que

$$\tau[s, t] = 1 + \max\{m \leq s : W_n^{0;m} > 0, \forall n \in [s, t]\} \quad (1.27)$$

Entonces si $m \leq s$ se tiene que:

$$\{\tau[s, t] < m\} \subset \bigcup_{i \in [s, t]} \{W_i^{0; m-1} = 0\} \quad (1.28)$$

Ahora por invariancia bajo traslaciones se tiene que:

$$\begin{aligned} P\left(\bigcup_{i \in [s, t]} \{W_i^{0; m-1} = 0\}\right) &= P\left(\bigcup_{i \in [s, t]} \{W_{i-m+1}^{0; 0} = 0\}\right) \\ &\leq \sum_{i=s}^t P(W_{i-m+1}^{0; 0} = 0) \\ &= \sum_{i=1}^{t-s} P(W_{s-m+i}^{0; 0} = 0) \end{aligned}$$

Utilizando la hipótesis y el lema A.0.3 (a) sabemos que la cadena $W_i^{0; 0}$ es recurrente no positiva, entonces cada uno de los términos $P(W_{s-m+i}^{0; 0} = 0)$ tiende a 0 cuando $m \rightarrow -\infty$. Entonces:

$$\begin{aligned} \lim_{m \rightarrow -\infty} P(\tau[s, t] < m) &\leq \lim_{m \rightarrow -\infty} \sum_{i=1}^{t-s} P(W_{s-m+i}^{0; 0} = 0) \\ &= \sum_{i=1}^{t-s} \lim_{m \rightarrow -\infty} P(W_{s-m+i}^{0; 0} = 0) \\ &= 0 \end{aligned}$$

y por lo tanto se cumple que

$$\sum_{i \in \mathbb{Z}} P(\tau[s, t] = i) = 1$$

Veamos ahora la segunda afirmación del lema. Observemos que:

$$P(\tau[s, \infty] < m) \leq P\left(\bigcup_{i=s-m+1}^{\infty} \{W_i^{0; 0} = 0\}\right) \leq \sum_{i=1}^{\infty} P(W_{s-m+i}^{0; 0} = 0) \quad (1.29)$$

Pero como la cadena es trascendente por hipótesis entonces cada término de la derecha en 1.29 tiende a cero cuando $m \rightarrow -\infty$. De esta manera:

$$\sum_{i \in \mathbb{Z}} P(\tau[s, \infty] = i) = 1$$

□

Sabemos por resultados anteriores que:

$$[0, a_k) \subset \bigcup_{l=0}^k B_l(\underline{w}) = \left[0, \sum_{y \in E} a_k(y; \underline{w})\right) \quad \forall \underline{w} \in E^{-\mathbb{N}^*} \quad (1.30)$$

Como consecuencia de esto tenemos que :

$$[0, a_k) \cap B_l(\underline{w}) = \emptyset \quad \forall l > k \quad (1.31)$$

Definimos entonces

$$B_{l;k}(w_{-1}, \dots, w_{-k}) := [0, a_k) \cap B_l(\underline{w}) \quad (1.32)$$

$$B_{l;k}(y; w_{-1}, \dots, w_{-k}) := [0, a_k) \cap B_l(y; \underline{w}) \quad (1.33)$$

que por 1.31 tiene sentido para todo $l, k \geq 0$. Similarmente a 1.21, para $\underline{u} \in [0, 1]^{\mathbb{Z}}$ definimos

$$\tau[n](\underline{u}) := \text{máx} \{m \leq n : u_j < a_{j-m} \forall j \in [m, n]\} \quad (1.34)$$

Proposición 1.3.4. *Sea y_0 un elemento arbitrario de E . Definimos la función $\Phi : [0, 1]^{\mathbb{Z}} \rightarrow E^{\mathbb{Z}}$ tal que $\Phi(\underline{u}) = \underline{x}$ recursivamente para $j = \tau[n](\underline{u})$ hasta $j = n$ de la siguiente manera: notamos $\tau := \tau[n](\underline{u})$ y definimos:*

$$\begin{aligned} x_\tau &:= \sum_{y \in E} y \mathbf{1}_{\{u_\tau \in B_0(y)\}} \\ x_{\tau+1} &:= \sum_{y \in E} y \mathbf{1}_{\{u_{\tau+1} \in B_0(y) \cup B_{1;1}(y; x_\tau)\}} \\ &\vdots \\ x_n &:= \sum_{y \in E} y \mathbf{1}_{\{u_n \in \bigcup_{j=0}^{n-\tau} B_{j;n-\tau}(y; x_{n-1}, \dots, x_\tau)\}} \end{aligned} \quad (1.35)$$

si $\tau[n](\underline{u})$ es finito, y $x_n = y_0$ si $\tau[n](\underline{u}) = -\infty$ (observar que cada suma en la definición anterior se reduce a un solo término). Entonces

1. Φ está bien definida y es medible.
2. Para cada $n \in \mathbb{Z}$ la componente n -ésima de $\Phi(\underline{u})$ que notaremos $\Phi(\underline{u})_n$ depende solo de los valores u_i en el intervalo $[\tau[n], n]$ si $\tau[n](\underline{u}) > -\infty$ y de $\{u_i : i < n\}$ en otro caso.
3. Si $P(\tau[n](U) > -\infty) = 1$ para todo $n \in \mathbb{Z}$ entonces la ley μ de $\Phi(U)$ es compatible con la especificación \mathcal{P} .

Demostración. Sea

$$A_n := \{\underline{u} \in [0, 1]^{\mathbb{Z}} : \tau[n](\underline{u}) > -\infty\}$$

1. En el conjunto A_n , la consistencia de la definición 1.35 se debe a los siguientes hechos:
 - (i) por definición, $\tau[n] \leq \tau[j]$ para todo $j \in [\tau[n], n]$ y
 - (ii) como vale 1.30 entonces esto muestra que el evento $\{U_n < a_k\}$ se cumple, luego solo hay que mirar los valores x_{n-1}, \dots, x_{n-k} para obtener el valor de x_n . Estos dos hechos implican que para todo valor $k \in [\tau[n], n]$ el valor x_k calculado usando 1.35

con k en lugar de n es el mismo valor que el que se obtiene como parte del cálculo recursivo 1.35 para x_n . La $\mathcal{F}(U_i : i \leq n)$ -medibilidad de x_n se debe directamente a la definición 1.35. Como los conjuntos A_n son $\mathcal{F}(U_i : i \leq n)$ -medibles los mapas x_n continúan siendo medibles luego de “juntarlos”. Concluimos entonces que el mapa $\Phi(\underline{u}) = (x_n : n \in \mathbb{Z})$ esta bien definido y es medible.

2. Se deduce directamente de la definición de la función.
3. Cuando se cumple A_n , 1.31 implica que 1.35 sea

$$x_n = \sum_{y \in E} y \mathbf{1}_{\{u_n \in \bigcup_{k \geq 0} B_k(y; x_{n-1}, \dots, x_{n-k})\}} \quad (1.36)$$

y esto implica que

$$P(\Phi(\underline{U})_n = y | U_i : i < n; A_n) = P(U_n \in \bigcup_{k \geq 0} B_k(y; \Phi(\underline{U})_{n-1}, \dots, \Phi(\underline{U})_{n-k}) | U_i : i < n; A_n) \quad (1.37)$$

Como $\Phi(\underline{U}_j)$ es $\mathcal{F}(U_i : i \leq j)$ -medible y cada U_n es independiente de $\mathcal{F}(U_i : i \leq n)$, si A_n tiene medida total entonces 1.37 es igual a

$$\begin{aligned} \left| \bigcup_{k \geq 0} B_k(y; \Phi(\underline{U})_{n-1}, \dots, \Phi(\underline{U})_{n-k}) \right| &= \sum_{k \geq 0} |B_k(y; \Phi(\underline{U})_{n-1}, \dots, \Phi(\underline{U})_{n-k})| \\ &= \mathcal{P}(y; \Phi(\underline{U})_{n-1}, \dots, \Phi(\underline{U})_{n-k}) \end{aligned}$$

En otras palabras:

$$P(\Phi(\underline{U})_n = y | U_i : i < n) = \mathcal{P}(y; \Phi(\underline{U})_i : i < n) \quad (1.38)$$

Como el lado derecho de 1.38 solo depende de $\Phi(\underline{U})_i$ con $i < n$, entonces el lado izquierdo tambien depende solo de esos valores, por lo tanto la ley de $\Phi(\underline{U})_n$ dado $\{\Phi(\underline{U})_i : i < n\}$ sigue siendo dada por \mathcal{P} , entonces la distribución μ de $\Phi(\underline{U})$ es compatible con \mathcal{P} .

□

Capítulo 2

Procesos de Poisson

En este capítulo comenzaremos con una definición constructiva de un proceso de Poisson. En la cuarta sección veremos algunas propiedades de dicho proceso que nos servirán para probar en la sección 6 que el proceso construido verifica las definiciones clásicas de proceso de Poisson. En las secciones 7 y 8 definiremos procesos de Poisson bi-dimensionales y veremos como construir un proceso de Poisson uno-dimensional a partir de uno bi-dimensional. En las dos últimas secciones introduciremos la noción de superposición de procesos de Poisson y procesos no homogéneos.

2.1. Procesos Puntuales

Sea E un subconjunto de un espacio euclídeo (para nosotros sera suficiente suponer que E esta contenido en $\mathbb{R}^d, d \geq 1$). Queremos distribuir puntos aleatoriamente en el conjunto E y tener una notación conveniente para la función que cuenta el número aleatorio de puntos que caen en un subconjunto acotado $A \subseteq E$.

Sean $\{X_n\}_{n \in \mathbb{N}}$ elementos aleatorios en el conjunto E que representan puntos en el espacio de estados E (i.e. $X_n : \Omega \rightarrow E, \forall n \in \mathbb{N}$). Definimos $\epsilon_{X_n} : \Omega \times \mathcal{P}(E) \rightarrow \{0, 1\}$ de la siguiente manera:

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

donde $\mathcal{P}(E) = \{A \subset E\}$ es el conjunto de los subconjuntos de E .

Definimos la medida de conteo

$$N : \Omega \times \mathcal{P}(E) \rightarrow \mathbb{N}$$

tal que:

$$N(\omega, A) := \sum_{n \in \mathbb{N}} \epsilon_{X_n(\omega)}(A) \tag{2.1}$$

De esta manera $N(A) : \Omega \rightarrow \mathbb{N}$ es una variable aleatoria que cuenta el número de puntos que caen en A . Llamamos a $N : \Omega \times \mathcal{P}(E) \rightarrow \mathbb{N}$ proceso puntual y a $\{X_n\}_{n \in \mathbb{N}}$ puntos.

Observación 2.1.1. Para facilitar la notación suprimiremos la dependencia que el proceso tiene con ω .

Ejemplo 2.1.2. Supongamos que queremos modelar el tiempo y la ubicación en los que ocurren los terremotos. En este caso una posible elección del espacio de estados podría ser $E = [0, \infty) \times \mathbb{R}^2$, y el proceso puntual podría ser representado por:

$$N(\cdot) = \sum_{n \in \mathbb{N}} \epsilon_{(T_n, (L_{n_1}, L_{n_2}))}(\cdot)$$

donde T_n representa el tiempo del n -ésimo terremoto y (L_{n_1}, L_{n_2}) representan la latitud y la longitud en las que ocurre el n -ésimo terremoto.

Para $t > 0$ y $B \subset \mathbb{R}^2$, $N([0, t] \times B)$ podría ser el número de terremotos ocurridos hasta el tiempo t en la región B . Si quisieramos ser mas precisos en la identificación de los terremotos, le podríamos agregar por ejemplo, la intensidad con la que éste ocurre. En este caso el espacio de estados podría ser $[0, \infty) \times \mathbb{R}^2 \times [0, \infty)$, y el proceso puntual sería

$$\tilde{N}(\cdot) = \sum_{n \in \mathbb{N}} \epsilon_{(T_n, (L_{n_1}, L_{n_2}), I_n)}(\cdot)$$

donde I_n indica la intensidad del n -ésimo terremoto.

Notar que la notación es lo suficientemente flexible como para poder hacerle suaves modificaciones si queremos modelar procesos con múltiples puntos en un solo lugar. Por ejemplo, consideremos como antes $\{X_n\}_{n \in \mathbb{N}}$ una sucesión aleatoria de elementos en E (los cuales podrían ahora representar “sitios”), y sea $\{\zeta_n\}_{n \in \mathbb{N}}$ una sucesión de valores enteros no negativos que representan el número aleatorio de puntos que hay en el sitio aleatorio X_n . Luego el proceso puntual con múltiples puntos en sus sitios puede ser representado como:

$$N(\cdot) = \sum_{n \in \mathbb{N}} \zeta_n \epsilon_{X_n}(\cdot)$$

Luego para una región $A \subset E$ tenemos que

$$N(A) = \sum_{n \in \mathbb{N}} \zeta_n \epsilon_{X_n}(A)$$

representa el número de puntos en A , ya que estamos contando el número de puntos en cada sitio de A .

Veamos un ejemplo. Supongamos que a cierto lugar arriban ómnibus de acuerdo con un proceso puntual con puntos $\{X_n\}$, es decir que el n -ésimo ómnibus llega en tiempo X_n . Si ζ_n representa el número de pasajeros que llega en el n -ésimo ómnibus, el total de pasajeros que llegan en tiempo t es:

$$\sum_{n \in \mathbb{N}} \zeta_n \epsilon_{X_n}([0, t])$$

Definición 2.1.3. Sea $N : \Omega \times \mathcal{P}(E) \rightarrow \mathbb{N}$ un proceso puntual. Definimos la media o intensidad de dicho proceso como: $\mu : \mathcal{P}(E) \rightarrow \mathbb{R}$ tal que

$$\mu(A) := \mathbf{E}(N(A)) \quad \forall A \subset \mathcal{P}(E)$$

En otras palabras $\mu(A)$ es el número esperado de puntos en la región A .

2.2. Procesos de Poisson

Sea N un proceso puntual con espacio de estados E . Sea $\mathcal{E} \subset \mathcal{P}(E)$ una σ -álgebra en E , es decir

1. $\mathcal{E} \neq \emptyset$
2. Si $A \in \mathcal{E}$ entonces $A^c \in \mathcal{E}$
3. Si $\{A_n\}_{n \in \mathbb{N}} \subset \mathcal{E}$ entonces $\bigcup_{n \in \mathbb{N}} A_n \in \mathcal{E}$.

Definición 2.2.1. Decimos que el proceso N es un proceso de Poisson con media μ si verifica las siguientes propiedades:

1. Para todo $A \in \mathcal{E}$ se cumple que :

$$P(N(A) = k) = \begin{cases} \frac{\mu(A)^k}{k!} e^{-\mu(A)} & \text{si } \mu(A) < \infty \\ 0 & \text{en otro caso.} \end{cases}$$

2. Si A_1, \dots, A_k son subconjuntos de E en \mathcal{E} disjuntos, entonces $N(A_1), \dots, N(A_k) : \Omega \rightarrow \mathbb{N}$ son variables aleatorias independientes.

En otras palabras N es un proceso de Poisson si el número aleatorio de puntos en el conjunto A , o sea $N(A) : \Omega \rightarrow \mathbb{N}$ tiene distribución de Poisson con parámetro $\mu(A)$, y el número de puntos en regiones disjuntas son v.a. independientes.

Observación 2.2.2. 1. Cuando $E = \mathbb{R}$ la propiedad (2) de la definición anterior se llama propiedad de incrementos independientes, ya que implica que si $t_1 < t_2 < \dots < t_k$ con $t_i \in \mathbb{R} \forall i = 1, \dots, k$ entonces $\{N((t_i, t_{i+1}]), i = 1, \dots, k-1\}$ son variables aleatorias independientes.

2. Cuando la media es un múltiplo de la medida de Lebesgue (es decir que es la longitud cuando $E = \mathbb{R}$, el área si $E = \mathbb{R}^2$, el volumen si $E = \mathbb{R}^3$, etc)decimos que el proceso es homogéneo.

Veamos ahora una definición constructiva de un proceso de Poisson. Comencemos considerando $\{A_i\}_{i \in \mathbb{Z}} \subset \mathcal{P}(\mathbb{R})$ tal que $A_i = [l_i, l_{i+1})$ con $l_i < l_{i+1} \forall i \in \mathbb{Z}$ una partición de \mathbb{R} . Para cada $i \in \mathbb{Z}$ llamamos $|A_i| := l_{i+1} - l_i$ a la longitud del intervalo A_i , y dado

$\lambda \in \mathbb{R}$, $\lambda > 0$ le asignamos a cada intervalo A_i una variable aleatoria $Y_i : \Omega \rightarrow \mathbb{N}$ con distribución de Poisson con parámetro $\lambda|A_i|$, es decir que:

$$P(Y_i = k) = \frac{(\lambda|A_i|)^k}{k!} e^{-\lambda|A_i|}$$

Asumimos que $\{Y_i, i \in \mathbb{Z}\}$ es una familia de variables aleatorias independientes. En segundo lugar asignamos a cada $i \in \mathbb{Z}$ una sucesión de v.a.i. $\{U_{i,j}, j \in \mathbb{N}^*\}$ donde la variable $U_{i,j}$ con $j \in \mathbb{N}^*$ tiene distribución uniforme en el intervalo A_i , es decir:

$$P(U_{i,j} \in A \cap A_i) = \frac{|A \cap A_i|}{|A_i|}, \forall A \subset \mathbb{R}$$

Definimos $S : \Omega \rightarrow \mathbb{R}$ el conjunto aleatorio dado por:

$$S := \bigcup_{i \in \mathbb{Z}} S_i$$

donde $S_i : \Omega \rightarrow \mathbb{R}$ con $i \in \mathbb{Z}$ esta definido por:

$$S_i = \begin{cases} \{U_{i,j} : 1 \leq j \leq Y_i\} & \text{si } Y_i \geq 1 \\ \phi & \text{si } Y_i = 0 \end{cases}$$

(omitimos la dependencia con ω en todas las notaciones).

Finalmente sea $\{\tilde{S}_n\}_{n \in \mathbb{Z}}$ la sucesión ordenada de los puntos de S , donde :

$$\tilde{S}_1 = \text{mín}\{s > 0 : s \in S\}$$

es el primer punto de S que está a la derecha del origen y

$$\tilde{S}_n = \begin{cases} \text{mín}\{s > \tilde{S}_{n-1} : s \in S\} & \text{si } n \geq 2 \\ \text{máx}\{s < \tilde{S}_{n+1} : s \in S\} & \text{si } n \leq 0 \end{cases}$$

Definimos ahora

$$N_S : \Omega \times \mathbb{R} \rightarrow \mathbb{N} \tag{2.2}$$

tal que:

$$N_S(\omega, A) := \sum_{n \in \mathbb{N}} \mathbf{1}_{\{\tilde{S}_n(\omega) \in A\}}$$

para todo $A \subset \mathbb{R}$, y entonces $N_S(A)$ es el número de puntos que hay en el conjunto $S \cap A$. Cuando no se preste a confusión escribiremos $N(A)$ en lugar de $N_S(A)$ cuando nos estemos refiriendo a la variable aleatoria $N_S(A) : \Omega \rightarrow \mathbb{N}$. Al proceso definido anteriormente lo llamaremos proceso de Poisson.

2.3. Definición formal de un proceso puntual

En esta sección analizaremos con mas detalle el espacio de estados de nuestro modelo y veremos que hay ciertos conjuntos en dicho espacio que determinan la distribución de un proceso puntual.

Sea E el espacio donde viven los puntos de nuestro modelo (en nuestro caso E será un subconjunto de \mathbb{R}^d con $d \geq 1$ como dijimos en la primera sección de este capítulo) y supongamos que \mathcal{E} es una σ -álgebra en E (por ejemplo podemos considerar la σ -álgebra generada por los conjuntos abiertos de E o lo que es equivalente, la generada por los “rectángulos” de E). Una manera de modelar una distribución aleatoria de puntos en E es dar $\{X_n\}_{n \in \mathbb{N}}$ elementos aleatorios en E y luego decir que el proceso puntual estocástico es la función de conteo que a la región $A \in \mathcal{E}$ le asigna el número de elementos aleatorios $\{X_n\}$ que están en A , como hicimos en la primer sección de este capítulo. A continuación repetiremos dicha construcción de una manera mas formal. Dados $x \in E$ y $A \in \mathcal{E}$ definimos la medida delta de la siguiente manera:

$$\epsilon_x(A) := \begin{cases} 1 & \text{si } x \in A \\ 0 & \text{si } x \notin A \end{cases}$$

Una medida puntual en E es una medida de la forma

$$m := \sum_i \epsilon_{x_i} \tag{2.3}$$

donde $x_i \in E$ y toda región acotada A contiene solo una cantidad finita de puntos x_i . Luego $m(A)$ es el número de x_i que están en la región A . Denotemos por $M_p := M_p(E)$ al conjunto de todas las medidas puntuales en E y consideremos los subconjuntos de M_p de la forma

$$\{m \in M_p : m(I) = k\}$$

donde $k \geq 0$ e I es un rectángulo acotado en E . Llamaremos $\mathcal{M}_p := \mathcal{M}_p(E)$ a la menor σ -álgebra que contiene dichos conjuntos (donde k e I varían). Podemos ahora definir un proceso puntual como un elemento aleatorio de (M_p, \mathcal{M}_p) , es decir, dado un espacio de probabilidad (Ω, \mathcal{F}, P) donde \mathcal{F} es la σ -álgebra de los eventos, N es un proceso puntual si es un mapa medible $N : (\Omega, \mathcal{F}) \rightarrow (M_p, \mathcal{M}_p)$, lo cual significa que para todo $\Lambda \in \mathcal{M}_p$ se cumple que

$$N^{-1}(\Lambda) := \{\omega \in \Omega : N(\omega) \in \Lambda\} \in \mathcal{F}$$

En particular, si N es un proceso puntual entonces

$$\{N(I) = k\} = N^{-1}(\{m \in M_p : m(I) = k\}) \in \mathcal{F}$$

ya que $\{m \in M_p : m(I) = k\} \in \mathcal{M}_p$.

Consideremos ahora un proceso puntual:

$$N : (\Omega, \mathcal{F}) \rightarrow (M_p, \mathcal{M}_p)$$

y P una medida de probabilidad en (Ω, \mathcal{F}) .

La distribución de N es la medida $P \circ N^{-1} = P(N \in (\cdot))$ en (M_p, \mathcal{M}_p) . Luego, la probabilidad de un evento que depende de N puede ser determinada si conocemos la distribución de N . Definimos las distribuciones de dimensión finita de N como la colección de funciones:

$$p_{I_1, \dots, I_k}(n_1, \dots, n_k) := P(N(I_1) = n_1, \dots, N(I_k) = n_k)$$

donde I_1, \dots, I_k son rectángulos acotados y n_1, \dots, n_k son enteros no negativos.

Teorema 2.3.1. *Las distribuciones de dimensión finita de un proceso puntual N determinan de manera única la distribución $P \circ N^{-1}$ de N .*

Demostración. Dados $k \geq 0$ e I un rectángulo acotado en E consideremos los conjuntos

$$E_{k,I} := \{m \in M_p : m(I) = k\} \quad (2.4)$$

Sea \mathcal{G} la clase de las intersecciones finitas de conjuntos de la forma 2.4, es decir

$$\mathcal{G} := \left\{ \bigcap_{i=1}^n E_{k_i, I_i} : k_1, \dots, k_n \in \mathbb{N}, I_{k_i} \text{ rectángulo acotado } \forall i = 1, \dots, n, n \in \mathbb{N} \right\}$$

Luego como \mathcal{G} es cerrado bajo intersecciones finitas y genera a la σ -álgebra \mathcal{M}_p se tiene que una medida en \mathcal{M}_p está únicamente determinada por los valores que toma en \mathcal{G} (Billingsley, 1986 pág. 38). En particular, $P \circ N^{-1}$ es una medida definida en \mathcal{M}_p y sus valores en \mathcal{G} están dados por las distribuciones de dimensión finita. □

2.4. Propiedades

El objetivo de esta sección es probar que efectivamente la construcción realizada en la sección (2.2) es un proceso de Poisson, es decir que el proceso definido por 2.2 verifica la definición 2.2.1. Comenzaremos viendo algunas propiedades de las variables aleatorias con distribución de Poisson que serán útiles para lograr el objetivo planteado.

Lema 2.4.1. *Si $0 < p_i < 1$ para todo $i \geq 0$ entonces:*

1. $\lim_{N \rightarrow \infty} \prod_{i=0}^N p_i = 0 \iff \sum_{i=0}^{\infty} q_i := \sum_{i=0}^{\infty} (1 - p_i) = \infty$
2. $\prod_{i=0}^{\infty} p_i > 0 \iff \sum_{i=0}^{\infty} q_i := \sum_{i=0}^{\infty} (1 - p_i) < \infty$

Demostración. Decimos que $a_n \sim b_n$ cuando $n \rightarrow \infty$ si $\lim_{n \rightarrow \infty} \frac{a_n}{b_n} = 1$. Entonces si $a_n \sim b_n$ se cumple que $\sum_{n \in \mathbb{N}} a_n < \infty \iff \sum_{n \in \mathbb{N}} b_n < \infty$.

Equivalentemente podemos decir que para cada $\epsilon > 0$ se tiene para todo n suficientemente

grande que $(1 - \epsilon)b_n < a_n < (1 + \epsilon)b_n$.

Utilizando lo anterior y que $-\log(1 - x) \sim x$ cuando $x \rightarrow 0$ tenemos entonces que

$$\begin{aligned} \prod_{i=0}^{\infty} p_i > 0 &\iff \sum_{i=0}^{\infty} -\log(1 - q_i) < \infty \\ &\iff \sum_{i=0}^{\infty} q_i < \infty \end{aligned}$$

Notar que tanto $\sum_{i=0}^{\infty} -\log(1 - q_i) < \infty$ como $\sum_{i=0}^{\infty} q_i < \infty$ implican que $q_i \rightarrow 0$ \square

Lema 2.4.2. 1. Sean X_1, X_2, \dots, X_n variables aleatorias independientes con distribución de Poisson de parámetro $\lambda_i, i = 1, \dots, n$ respectivamente. Se cumple que la variable aleatoria $X_1 + \dots + X_n$ tiene distribución de Poisson de parámetro $\lambda_1 + \dots + \lambda_n$.

2. Sea $\{X_n\}_{n \in \mathbb{N}}$ una sucesión de variables aleatorias independientes con distribución de Poisson de parámetro λ_n respectivamente. Si $\sum_{n \in \mathbb{N}} \lambda_n < \infty$ entonces la variable aleatoria $\sum_{n \in \mathbb{N}} X_n$ tiene distribución de Poisson con parámetro $\sum_{n \in \mathbb{N}} \lambda_n$.

Demostración. 1. Haremos inducción en la cantidad de variables aleatorias n .

Caso $n=2$: Sea $k \in \mathbb{N}$. Usando la fórmula de la probabilidad total tenemos que

$$P(X_1 + X_2 = k) = P(X_1 = k - X_2) = \sum_{h=0}^{\infty} P(X_1 = k - h, X_2 = h)$$

Como X_1 y X_2 son independientes y tienen distribución de Poisson, y observando que $X_1 \geq 0$ tenemos que lo anterior es equivalente a

$$\begin{aligned} \sum_{h=0}^k P(X_1 = k - h)P(X_2 = h) &= \sum_{h=0}^k \frac{\lambda_1^{k-h}}{(k-h)!} e^{-\lambda_1} \frac{\lambda_2^h}{h!} e^{-\lambda_2} \\ &= e^{-(\lambda_1 + \lambda_2)} \frac{\lambda_1^k}{k!} \sum_{h=0}^k C_h^k \left(\frac{\lambda_2}{\lambda_1} \right)^h 1^{h-k} \\ &= e^{-(\lambda_1 + \lambda_2)} \frac{\lambda_1^k}{k!} \left(1 + \frac{\lambda_2}{\lambda_1} \right)^k \\ &= e^{-(\lambda_1 + \lambda_2)} \frac{\lambda_1^k (\lambda_1 + \lambda_2)^k}{k! \lambda_1^k} \\ &= e^{-(\lambda_1 + \lambda_2)} \frac{(\lambda_1 + \lambda_2)^k}{k!} \end{aligned}$$

con lo cual queda probado que $X_1 + X_2$ es una variable aleatoria con distribución de Poisson de parámetro $\lambda_1 + \lambda_2$.

Supongamos ahora que el resultado se cumple para n variables. Entonces utilizando nuevamente la fórmula de probabilidad total y la independencia entre las variables tenemos que:

$$\begin{aligned} P\left(\sum_{i=1}^{n+1} X_i = k\right) &= P\left(\sum_{i=1}^n X_i = k - X_{n+1}\right) \\ &= \sum_{h=0}^k P\left(\sum_{i=1}^n X_i = k - h, X_{n+1} = h\right) \\ &= \sum_{h=0}^k P\left(\sum_{i=1}^n X_i = k - h\right) P(X_{n+1} = h) \end{aligned}$$

Aplicando la hipótesis de inducción y utilizando que X_{n+1} es de Poisson con parámetro λ_{n+1} obtenemos que lo anterior es:

$$\begin{aligned} \sum_{h=0}^k \frac{(\sum_{i=1}^n \lambda_i)^{k-h}}{(k-h)!} e^{-(\sum_{i=1}^n \lambda_i)} \frac{\lambda_{n+1}^h}{h!} e^{-\lambda_{n+1}} &= \\ e^{-(\sum_{i=1}^n \lambda_i)} e^{-\lambda_{n+1}} \sum_{h=0}^k \left(\sum_{i=1}^n \lambda_i\right)^{k-h} \lambda_{n+1}^h \frac{\mathcal{C}_h^k}{k!} &= \\ \frac{(\sum_{i=1}^{n+1} \lambda_i)^k}{k!} e^{-(\sum_{i=1}^{n+1} \lambda_i)} & \end{aligned}$$

con lo cual queda probada la parte (1) del lema.

- Definimos las variables $Y_n : \Omega \rightarrow \mathbb{N}$ con $n \in \mathbb{N}$ tal que $Y_n = X_1 + \dots + X_n$. Sea $Y : \Omega \rightarrow \mathbb{N} \cup \{+\infty\}$ tal que $Y(\omega) = \sum_{i \in \mathbb{N}} X_i(\omega)$.

Consideremos el conjunto $A_\infty = \{\omega \in \Omega : Y(\omega) = \infty\}$ y la variable $\widehat{Y} : \Omega \setminus A_\infty \rightarrow \mathbb{N}$ donde $\widehat{Y} = Y|_{\Omega \setminus A_\infty}$. Observar que si $P(A_\infty) = 0$ entonces las variables Y e \widehat{Y} tienen la misma distribución. Consideremos ahora para cada $k \in \mathbb{N}$ y cada $n \in \mathbb{N}^*$ los conjuntos:

$$\Omega_{n,k} = \left\{ \omega \in \Omega : \sum_{i=1}^n X_i(\omega) = k \text{ y } X_m(\omega) = 0 \text{ si } m > n \right\} \quad (2.5)$$

$$\Omega_{\infty,k} = \{ \omega \in \Omega : \widehat{Y}(\omega) = k \} \quad (2.6)$$

Observemos que $\Omega_{n,k} \subseteq \Omega_{n+1,k}$ para todo $n \in \mathbb{N}$ y que $\bigcup_{n \in \mathbb{N}} \Omega_{n,k} = \Omega_{\infty,k}$. Entonces

utilizando la parte (1) del lema tenemos que :

$$\begin{aligned}
 P(\Omega_{\infty,k}) &= \lim_{n \rightarrow \infty} P(\Omega_{n,k}) \\
 &= \lim_{n \rightarrow \infty} \left[\frac{e^{-\sum_{i \in \mathbb{N}} \lambda_i}}{k!} (\lambda_1 + \dots + \lambda_n)^k \right] \\
 &= \frac{e^{-\sum_{i \in \mathbb{N}} \lambda_i}}{k!} \lim_{n \rightarrow \infty} (\lambda_1 + \dots + \lambda_n)^k \\
 &= \frac{e^{-\sum_{i \in \mathbb{N}} \lambda_i}}{k!} \left(\sum_{i \in \mathbb{N}} \lambda_i \right)^k
 \end{aligned}$$

donde la última igualdad es válida por la continuidad de la función $f(x) = x^k$. Esto implica que la variable \hat{Y} tiene distribución de Poisson de parámetro $\sum_{i \in \mathbb{N}} \lambda_i$. Luego para demostrar el lema solo queda probar que la $P(A_{\infty}) = 0$.

$$\begin{aligned}
 P(A_{\infty}) &= P(\{\omega \in \Omega : Y(\omega) = \infty\}) \\
 &= P(\{\omega \in \Omega : X_i(\omega) \neq 0 \text{ para infinitos } i \in \mathbb{N}\})
 \end{aligned}$$

Consideremos los conjuntos

$$A_n := \{\omega \in \Omega : X_n(\omega) \neq 0\}$$

Entonces

$$\sum_{n \in \mathbb{N}} P(A_n) = \sum_{n \in \mathbb{N}} P(X_n \neq 0) = \sum_{n \in \mathbb{N}} 1 - e^{-\lambda_n} \quad (2.7)$$

Por otro lado sabemos que $\sum_{n \in \mathbb{N}} \lambda_n < \infty$ entonces $\prod_{n \in \mathbb{N}} e^{-\lambda_n} = e^{-\sum_{n \in \mathbb{N}} \lambda_n} > 0$ y usando el lema 2.4.1 deducimos de 2.7 que $\sum_{n \in \mathbb{N}} P(A_n) < \infty$. De esta manera concluimos usando el lema de Borel-Cantelli que $P(A_{\infty}) = 0$ terminando así la demostración del lema. □

Proposición 2.4.3. *Para cada intervalo $A \subset \mathbb{R}$ acotado, la variable aleatoria $N(A) : \Omega \rightarrow \mathbb{N}$ tiene distribución de Poisson con media $\lambda|A|$.*

Demostración. Observemos en primer lugar que como los intervalos A_i con $i \in \mathbb{Z}$ son disjuntos, entonces los intervalos $A \cap A_i$ también lo son. Esto implica que las variables aleatorias $N(A \cap A_i)$ son independientes puesto que, dados $i \neq j$ en \mathbb{Z} y $k, h \in \mathbb{N}$ tenemos que:

$$\begin{aligned}
 P(N(A \cap A_i) = k, N(A \cap A_j) = h) &= P \left(\sum_{n \geq 0} \mathbf{1}_{\{\tilde{S}_n \in A \cap A_i\}} = k, \sum_{n \geq 0} \mathbf{1}_{\{\tilde{S}_n \in A \cap A_j\}} = h \right) = \\
 P(\exists n_1, \dots, n_k \in \mathbb{Z} : \tilde{S}_{n_l} \in A \cap A_i \forall l = 1, \dots, k; \exists m_1, \dots, m_h \in \mathbb{Z} : \tilde{S}_{m_r} \in A \cap A_j, \forall r = 1, \dots, h) &=
 \end{aligned}$$

$P(\exists r_1, \dots, r_k \in \mathbb{N} : U_{i,r_l} \in A \cap A_i \forall l = 1, \dots, k; \exists s_1, \dots, s_h \in \mathbb{N} : U_{j,s_t} \in A \cap A_j, \forall t = 1, \dots, h)$

Ahora teniendo en cuenta que las variables $U_{i,j}$ son independientes obtenemos que lo anterior es:

$$\begin{aligned} & P(U_{i,r_l} \in A \cap A_i, \forall l = 1, \dots, k) P(U_{j,s_t} \in A \cap A_j, \forall t = 1, \dots, h) = \\ & P\left(\sum_{r=1}^{+\infty} \mathbf{1}_{\{U_{i,r} \in A \cap A_i\}} = k\right) P\left(\sum_{r=1}^{+\infty} \mathbf{1}_{\{U_{j,r} \in A \cap A_j\}} = h\right) = \\ & P(N(A \cap A_i) = k) P(N(A \cap A_j) = h). \end{aligned}$$

Entonces $\{N(A \cap A_i)\}_{i \in \mathbb{Z}}$ son variables independientes.

Sea $k \in \mathbb{N}$. Particionando el espacio según los posibles valores de la variable Y_i tenemos que:

$$P(N(A \cap A_i) = k) = \sum_{h=0}^{+\infty} P(Y_i = h, N(A \cap A_i) = k) \quad (2.8)$$

Observemos ahora que si $h < k$ e $Y_i = h < k$ entonces en $S \cap (A \cap A_i)$ van a haber a lo sumo $Y_i = h$ puntos (por la definición de S_i) y por lo tanto $N(A \cap A_i)$ no puede tomar el valor k . Entonces 2.8 resulta ser:

$$\sum_{h=k}^{+\infty} P(Y_i = h, N(A \cap A_i) = k) = \sum_{h=k}^{+\infty} P\left(Y_i = h, \sum_{j=1}^{+\infty} \mathbf{1}_{\{U_{i,j} \in A \cap A_i\}} = k\right) \quad (2.9)$$

Teniendo en cuenta que solo caen en el conjunto S los $U_{i,j}$ con $1 \leq j \leq Y_i$, y que las variables $U_{i,j}$ son independientes de Y_i 2.9 es:

$$\begin{aligned} & \sum_{h=k}^{+\infty} P\left(Y_i = h, \sum_{j=1}^h \mathbf{1}_{\{U_{i,j} \in A \cap A_i\}} = k\right) = \\ & \sum_{h=k}^{+\infty} P(Y_i = h) P\left(\sum_{j=1}^h \mathbf{1}_{\{U_{i,j} \in A \cap A_i\}} = k\right) \end{aligned} \quad (2.10)$$

Afirmación (1): $\mathbf{1}_{\{U_{i,j} \in A \cap A_i\}}$ es una variable con distribución de Bernoulli con parámetro $\frac{|A \cap A_i|}{|A_i|}$.

demostración de la afirmación (1):

$$P(\mathbf{1}_{\{U_{i,j} \in A \cap A_i\}} = 1) = P(U_{i,j} \in A \cap A_i) = \frac{|A \cap A_i|}{|A_i|} \quad \forall j \in \mathbb{N}^*$$

ya que las variables $U_{i,j}$ son uniformes en el intervalo A_i .

Además las variables $\{\mathbf{1}_{\{U_{i,j} \in A \cap A_i\}}\}_{j \in \mathbb{N}^*}$ son independientes ya que las $U_{i,j}$ lo son.

Afirmación (2): $\sum_{j=1}^h \mathbf{1}_{\{U_{i,j} \in A \cap A_i\}}$ es una variable aleatoria con distribución binomial con parámetros h y $\frac{|A \cap A_i|}{|A_i|}$.

demostración de la afirmación (2):

$$P(\sum_{j=1}^h \mathbf{1}_{\{U_{i,j} \in A \cap A_i\}} = k) =$$

$$P(\exists j_1, \dots, j_k \in \{1, \dots, h\} : U_{i,j_l} \in A \cap A_i \forall l = 1, \dots, k, U_{i,j} \notin A \cap A_i \text{ si } j \in \mathbb{N}^* \setminus \{j_1, \dots, j_k\})$$

Entonces, una vez elegidos los j_1, \dots, j_k en \mathbb{N}^* :

$$P(U_{i,j_l} \in A \cap A_i \forall l = 1, \dots, k, U_{i,j} \notin A \cap A_i \text{ si } j \in \mathbb{N}^* \setminus \{j_1, \dots, j_k\}) =$$

$$\begin{aligned} & \prod_{l=1}^k P(U_{i,j_l} \in A \cap A_i) \prod_{j \in \{1, \dots, h\} \setminus \{j_1, \dots, j_k\}} P(U_{i,j} \notin A \cap A_i) = \\ & \left(\prod_{l=1}^k \frac{|A \cap A_i|}{|A_i|} \right) \left(\prod_{j \in \{1, \dots, h\} \setminus \{j_1, \dots, j_k\}} \left(1 - \frac{|A \cap A_i|}{|A_i|} \right) \right) = \\ & \left(\frac{|A \cap A_i|}{|A_i|} \right)^k \left(1 - \frac{|A \cap A_i|}{|A_i|} \right)^{h-k} \end{aligned}$$

Para terminar con la afirmación observar que tenemos \mathcal{C}_k^h maneras diferentes de elegir los $j_1, \dots, j_k \in \{1, \dots, h\}$ por lo tanto:

$$P\left(\sum_{j=1}^h \mathbf{1}_{\{U_{i,j} \in A \cap A_i\}} = k\right) = \mathcal{C}_k^h \left(\frac{|A \cap A_i|}{|A_i|} \right)^k \left(1 - \frac{|A \cap A_i|}{|A_i|} \right)^{h-k}$$

con lo cual queda probada la afirmación (2).

Aplicando ahora la afirmación (2) 2.10 resulta ser:

$$\sum_{h=k}^{+\infty} P(Y_i = h) \mathcal{C}_k^h \left(\frac{|A \cap A_i|}{|A_i|} \right)^k \left(1 - \frac{|A \cap A_i|}{|A_i|} \right)^{h-k} \quad (2.11)$$

y como $Y_i \sim \text{Poisson}(\lambda|A_i|)$ lo anterior es:

$$\begin{aligned} & \sum_{h=k}^{+\infty} \frac{(\lambda|A_i|)^h}{h!} e^{-\lambda|A_i|} \frac{h!}{(h-k)!k!} \left(\frac{|A \cap A_i|}{|A_i|} \right)^k \left(1 - \frac{|A \cap A_i|}{|A_i|} \right)^{h-k} \\ & = \frac{e^{-\lambda|A_i|}}{k!} \left(\frac{|A \cap A_i|}{|A_i|} \right)^k \sum_{h=k}^{+\infty} \frac{\lambda^h |A_i|^h}{(h-k)!} \left(1 - \frac{|A \cap A_i|}{|A_i|} \right)^{h-k} \\ & = \frac{e^{-\lambda|A_i|}}{k!} \left(\frac{|A \cap A_i|}{|A_i|} \right)^k \sum_{u=0}^{+\infty} \frac{\lambda^{u+k} |A_i|^{u+k}}{u!} \left(1 - \frac{|A \cap A_i|}{|A_i|} \right)^u \end{aligned}$$

$$\begin{aligned}
 &= \frac{e^{-\lambda|A_i|}}{k!} \left(\frac{|A \cap A_i|}{|A_i|} \right)^k \lambda^k |A_i|^k \sum_{u=0}^{+\infty} \frac{\lambda^u |A_i|^u}{u!} \left(1 - \frac{|A \cap A_i|}{|A_i|} \right)^u \\
 &= e^{-\lambda|A_i|} \frac{|A \cap A_i|^k}{k!} \lambda^k e^{\lambda(|A_i| - |A \cap A_i|)} \\
 &= \frac{(\lambda|A \cap A_i|)^k}{k!} e^{-\lambda|A \cap A_i|}
 \end{aligned}$$

Entonces $N(A \cap A_i) \sim \text{Poisson}(\lambda|A \cap A_i|) \forall i \in \mathbb{Z}$.

Por último observar que $N(A) = \sum_{i \in \mathbb{Z}} N(A \cap A_i)$ en virtud de que los A_i son disjuntos, por lo tanto $N(A)$ es suma de v.a.i. con distribución de Poisson de media $\lambda|A \cap A_i|$. Luego, como $\biguplus_{i \in \mathbb{Z}} A_i = \mathbb{R}$ entonces $\biguplus_{i \in \mathbb{Z}} (A \cap A_i) = A$ y por lo tanto $\sum_{i \in \mathbb{Z}} |A \cap A_i| = |A|$. Aplicando finalmente el lema 2.4.2 queda probado que $N(A) \sim \text{Poisson}(\lambda|A|)$ como queríamos. \square

Proposición 2.4.4. *Para cada familia finita de intervalos disjuntos B_1, \dots, B_L las variables aleatorias $\{N(B_i)\}_{i=1}^L$ son independientes y tienen distribución de Poisson con media $\lambda|B_l|, l = 1, \dots, L$ respectivamente.*

Demostración. Sea $i \in \mathbb{Z}$ fijo y veamos que dado $N(A_i) = h_i$ las v.a. $\{N(B_l \cap A_i) : l = 1, \dots, L\}$ tienen distribución multinomial, es decir que para valores enteros $k_{l,i} \geq 0$ tales que $\sum_{l=1}^L k_{l,i} = h_i$ y valores reales $b_{l,i} \geq 0$ tales que $\sum_{l=1}^L b_{l,i} = 1$ se cumple que:

$$P(N(B_l \cap A_i) = k_{l,i}, l = 1, \dots, L | N(A_i) = h_i) = h_i! \frac{\prod_{l=1}^L b_{l,i}^{k_{l,i}}}{\prod_{l=1}^L k_{l,i}!} \quad (2.12)$$

Definimos

$$B_{L+1} := \left(\bigcup_{l=1}^L B_l \right)^c \quad (2.13)$$

y consideramos

$$b_{l,i} := \frac{|B_l \cap A_i|}{|A_i|} \quad \forall 1 \leq l \leq L \quad (2.14)$$

$$b_{L+1,i} := 1 - \sum_{l=1}^L b_{l,i} \quad (2.15)$$

Sean $k_{l,i} \geq 0, l = 1, \dots, L+1$ enteros arbitrarios tales que $\sum_{l=1}^{L+1} k_{l,i} = h_i$, entonces

$$\begin{aligned}
 &P(N(B_l \cap A_i) = k_{l,i}, l = 1, \dots, L+1 | N(A_i) = h_i) = \\
 &\frac{P(N(B_l \cap A_i) = k_{l,i}, l = 1, \dots, L+1, N(A_i) = h_i)}{P(N(A_i) = h_i)}
 \end{aligned} \quad (2.16)$$

Observemos que

$$\biguplus_{l=1}^{L+1} B_l \cap A_i = \left(\biguplus_{l=1}^L B_l \cap A_i \right) \cup (B_{L+1} \cap A_i) = A_i$$

donde esta última igualdad es por la definición de B_{L+1} .
 Luego si $N(B_l \cap A_i) = k_{l,i}$, $\forall l = 1, \dots, L+1$ entonces

$$N(A_i) = \sum_{l=1}^{L+1} N(B_l \cap A_i) = \sum_{l=1}^{L+1} k_{l,i} = h_i \quad (2.17)$$

Debido a 2.17 tenemos que 2.16 es:

$$\frac{P(N(B_l \cap A_i) = k_{l,i}, l = 1, \dots, L+1)}{P(N(A_i) = h_i)} \quad (2.18)$$

Como los intervalos $\{B_l \cap A_i\}_{l=1, \dots, L+1}$ son disjuntos resulta que las variables aleatorias $\{N(B_l \cap A_i)\}_{l=1, \dots, L+1}$ son independientes, luego teniendo en cuenta esto y la proposición 2.4.3 resulta que 2.18 es:

$$\begin{aligned} \prod_{l=1}^{L+1} \frac{P(N(B_l \cap A_i) = k_{l,i})}{P(N(A_i) = h_i)} &= \frac{h_i!}{e^{-\lambda|A_i|} (\lambda|A_i|)^{h_i}} \prod_{l=1}^{L+1} \frac{e^{-\lambda|B_l \cap A_i|} (\lambda|B_l \cap A_i|)^{k_{l,i}}}{k_{l,i}!} \\ &= \frac{h_i!}{\lambda^{h_i} |A_i|^{h_i}} e^{\lambda|A_i|} e^{-\lambda \sum_{l=1}^{L+1} |B_l \cap A_i|} \lambda^{\sum_{l=1}^{L+1} k_{l,i}} \prod_{l=1}^{L+1} \frac{|B_l \cap A_i|^{k_{l,i}}}{k_{l,i}!} \\ &= \frac{h_i!}{\prod_{l=1}^{L+1} k_{l,i}!} \prod_{l=1}^{L+1} b_{l,i}^{k_{l,i}} \end{aligned}$$

como queríamos probar. Luego como las variables $\{N(A_i)\}_{i \in \mathbb{Z}}$ son independientes se deduce de 2.12 que $\{N(B_l \cap A_i) \mid l = 1, \dots, L, i \in \mathbb{Z}\}$ son variables independientes y que $N(B_l \cap A_i) \sim \text{Poisson}(\lambda|B_l \cap A_i|)$.

Para terminar solo queda observar que dado $l = 1, \dots, L$ se verifica que $N(B_l) = \sum_{i \in \mathbb{Z}} N(B_l \cap A_i)$ y que $\sum_{i \in \mathbb{Z}} \lambda|B_l \cap A_i| = \lambda|B_l|$ con lo cual el lema 2.4.2 nos dice que:

$$N(B_l) \sim \text{Poisson}(\lambda|B_l|) \quad \forall l = 1, \dots, L$$

□

Proposición 2.4.5. *Para todo intervalo A la distribución condicional de los puntos en $S \cap A$ dado $N(A) = n$ es la misma que la de n v.a.i. distribuídas uniformemente en el intervalo A .*

Demostración. Usaremos el teorema 2.3.1. Sean B_1, \dots, B_L una partición de A y sean n_1, \dots, n_L enteros no negativos tales que $n_1 + \dots + n_L = n$. Por 2.12 sabemos que:

$$P(N(B_l) = n_l, l = 1, \dots, L \mid N(A) = n) = \frac{n!}{n_1! \dots n_L!} \prod_{l=1}^L \left(\frac{|B_l|}{|A|} \right)^{n_l}$$

Entonces, por el teorema 2.3.1 es suficiente probar que si U_1, \dots, U_n son v.a.i. uniformemente distribuidas en A y $M(B) = \sum_i \mathbf{1}_{\{U_i \in B\}}$ se cumple que

$$P(M(B_l) = n_l, l = 1, \dots, L) = \frac{n!}{n_1! \dots n_L!} \prod_{l=1}^L \left(\frac{|B_l|}{|A|} \right)^{n_l}$$

Pero teniendo en cuenta que las variables aleatorias $M(B_l)$ son v.a. binomiales con parámetros n y $\frac{|B_l|}{|A|}$ (esto se debe a la afirmación (2) hecha en la proposición 2.4.3) tenemos que:

$$\begin{aligned} P(M(B_l) = n_l : l = 1, \dots, L) &= \\ \mathcal{C}_{n_1}^n P(U_{n_1} \in B_1 \forall i = 1, \dots, n_1) \cdot \mathcal{C}_{n_2}^{n-n_1} P(U_{n_1} \in B_2 \forall i = n_1 + 1, \dots, n_1 + n_2) \dots \\ \dots \mathcal{C}_{n_L}^{n-(n_1+n_2+\dots+n_{L-1})} P(U_{n_i} \in B_L \forall i = n_1 + \dots + n_{L-1}, \dots, n) &= \\ \frac{n!}{(n-n_1)!n_1!} \left(\frac{|B_1|}{|A|} \right)^{n_1} \cdot \frac{(n-n_1)!}{(n-n_1-n_2)!n_2!} \left(\frac{|B_2|}{|A|} \right)^{n_2} \dots \\ \dots \frac{(n-n_1-\dots-n_{L-1})!}{n_L!} \left(\frac{|B_L|}{|A|} \right)^{n_L} &= \\ \frac{n!}{n_1!n_2! \dots n_L!} \prod_{l=1}^L \left(\frac{|B_l|}{|A|} \right)^{n_l} & \end{aligned}$$

□

Observación 2.4.6. De las proposiciones 2.4.4 y 2.4.5 se deduce que la construcción del proceso 2.2 no depende de la elección de los intervalos A_i que conforman la partición de \mathbb{R} .

2.5. Resultados

En esta sección presentaremos una construcción alternativa de un proceso de Poisson uno-dimensional que utilizaremos luego para probar ciertos resultados acerca de la distribución de los tiempos entre eventos sucesivos de un proceso de Poisson.

Sean T_1 una variable aleatoria con distribución exponencial de parámetro λ y $N_1(\cdot)$ un proceso de Poisson con parámetro λ , independiente de la variable T_1 . Definimos el proceso $N(\cdot)$ por:

$$N(A) := \mathbf{1}_{\{T_1 \in A\}} + N_1(A - T_1) \tag{2.19}$$

para $A \subset \mathbb{R}$, donde $A - t = \{x \in \mathbb{R} : x + t \in A\}$. En otras palabras, el proceso $N(\cdot)$ es obtenido fijando el primer evento según la v.a. T_1 y agregando luego de ese instante un proceso de Poisson independiente.

El siguiente resultado nos asegura que el proceso definido antes verifica las propiedades de proceso de Poisson.

Teorema 2.5.1. *El proceso definido por 2.19 es un proceso de Poisson.*

Demostración. En virtud del teorema 2.3.1 es suficiente ver que el proceso 2.19 verifica la proposición 2.4.4. Por lo tanto tenemos que calcular:

$$P(N(B_l) = k_l : l = 1, \dots, L) \quad (2.20)$$

donde B_1, \dots, B_L son intervalos arbitrarios en \mathbb{R} , k_1, \dots, k_L valores enteros no nulos, $L \geq 1$. Para facilitar la notación, probaremos el caso $L = 1$ y $B_1 = [a, c]$ y se deja como ejercicio la generalización al caso $L \geq 1$.

Condicionando el valor del proceso a los valores que toma la variable T_1 tenemos que:

$$\begin{aligned} P(N([a, c]) = k) &= P(N([a, c]) = k, T_1 < a) \\ &+ P(N([a, c]) = k, T_1 \in [a, c]) + P(N([a, c]) = k, T_1 > c) \end{aligned} \quad (2.21)$$

Veamos primero el caso $k = 0$. En este caso los términos de 2.21 resultan ser:

$$\begin{aligned} P(N([a, c]) = 0, T_1 < a) &= \int_0^a \lambda e^{-\lambda t} P(N_1([a-t, c-t]) = 0) dt \\ &= \int_0^a \lambda e^{-\lambda t} e^{-\lambda(c-a)} dt \\ &= \int_0^a \lambda e^{-\lambda(c-a+t)} dt \\ &= e^{-\lambda(c-a)} (1 - e^{-\lambda a}) \end{aligned} \quad (2.22)$$

Por otro lado

$$P(N([a, c]) = 0, T_1 \in [a, c]) = 0 \quad (2.23)$$

ya que por la definición del proceso $N(\cdot)$ el hecho de que $T_1 \in [a, c]$ implica que $N([a, c]) \geq 1$. Por último tenemos que

$$\begin{aligned} P(N([a, c]) = 0, T_1 > c) &= \int_c^\infty \lambda e^{-\lambda t} dt \\ &= e^{-\lambda c} \end{aligned} \quad (2.24)$$

ya que como $T_1 > c$ entonces $N_1([a-t, c-t]) = 0$. Luego en el caso $k = 0$ obtenemos sumando 2.22, 2.23 y 2.24 que:

$$P(N([a, c]) = 0) = e^{-\lambda(c-a)} \quad (2.25)$$

como queríamos ver.

Veamos ahora el caso $k > 0$. Estudiando cada caso, análogamente a lo hecho para el caso $k = 0$ tenemos que:

$$\begin{aligned} P(N([a, c]) = k, T_1 < a) &= \int_0^a \lambda e^{-\lambda t} P(N_1([a-t, c-t]) = k) dt \\ &= \int_0^a \lambda e^{-\lambda t} e^{-\lambda(c-a)} \frac{\lambda^k (c-a)^k}{k!} dt \\ &= \frac{\lambda^k (c-a)^k}{k!} e^{-\lambda(c-a)} (1 - e^{-\lambda a}) \end{aligned} \quad (2.26)$$

Para calcular el valor del segundo término de la derecha de 2.21, observemos que como $T_1 \in [a, c]$ entonces $a - t \leq 0$ para todo $t \in [a, c]$, y entonces:

$$\begin{aligned} P(N([a, c]) = k, T_1 \in [a, c]) &= \int_a^c \lambda e^{-\lambda t} P(N_1([0, c-t]) = k-1) dt \\ &= \int_a^c \lambda e^{-\lambda t} e^{-\lambda(c-t)} \frac{(\lambda(c-t))^{k-1}}{(k-1)!} dt \\ &= \frac{\lambda^k}{k!} e^{-\lambda c} (c-a) \end{aligned} \quad (2.27)$$

Por último, el último término se anula ya que como $T_1 > c$ entonces $[a-t, c-t] \subset (-\infty, 0]$ para todo $t \geq c$. De esta manera concluimos que

$$P(N([a, c]) = k) = \frac{\lambda^k (c-a)^k}{k!} e^{-\lambda(c-a)} \quad (2.28)$$

para todo $k \in \mathbb{N}$, con lo cual queda probado el teorema. \square

Corolario 2.5.2. *Sea $N(\cdot)$ un proceso de Poisson uno-dimensional de parámetro λ . Consideremos S_1, S_2, \dots los tiempos ordenados de ocurrencia de los eventos del proceso. Entonces las variables aleatorias $T_n := S_n - S_{n-1}$, $n \geq 2$ son independientes e idénticamente distribuidas con distribución exponencial de parámetro λ .*

Demostración. El teorema 2.5.1 nos dice que un proceso de Poisson $N_0(\cdot)$ se puede construir colocando el instante del primer evento según una v.a. exponencial T_1 y a continuación un proceso de Poisson $N_1(\cdot)$ independiente de T_1 . De la misma manera, el proceso $N_1(\cdot)$ puede ser construido con una v.a. exponencial T_2 y un proceso de Poisson $N_2(\cdot)$ independiente. Iterando esta construcción tenemos que los instantes entre eventos sucesivos es una sucesión de v.a.i. T_1, T_2, \dots distribuidas exponencialmente. \square

2.6. Propiedad de los estadísticos de orden

Como vimos, una manera de construir un proceso de Poisson en una región acotada A es “salpicar” una cantidad de Poisson de v.a.i. con distribución uniforme sobre A . Luego vimos que condicionados a que en la región A hayan n puntos, esos puntos están distribuidos como n v.a.i. uniformemente distribuidas en A . En esta sección supondremos que el espacio de estados E es $[0, +\infty)$ y veremos algunos resultados válidos en este caso.

Definición 2.6.1. *Sean X_1, \dots, X_n v.a.i.i.d. en un espacio Ω . Definimos nuevas variables aleatorias $X_{(1)}, \dots, X_{(n)}$ con dominio Ω que llamaremos estadísticos de orden de las variables X_1, \dots, X_n de la siguiente manera:*

$$\begin{aligned} X_{(1)}(\omega) &:= \min\{X_1(\omega), \dots, X_n(\omega)\} \\ X_{(i)}(\omega) &:= \min(\{X_1(\omega), \dots, X_n(\omega)\} \setminus \{X_{(1)}(\omega), \dots, X_{(i-1)}(\omega)\}) \text{ si } i = 2, \dots, n \end{aligned}$$

Veremos ahora cual es la densidad conjunta de los estadísticos de orden cuando las v.a son uniformes en el intervalo $(0, t) \subset \mathbb{R}$.

Lema 2.6.2. Sean U_1, \dots, U_n v.a.i.i.d. con distribución uniforme en el intervalo $(0, t)$, y sean $U_{(1)}, \dots, U_{(n)}$ los estadísticos de orden asociados. Entonces la densidad conjunta de los estadísticos de orden es:

$$f_{U_{(1)}, \dots, U_{(n)}}(u_1, \dots, u_n) = \begin{cases} \frac{n!}{t^n} & \text{si } 0 < u_1 < \dots < u_n < t \\ 0 & \text{en otro caso} \end{cases} \quad (2.29)$$

Demostración. Sea $\Pi := \{f : \{1, \dots, n\} \rightarrow \{1, \dots, n\} : f \text{ es función biyectiva}\}$ la familia de las permutaciones de n elementos. Si $\pi \in \Pi$ entonces $(U_{(1)}, \dots, U_{(n)}) = (U_{\pi(1)}, \dots, U_{\pi(n)})$ en el conjunto $\{U_{\pi(1)} < \dots < U_{\pi(n)}\}$. Luego para cada función acotada $g(u_1, \dots, u_n)$ tenemos que

$$\mathbf{E}(g(U_{(1)}, \dots, U_{(n)})) = \sum_{\pi \in \Pi} \mathbf{E}(g(U_{\pi(1)}, \dots, U_{\pi(n)})) \mathbf{1}_{\{U_{\pi(1)} < \dots < U_{\pi(n)}\}}$$

Como la densidad conjunta de $U_{\pi(1)}, \dots, U_{\pi(n)}$ es:

$$f_{U_{\pi(1)}, \dots, U_{\pi(n)}}(u_1, \dots, u_n) = f_{U_1, \dots, U_n}(u_1, \dots, u_n)$$

y U_1, \dots, U_n son uniformes tenemos que

$$f_{U_{\pi(1)}, \dots, U_{\pi(n)}}(u_1, \dots, u_n) = \begin{cases} \frac{1}{t^n} & \text{si } (u_1, \dots, u_n) \in [0, 1]^n \\ 0 & \text{en otro caso} \end{cases}$$

y entonces

$$\begin{aligned} \mathbf{E}(g(U_{(1)}, \dots, U_{(n)})) &= \sum_{\pi \in \Pi} \int_{\{0 < u_1 < \dots < u_n < t\}} g(u_1, \dots, u_n) \frac{1}{t^n} du_1 \dots du_n \\ &= \int_{[0, 1]^n} g(u_1, \dots, u_n) \left(\frac{n!}{t^n} \mathbf{1}_{\{u_1 < \dots < u_n\}}(u_1, \dots, u_n) \right) du_1 \dots du_n \end{aligned}$$

lo cual muestra que la densidad de $(U_{(1)}, \dots, U_{(n)})$ es como queríamos probar

$$\frac{n!}{t^n} \mathbf{1}_{\{u_1 < \dots < u_n\}}(u_1, \dots, u_n)$$

□

Probaremos ahora que un proceso de Poisson homogéneo en $[0, +\infty)$ verifica la propiedad de los estadísticos de orden, lo cual significa que, condicionado a que hayan n puntos en $(0, t]$ las ubicaciones de esos puntos están distribuídas como los estadísticos de orden de una muestra de n v.a. distribuídas uniformemente en el intervalo $(0, t)$.

Teorema 2.6.3. Si N es un proceso de Poisson homogéneo en $[0, +\infty)$ con media λ entonces, condicionado a que $N((0, t]) = n$, la cantidad de puntos del proceso en el intervalo $[0, t]$ en orden creciente esta distribuída como los estadísticos de orden de una muestra de tamaño n de v.a. uniformemente distribuídas en el intervalo $[0, t]$, es decir que la distribución condicionada del vector (S_1, \dots, S_n) dado que $N([0, t]) = n$ es la misma que la ley de $(U_{(1)}, \dots, U_{(n)})$, donde esto último son los estadísticos de orden de las variables (U_1, \dots, U_n) uniformemente distribuídas en $[0, t]$.

Demostración. Supongamos que $a_1 < b_1 < a_2 < b_2 < \dots < a_n < b_n < t$, entonces

$$\begin{aligned} P(S_i \in (a_i, b_i], i = 1, \dots, n \mid N((0, t]) = n) = \\ \frac{P(a_i < S_i < b_i, \forall i = 1, \dots, n, N((0, t]) = n)}{P(N((0, t]) = n)} \end{aligned} \quad (2.30)$$

Ahora, concentrándonos solamente en el numerador de 2.30 y aplicando la proposición 2.4.4 tenemos que dicho numerador es:

$$\begin{aligned} P\left(N((0, a_1]) = 0, N((a_1, b_1]) = 1, \bigcap_{i=1}^{n-1} [N((b_i, a_{i+1}]) = 0, N((a_{i+1}, b_{i+1}]) = 1], N((b_n, t]) = 0\right) \\ = e^{-\lambda a_1} e^{-\lambda(b_1 - a_1)} \lambda(b_1 - a_1) \left[\prod_{i=1}^{n-1} e^{-\lambda(a_{i+1} - b_i)} e^{-\lambda(b_{i+1} - a_{i+1})} \lambda(b_{i+1} - a_{i+1}) \right] e^{-\lambda(t - b_n)} \\ = e^{-\lambda t} \lambda^n \prod_{i=1}^n (b_i - a_i) \end{aligned}$$

Entonces 2.30 resulta ser:

$$e^{-\lambda t} \lambda^n \prod_{i=1}^n (b_i - a_i) \frac{n!}{e^{-\lambda t} (\lambda t)^n} = \frac{n!}{t^n} \prod_{i=1}^n (b_i - a_i) \quad (2.31)$$

Dividiendo por el término $\prod_{i=1}^n (b_i - a_i)$ que aparece en 2.31 y haciendo $b_i \downarrow a_i \forall i = 1, \dots, n$, el lado izquierdo de la ecuación 2.30 se convierte en la densidad condicional dado $N((0, t]) = n$, entonces:

$$f_{S_1, \dots, S_n \mid N((0, t])}(a_1, \dots, a_n) = \frac{n!}{t^n}$$

que por el lema 2.6.2 es lo que queríamos probar. \square

2.7. Definiciones alternativas de procesos de Poisson

Al comienzo de este capítulo vimos una definición clásica de un proceso de Poisson. Inmediatamente después dimos una definición constructiva del mismo. En esta sección queremos comparar la definición constructiva con la clásica.

Consideremos procesos puntuales definidos en toda la recta real \mathbb{R} . Notaremos $N(t)$ para $N((0, t])$.

Definición 2.7.1. 1. Decimos que un proceso uno-dimensional $N(\cdot)$ tiene incrementos estacionarios si la distribución de la variable $N([s, t])$ es la misma que la de la variable $N([s + r, t + r]) \forall r \in \mathbb{R}$.

2. Decimos que el proceso $N(\cdot)$ tiene incrementos independientes si para $A, B \subset \mathbb{R}$ tales que $A \cap B = \emptyset$ se verifica que las variables $N(A)$ y $N(B)$ son independientes.

Definición 2.7.2. Sea $f : \mathbb{R} \rightarrow \mathbb{R}$ una función. Decimos que la función f es un $o(h)$ si:

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

Daremos a continuación dos definiciones de proceso de Poisson y probaremos que son equivalentes.

Definición 2.7.3. Decimos que un proceso $N(\cdot) : \Omega \times \mathcal{P}(\mathbb{R}) \rightarrow \mathbb{N}$ es un proceso de Poisson de parámetro λ si:

1. $N(0) = 0$
2. El proceso tiene incrementos estacionarios e independientes.
3. Las variables aleatorias $N([s, t]) : \Omega \rightarrow \mathbb{N}$ tienen distribución de Poisson de parámetro $\lambda(t - s)$.

Definición 2.7.4. Decimos que un proceso $N(\cdot) : \Omega \times \mathcal{P}(\mathbb{R}) \rightarrow \mathbb{N}$ es un proceso de Poisson de parámetro λ si:

1. $N(0) = 0$
2. El proceso tiene incrementos estacionarios e independientes.
3. $P(N([t, t + h]) = 1) = \lambda h + o(h)$.
4. $P(N([t, t + h]) \geq 2) = o(h)$.

Observación 2.7.5. Cada una de las definiciones anteriores define unívocamente un proceso como consecuencia del teorema 2.3.1 (ya que se está diciendo cual es la distribución de los intervalos $[s, t]$).

Proposición 2.7.6. Las definiciones 2.7.3 y 2.7.4 son equivalentes.

Demostración. Veamos primero que 2.7.3 implica 2.7.4.

Sabemos que la variable $N([t, t + h])$ tiene distribución de Poisson con parámetro λh debido a la proposición 2.4.3. Luego:

$$P(N([t, t + h]) = 1) = \lambda h e^{-\lambda h} = \lambda h \left(1 + \sum_{i=1}^{\infty} \frac{(-\lambda h)^i}{i!} \right) \quad (2.32)$$

usando el desarrollo en serie de $e^{-\lambda h}$. Luego si notamos $f(h) = \lambda h \sum_{i=1}^{\infty} \frac{(-\lambda h)^i}{i!}$ tenemos que:

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

y por lo tanto f es un $o(h)$ concluyendo que:

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

Haciendo un razonamiento análogo tenemos que:

$$\begin{aligned} P(N([t, t+h]) \geq 2) &= \sum_{k=2}^{\infty} P(N([t, t+h]) = k) \\ &= \sum_{k=2}^{\infty} e^{-\lambda h} \frac{(\lambda h)^k}{k!} \\ &= e^{-\lambda h} h^2 \sum_{k=2}^{\infty} \frac{\lambda^k h^{k-2}}{k!} \end{aligned}$$

Luego tenemos que:

$$\lim_{h \rightarrow 0} \frac{P(N([t, t+h]) \geq 2)}{h} = \lim_{h \rightarrow 0} e^{-\lambda t} h \sum_{k=2}^{\infty} \frac{\lambda^k h^{k-2}}{k!} = 0$$

y por lo tanto

$$P(N([t, t+h]) \geq 2) = o(h)$$

Queda ver ahora que 2.7.4 implica 2.7.3. Notemos $N(t)$ el número total de eventos ocurridos hasta el tiempo t y sea $P_n(t) = P(N(t) = n)$. Podemos deducir una ecuación diferencial para $P_0(t)$ de la siguiente manera:

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

donde las dos últimas igualdades se deben a las condiciones (3) y (4) de la definición 2.7.4. Luego:

$$\begin{aligned} \lim_{h \rightarrow 0} \frac{P_0(t+h) - P_0(t)}{h} &= \lim_{h \rightarrow 0} \frac{-\lambda h P_0(t) - P_0(t) o(h)}{h} \\ &= \lim_{h \rightarrow 0} \left(-\lambda P_0(t) + P_0(t) \frac{o(h)}{h} \right) \\ &= -\lambda P_0(t) \end{aligned}$$

Por lo tanto hemos deducido que :

$$P_0'(t) = -\lambda P_0(t)$$

lo cual es equivalente a :

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

Integrando obtenemos que:

$$\ln(P_0(t)) = -\lambda t + c$$

para cierta constante $c \in \mathbb{R}$ y entonces

$$P_0(t) = ke^{-\lambda t}$$

donde $k \in \mathbb{R}$. Como $P_0(0) = P(N(0) = 0) = 1$ entonces tenemos que

$$P_0(t) = e^{-\lambda t} \tag{2.33}$$

Análogamente, para $n \geq 1$ tenemos que:

$$\begin{aligned} P_n(t+h) &= P(N(t+h) = n) \\ &= P(N(t) = n, N(t+h) - N(t) = 0) \\ &+ P(N(t) = n-1, N(t+h) - N(t) = 1) \\ &+ P(N(t+h) = n, N(t+h) - N(t) \geq 2) \end{aligned}$$

Observemos ahora que :

$$P(N(t+h) = n, N(t+h) - N(t) \geq 2) = P(N(t+h) = n, N(h) \geq 2) \leq P(N(h) \geq 2)$$

Entonces, aplicando la condición (4) de la definición 2.7.4 tenemos que:

$$\begin{aligned} \lim_{h \rightarrow 0} \frac{P(N(t+h) = n, N(t+h) - N(t) \geq 2)}{h} &\leq \\ \lim_{h \rightarrow 0} \frac{P(N(h) \geq 2)}{h} &= 0 \end{aligned}$$

y entonces concluimos que $P(N(t+h) = n, N(t+h) - N(t) \geq 2)$ es un $o(h)$. Luego usando la independencia entre los incrementos del proceso obtenemos que

$$\begin{aligned} P_n(t+h) &= P(N(t) = n)P(N(t+h) - N(t) = 0) \\ &+ P(N(t) = n-1)P(N(t+h) - N(t) = 1) + o(h) \\ &= P_n(t)(1 - \lambda h) + P_{n-1}(t)\lambda h + \hat{o}(h) \end{aligned}$$

con lo que deducimos que

$$\begin{aligned} \lim_{h \rightarrow 0} \frac{P_n(t+h) - P_n(t)}{h} &= \lim_{h \rightarrow 0} \left(-\lambda P_n(t) + \lambda P_{n-1}(t) + \frac{\hat{o}(h)}{h} \right) \\ &= -\lambda P_n(t) + \lambda P_{n-1}(t) \end{aligned}$$

Entonces

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

o equivalentemente

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

lo que es lo mismo que:

$$\frac{d}{dt} \left(e^{\lambda t} P_n(t) \right) = e^{\lambda t} \lambda P_{n-1}(t) \quad (2.34)$$

Ahora por 2.33 tenemos que

$$\frac{d}{dt} \left(e^{\lambda t} P_1(t) \right) = e^{\lambda t} \lambda P_0(t) = \lambda$$

Entonces

$$P_1(t) = (\lambda t + \widehat{c}) e^{-\lambda t}$$

con $\widehat{c} \in \mathbb{R}$ y como $P_1(0) = P(N(0) = 1) = 0$ entonces

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

Razonaremos ahora por inducción para probar que

$$P_n(t) = \frac{(\lambda t)^n}{n!} e^{-\lambda t}, \forall n \in \mathbb{N}$$

Supongamos entonces que la igualdad anterior es válida para $n - 1$. Por 2.34 sabemos que

$$\frac{d}{dt} \left(e^{\lambda t} P_n(t) \right) = e^{\lambda t} \lambda P_{n-1}(t)$$

entonces por hipótesis de inducción resulta que

$$\frac{d}{dt} \left(e^{\lambda t} P_n(t) \right) = \lambda \frac{(\lambda t)^{n-1}}{(n-1)!}$$

Integrando obtenemos que

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

con $\widetilde{c} \in \mathbb{R}$. Recordando que $P_n(0) = P(N(0) = n) = 0$ obtenemos que

$$P_n(t) = \frac{(\lambda t)^n}{n!} e^{-\lambda t}$$

como queríamos ver. Resulta entonces que el número de eventos en cualquier intervalo de longitud t tiene distribución de Poisson con parámetro λt lo cual completa la prueba. \square

Teorema 2.7.7. *El proceso*

$$N(\cdot) : \Omega \times \mathcal{P}(\mathbb{R}) \rightarrow \mathbb{N}$$

definido en 2.2 es un proceso de Poisson.

Demostración. Por la proposición 2.4.3 sabemos que la distribución de los puntos en el intervalo $[s, t]$ es de Poisson con parámetro $\lambda(t-s)$ y por la proposición 2.4.4 sabemos que el proceso tiene incrementos independientes y estacionarios, por lo tanto dicho proceso verifica las propiedades de la definición 2.7.3. \square

2.8. Procesos de Poisson en dos o mas dimensiones

En esta sección vamos a construir un proceso de Poisson en dos dimensiones. Esta construcción será útil en los siguientes capítulos para construir procesos de Poisson uno-dimensionales no homogéneos y superposición de procesos de Poisson. Construiremos de manera análoga a lo hecho en dimensión uno, un conjunto aleatorio de puntos en \mathbb{R}^2 . La misma construcción que haremos a continuación sirve para construir procesos en \mathbb{R}^d con $d \geq 2$.

Recordemos la constucción realizada en las primeras secciones de este capítulo. Comenzamos considerando una partición de \mathbb{R}^2 en rectángulos finitos A_i (a modo de ejemplo, A_i podría ser $[n, n+1] \times [m, m+1]$ con $n, m \in \mathbb{Z}$). Denotamos por $|A_i|$ el área del rectángulo A_i . A cada $i \in \mathbb{Z}$ le asignamos una variable aleatoria $Y_i : \Omega \rightarrow \mathbb{N}$ con distribución de Poisson de parámetro $\lambda|A_i|$ para cierto $\lambda \in \mathbb{R}^+$ y asumimos que estas variables son independientes. Luego, para cada $i \in \mathbb{Z}$ consideramos una sucesión de v.a.i. $\{U_{i,j}\}_{j \geq 1}$ distribuídas uniformemente en el rectángulo A_i . Definimos a continuación el conjunto aleatorio

$$S := \bigcup_{i \in \mathbb{Z}} \left(\bigcup_{j=1}^{Y_i} \{U_{i,j}\} \right)$$

con la convención de que $\bigcup_{j=1}^0 \{U_{i,j}\} = \phi$. En otras palabras lo que hacemos es colocar en cada A_i una cantidad Y_i de puntos distribuídos de manera uniforme. Hasta ahora repetimos el procedimiento hecho para dimensión uno. La diferencia ahora es que no hay una manera satisfactoria de ordenar los puntos en \mathbb{R}^2 , pero eso no será importante. Sea ahora $A \subset \mathbb{R}^2$ un conjunto medible. Definimos el proceso de Poisson bidimensional :

$$M : \Omega \times \mathcal{P}(\mathbb{R}^2) \rightarrow \mathbb{N} \tag{2.35}$$

tal que:

$$M(\omega, A) = M_{S(\omega)}(A) := \text{cantidad de puntos en el conjunto } S(\omega) \cap A$$

Entonces $M(A) : \Omega \rightarrow \mathbb{N}$ cuenta la cantidad de puntos en el conjunto $S \cap A$. Para evitar confusiones usaremos la letra M para los procesos de Poisson bidimensionales y la letra N para los uno-dimensionales.

Las siguientes propiedades se prueban con las mismas ideas usadas para el caso uno-dimensional, por lo cual omitiremos sus demostraciones.

Proposición 2.8.1. *Para cada conjunto finito $A \subset \mathbb{R}^2$ medible, la variable aleatoria $M(A) : \Omega \rightarrow \mathbb{N}$ tiene distribución de Poisson de parámetro $\lambda|A|$.*

Proposición 2.8.2. *Para cada familia finita de conjuntos disjuntos medibles B_1, \dots, B_L en \mathbb{R}^2 las variables aleatorias $M(B_l) : \Omega \rightarrow \mathbb{N}$ con $l = 1, \dots, L$ tienen distribución de Poisson con parámetro $\lambda|B_l|$ y son independientes.*

Proposición 2.8.3. *Para cualquier conjunto medible $A \subset \mathbb{R}^2$ la distribución condicional de los puntos de $S \cap A$ dado $M(A) = n$ es la misma que la de n v.a.i. distribuídas uniformemente en el conjunto A .*

Ejemplo 2.8.4. *Dada la construcción en 2 dimensiones podemos calcular la ley de la variable aleatoria*

$$V : \Omega \rightarrow \mathbb{R}$$

definida por

$$V(\omega) := \inf\{|x| : x \in S(\omega)\}$$

que mide la distancia del punto que esta mas cerca del origen.

Sea $B(0, b) := \{(x, y) \in \mathbb{R}^2 : \sqrt{x^2 + y^2} < b\}$ la bola de centro el origen y radio $b \in \mathbb{R}^+$.

Entonces:

$$P(V > b) = P(\{\omega \in \Omega : V(\omega) > b\}) \quad (2.36)$$

Observemos ahora que pedir que $V > b$ es pedir que el ínfimo de las distancias de los puntos de $S(\omega)$ al origen sea mayor que $b \in \mathbb{R}^+$, lo cual es equivalente a pedir que en $B(0, b)$ no haya puntos de $S(\omega)$. Entonces 2.36 es equivalente a :

$$P(M(B(0, b)) = 0)$$

Pero sabemos que la variable $M(B(0, b)) : \Omega \rightarrow \mathbb{N}$ tiene distribución de Poisson con parámetro $\lambda(\text{área}(B(0, b))) = \lambda b^2 \pi$. Entonces resulta que :

$$P(V > b) = e^{-\lambda \pi b^2}$$

2.9. Proyecciones

En esta sección mostraremos una forma de construir procesos de Poisson uno-dimensionales a partir de procesos bidimensionales.

Sea $M(\cdot) : \Omega \times \mathcal{P}(\mathbb{R}^2) \rightarrow \mathbb{N}$ un proceso de Poisson bidimensional con parámetro $\lambda = 1$. Queremos construir un proceso de dimensión uno $N(\cdot) : \Omega \times \mathcal{P}(\mathbb{R}) \rightarrow \mathbb{N}$ con parámetro λ como función de $M(\cdot)$. Definimos para cada intervalo $I \subset \mathbb{R}$:

$$N(I) : \Omega \rightarrow \mathbb{N} \text{ tal que } N(I) := M(I \times [0, \lambda]) \quad (2.37)$$

Entonces el número de puntos del proceso $N(\cdot)$ en el intervalo I es el número de puntos del proceso $M(\cdot)$ en el rectángulo $I \times [0, \lambda]$.

La siguiente proposición muestra que el proceso definido por 2.37 es efectivamente un proceso de Poisson de dimensión uno.

Proposición 2.9.1. *El proceso $N(\cdot) : \Omega \times \mathcal{P}(\mathbb{R}) \rightarrow \mathbb{N}$ definido por 2.37 es un proceso de Poisson uno-dimensional de parámetro λ .*

Demostración. Observemos en primer lugar que la variable aleatoria $N([a, b]) : \Omega \rightarrow \mathbb{N}$ tiene distribución de Poisson con parámetro $\lambda(b - a)$ por la proposición 2.8.1. Veamos en segundo lugar que el proceso tiene incrementos independientes. Sean $I, J \subset \mathbb{R}$ tal que $I \cap J = \emptyset$. Entonces como el proceso $M(\cdot)$ es de Poisson (por lo tanto tiene incrementos independientes) y $(I \times [0, \lambda]) \cap (J \times [0, \lambda]) = \emptyset$ tenemos que:

$$\begin{aligned} P(N(I) = k, N(J) = h) &= P(M(I \times [0, \lambda]) = k, M(J \times [0, \lambda]) = h) \\ &= P(M(I \times [0, \lambda]) = k) P(M(J \times [0, \lambda]) = h) \\ &= P(N(I) = k) P(N(J) = h) \end{aligned}$$

lo cual significa que tiene incrementos independientes.

Para terminar veamos que tiene incrementos estacionarios. Sea $[a, b] \subset \mathbb{R}$ un intervalo y $k \in \mathbb{N}$, entonces:

$$\begin{aligned} P(N([a, b]) = k) &= P(M([a, b] \times [0, \lambda]) = k) = \\ &= \frac{(\lambda(b - a))^k}{k!} e^{-\lambda(b-a)} \end{aligned} \tag{2.38}$$

donde esta última igualdad se debe a la proposición 2.8.1.

Por otro lado, dado $r \in \mathbb{R}$, haciendo un razonamiento análogo tenemos que:

$$\begin{aligned} P(N([a + r, b + r]) = k) &= P(M([a + r, b + r] \times [0, \lambda]) = k) = \\ &= \frac{(\lambda(b - a))^k}{k!} e^{-\lambda(b-a)} \end{aligned} \tag{2.39}$$

Luego, como 2.38 y 2.39 son iguales, el proceso tiene incrementos estacionarios. De esta manera se verifica para el proceso $N(\cdot)$ la definición 2.7.3. \square

Una pregunta que puede surgir es si dos puntos de un proceso bidimensional pueden ser proyectados en un mismo punto de un proceso uno-dimensional. La respuesta a dicha pregunta esta dada en el siguiente lema.

Lema 2.9.2. *Sea $I \subset \mathbb{R}$ un intervalo finito. El evento “dos puntos de un proceso de Poisson bidimensional son proyectados sobre un mismo punto del intervalo I ” tiene probabilidad 0.*

Demostración. Supongamos sin pérdida de generalidad que $I = [0, 1]$. Particionamos I en intervalos de longitud $\delta < 1$: $I_n^\delta := ((n - 1)\delta, n\delta]$. Entonces $\{I_n^\delta\}_{n=1}^{\lceil \frac{|I|}{\delta} \rceil}$ es una partición de I . Ahora tenemos que:

$$\begin{aligned} P(\text{dos puntos sean proyectados en un mismo punto}) &\leq \\ &P(\text{dos puntos pertenecen al mismo intervalo } I_n^\delta) = \\ &P\left(\bigcup_{n=1}^{\lceil \frac{|I|}{\delta} \rceil} \{M(I_n^\delta \times [0, \lambda]) \geq 2\}\right) \end{aligned} \tag{2.40}$$

Aplicando la proposición 2.8.1 y usando que el área de $I_n^\delta \times [0, \lambda]$ es $\lambda\delta$ para todo n tenemos que 2.40 está acotada superiormente por:

$$\sum_{n=1}^{\lfloor \frac{|I|}{\delta} \rfloor} \left(\sum_{j=2}^{+\infty} \frac{(\lambda\delta)^j}{j!} e^{-\lambda\delta} \right) = \left\lfloor \frac{|I|}{\delta} \right\rfloor e^{-\lambda\delta} \sum_{j=2}^{+\infty} \frac{(\lambda\delta)^j}{j!}$$

Sea ahora $f : \mathbb{R} \rightarrow \mathbb{R}$ tal que $f(\delta) = e^{-\lambda\delta} \sum_{j=2}^{+\infty} \frac{(\lambda\delta)^j}{j!}$. Se cumple entonces que $\lim_{\delta \rightarrow 0} \frac{f(\delta)}{\delta} = 0$ y por lo tanto f es un $o(\delta)$. Luego tenemos que:

$$P(\text{dos puntos sean proyectados en un mismo punto}) \leq \frac{|I|}{\delta} o(\delta) \rightarrow 0 \text{ cuando } \delta \rightarrow 0$$

lo cual termina la demostración. \square

Observación 2.9.3. *La construcción de procesos unidimensionales a partir de bidimensionales tiene la ventaja que nos permite construir simultáneamente procesos con diferentes parámetros, de tal manera que uno tenga siempre mayor o igual cantidad de puntos que el otro, como se ve en el siguiente lema.*

Lema 2.9.4. *Sea $M(\cdot)$ un proceso de Poisson bidimensional, y sean $\lambda_1, \lambda_2 \in \mathbb{R}^+$ tales que $\lambda_1 \geq \lambda_2$. Definimos los procesos de Poisson N_i con $i = 1, 2$ tales que :*

$$N_i(I) = M(I \times [0, \lambda_i])$$

para $i = 1, 2$ y para todo $I \subset \mathbb{R}$. Entonces se cumple que:

$$N_1(I) \geq N_2(I) \quad \forall I \subset \mathbb{R}$$

Demostración. Por la definición 2.37 es claro que N_1 proyecta mayor o igual cantidad de puntos que N_2 . \square

2.10. Superposición de procesos de Poisson

Supongamos que a un banco llegan clientes hombres de acuerdo a un proceso de Poisson de parámetro λp y clientes mujeres de acuerdo con un proceso de Poisson de parámetro $\lambda(1-p)$ para cierto $0 < p < 1$. Una manera de modelar el proceso de llegadas, distinguiendo los arribos según sexo es construir un proceso de Poisson bidimensional $M(\cdot)$ y definir los procesos $N_{1(\cdot)}$ y $N_{2(\cdot)}$ de la siguiente manera:

$$N_1(I) : \Omega \rightarrow \mathbb{N} \text{ tal que } N_1(I) := M(I \times [0, \lambda p]) \quad (2.41)$$

$$N_2(I) : \Omega \rightarrow \mathbb{N} \text{ tal que } N_2(I) := M(I \times [\lambda p, \lambda]) \quad (2.42)$$

para todo $I \subset \mathbb{R}$. De esta manera los puntos que están en la faja $\mathbb{R} \times [0, \lambda]$ son proyectados e indican los instantes de llegada de los clientes, independientemente del sexo de cada uno, los puntos que provienen de la faja $\mathbb{R} \times [0, \lambda p]$ los marcamos con 1 y determinan

los momentos de llegada de los clientes hombres, y los puntos que provienen de la faja $\mathbb{R} \times [\lambda p, \lambda]$ los marcamos con un 0 e indican los momentos de llegada de los clientes mujeres. El proceso $N(\cdot)$ tal que $N(I) := N_1(I) + N_2(I)$ lo podemos escribir como $N(I) = M(I \times [0, \lambda])$, por lo tanto es un proceso de Poisson de parámetro λ .

Sean S_n con $n \geq 1$ los tiempos de llegada de los clientes independientemente del sexo, es decir S_1 es el tiempo de arribo al banco del primer cliente y S_n es el tiempo de arribo al banco del n -ésimo cliente. Cada uno de los puntos anteriores tiene una marca que es 0 o 1 dependiendo de si es hombre o mujer respectivamente. Sean las variables $G(S_i) : \Omega \rightarrow \{0, 1\}$ con $i \geq 1$ tal que

$$G(S_i) := \begin{cases} 1 & \text{si } S_i \text{ está marcado con 1} \\ 0 & \text{si } S_i \text{ está marcado con 0} \end{cases} \quad (2.43)$$

Proposición 2.10.1. *Las variables aleatorias $G(S_i) : \Omega \rightarrow \{0, 1\}$ son i.i.d. con ley :*

$$P(G(S_i) = 1) = p \quad P(G(S_i) = 0) = 1 - p$$

para todo $i \geq 1$.

Demostración. Construiremos para empezar n v.a.i.i.d uniformes en el rectángulo $[0, t] \times [0, \lambda]$. Sean $\{V_n\}_{n \geq 1}$ y $\{W_n\}_{n \geq 1}$ dos sucesiones independientes de v.a. uniformemente distribuidas en el intervalo $[0, 1]$ e independientes entre si. Para cada n , sea $\pi_n : \{1, 2, \dots, n\} \rightarrow \{1, 2, \dots, n\}$ una función biyectiva (es decir que π_n es una permutación de n elementos) definida por: para $i = 1, \dots, n - 1$ se cumple que

$$V_{\pi_n(i)} \leq V_{\pi_n(i+1)}$$

En otras palabras, si ordenamos de manera creciente las primeras n variables V_1, \dots, V_n entonces $\pi_n(i)$ es el subíndice de la i -ésima variable.

Para cada n fijo, construimos una sucesión de v.a. en $[0, t] \times [0, \lambda]$ de la siguiente manera: definimos $U_i : \Omega \rightarrow [0, t] \times [0, \lambda]$ tal que

$$U_i := (tV_{\pi_n(i)}, \lambda W_i) \quad \forall i = 1, \dots, n \quad (2.44)$$

Veamos que las variables $\{U_i\}_{i=1, \dots, n}$ son independientes y distribuidas uniformemente en el rectángulo $[0, t] \times [0, \lambda]$. Como las variables V_i y W_j son variables independientes tenemos que:

$$\begin{aligned} P(U_i \in [a, b] \times [c, d]) &= P(tV_{\pi_n(i)} \in [a, b], \lambda W_i \in [c, d]) \\ &= P(tV_{\pi_n(i)} \in [a, b]) P(\lambda W_i \in [c, d]) \\ &= P\left(\frac{a}{t} \leq V_{\pi_n(i)} \leq \frac{b}{t}\right) P\left(\frac{c}{\lambda} \leq W_i \leq \frac{d}{\lambda}\right) \end{aligned}$$

Luego, como conocemos la distribución de las variables V_i y de las W_j , es fácil ver que :

$$U_i \sim \text{Uniforme}([0, t] \times [0, \lambda]) \quad \forall i = 1, \dots, n$$

Queda por ver que son independientes, lo cual se deduce fácilmente del hecho de que $\{V_i\}_{i \geq 1}$ y $\{W_j\}_{j \geq 1}$ son familias de v.a.i. que además son independientes entre si. Entonces construimos $\{U_i\}_{i=1, \dots, n}$ una familia de n v.a.i.i.d. uniformes en $[0, t] \times [0, \lambda]$. Observar que si ordenamos las variables $\{U_i\}_{i=1, \dots, n}$ de manera creciente de acuerdo con la abscisa ($V_{\pi_n(i)}$), entonces W_i es la ordenada del i-ésimo de los U_j .

Sea ahora $L \in \mathbb{N}$ fijo y para $1 \leq k \leq L$ sean $a_k \in \{0, 1\}$ arbitrarios. Definimos el conjunto $A_L := \{G(S_1) = a_1, \dots, G(S_L) = a_L\}$. Usando una versión discreta del teorema 2.3.1, para probar que las variables $G(S_i)$ con $i \geq 1$ tienen distribución de Bernoulli de parámetro p alcanza con probar que :

$$P(A_L) = p^{\sum a_k} (1 - p)^{\sum (1 - a_k)}$$

Sea $t > 0$ fijo. Entonces

$$P(A_L) = \sum_{n=L}^{+\infty} P(A_L, N((0, t)) = n) + P(A_L, N((0, t)) < L) \quad (2.45)$$

Ahora

$$P(A_L, N((0, t)) = n) = P(A_L | N((0, t)) = n) P(N((0, t)) = n)$$

Por definición $N((0, t)) = M([0, t] \times [0, \lambda])$ y por la proposición 2.8.3 la distribución de los puntos en ese rectángulo dado que $M([0, t] \times [0, \lambda]) = n$ es la misma que la distribución de n v.a.i.i.d. uniformes en el rectángulo. Para $1 \leq k \leq L$ sean

$$I_k = \begin{cases} [0, p \lambda] & \text{si } a_k = 1 \\ [p \lambda, \lambda] & \text{si } a_k = 0 \end{cases}$$

Entonces, usando la construcción 2.44 y la proposición 2.8.3 tenemos que:

$$\begin{aligned} P(A_L | N((0, t)) = n) &= P(A_L | M([0, t] \times [0, \lambda]) = n) \\ &= P(U_{\pi_n(k)} \in [0, t] \times I_k, 1 \leq k \leq n) \\ &= P(\lambda W_k \in I_k, 1 \leq k \leq n) \\ &= p^{\sum a_k} (1 - p)^{\sum (1 - a_k)} \end{aligned}$$

donde la última igualdad se debe a que :

$$P(\lambda W_k \in I_k) = \frac{|I_k|}{\lambda} = \begin{cases} p & \text{si } a_k = 1 \\ 1 - p & \text{si } a_k = 0 \end{cases}$$

luego como esta igualdad no depende de n tenemos que:

$$\begin{aligned} P(A_L) &= \sum_{n=L}^{+\infty} p^{\sum a_k} (1 - p)^{\sum (1 - a_k)} P(N((0, t)) = n) + P(A_L, N((0, t)) < L) \\ &= p^{\sum a_k} (1 - p)^{\sum (1 - a_k)} \sum_{n=L}^{+\infty} P(N((0, t)) = n) + P(A_L, N((0, t)) < L) \\ &= p^{\sum a_k} (1 - p)^{\sum (1 - a_k)} P(N((0, t)) \geq L) + P(A_L, N((0, t)) < L) \end{aligned}$$

Por la ley de los grandes números para un proceso de Poisson uno-dimensional [1] se cumple que:

$$\frac{N((0, t))}{t} \text{ converge a } \lambda$$

por lo tanto tenemos que:

$$\lim_{t \rightarrow +\infty} P(N((0, t)) \geq L) = 1 \text{ y } \lim_{t \rightarrow +\infty} P(A_L, N((0, t)) < L) = 0$$

Luego

$$P(A_L) = p^{\sum a_k} (1 - p)^{\sum (1 - a_k)}$$

como queríamos ver. Concluimos entonces que las variables $G(S_i)$, $i \geq 1$ son i.i.d. con distribución de Bernoulli de parámetro p . \square

Análogamente a lo hecho en la sección (2.5), veamos ahora una construcción alternativa de un proceso de Poisson bi-dimensional $M(\cdot)$ con parámetro 1 en la faja $[0, \infty) \times [0, \lambda]$. Esta construcción será utilizada en la prueba del siguiente teorema.

Sean T_1 una v.a. exponencial de parámetro λ , W_1 una v.a. uniforme en el intervalo $[0, 1]$ y $M_1(\cdot)$ un proceso de Poisson bi-dimensional con parámetro 1 definido en la faja $[0, \infty) \times [0, \lambda]$. Asumimos independencia entre T_1 , W_1 y $M_1(\cdot)$. Definimos el proceso $M(\cdot)$ como:

$$M(A) := \mathbf{1}_{\{(T_1, W_1) \in A\}} + M_1(A - T_1) \tag{2.46}$$

donde $A \subset \mathbb{R}^2$ y $A - t = \{(x, y) \in \mathbb{R}^2 : (x + t, y) \in A\}$.

Argumentos similares a los utilizados en el teorema 2.5.1, el corolario 2.5.2, y la proposición 2.10.1, prueban el siguiente teorema.

Teorema 2.10.2. *Sea $M(\cdot)$ un proceso de Poisson bidimensional con parámetro 1. Sean $\{S_i\}_{i \geq 1}$ los tiempos ordenados de ocurrencia de los eventos en la tira $[0, \lambda]$ y sean $\{W_i\}_{i \geq 1}$ las segundas coordenadas de dichos eventos. Entonces $\{S_{i+1} - S_i\}_{i \geq 1}$ son v.a.i.i.d. distribuídas exponencialmente con parámetro λ y $\{W_i\}_{i \geq 1}$ son v.a.i.i.d. con distribución uniforme en $[0, \lambda]$. Mas aún $\{S_{i+1} - S_i\}_{i \geq 1}$ y $\{W_i\}_{i \geq 1}$ son independientes.*

2.11. Procesos no homogéneos

Sea $\lambda : \mathbb{R} \rightarrow \mathbb{R}$ una función no negativa y derivable a trozos. Supongamos que para cada intervalo finito $I \subset \mathbb{R}$ la función λ tiene una cantidad finita de discontinuidades en I . Queremos construir un proceso puntual con incrementos independientes y “tasa instantánea” $\lambda(t)$, esto significa que queremos un proceso $N(\cdot) : \Omega \times \mathcal{P}(\mathbb{R}) \rightarrow \mathbb{N}$ tal que en los puntos de continuidad de la función λ se verifique que:

$$P(N([t, t + h]) = 1) = h\lambda(t) + o(h) \tag{2.47}$$

$$P(N([t, t + h]) \geq 2) = o(h) \tag{2.48}$$

Para eso, consideremos un proceso bidimensional $M(\cdot) : \Omega \times \mathcal{P}(\mathbb{R}^2) \rightarrow \mathbb{N}$ y definimos el proceso uno-dimensional $N(\cdot) : \Omega \times \mathcal{P}(\mathbb{R}) \rightarrow \mathbb{N}$ tal que

$$N(I) := M(F(I)) \quad \forall I \subset \mathbb{R} \quad (2.49)$$

donde $F(I) := \{(x, y) \in \mathbb{R}^2 : x \in I, 0 \leq y \leq \lambda(x)\}$. De esta manera $N(\cdot)$ es el proceso que proyecta los puntos de $M(\cdot)$ que están “por debajo” de la función λ .

Lema 2.11.1. *El proceso 2.49 verifica las condiciones 2.47 y 2.48*

Demostración. Sea $t \in \mathbb{R}$ un punto de continuidad de la función λ . Por la definición de $N(\cdot)$ se cumple que:

$$P(N([t, t+h]) = 1) = P(M(F([t, t+h]))) = 1)$$

Sean

$$y_0 := y_0(t, h) = \inf\{\lambda(x) : x \in [t, t+h]\}$$

y

$$y_1 := y_1(t, h) = \sup\{\lambda(x) : x \in [t, t+h]\}$$

Entonces es claro que:

$$M([t, t+h] \times [0, y_0]) \leq M(F([t, t+h])) \leq M([t, t+h] \times [0, y_1])$$

Por otro lado, como la función λ es continua en t tenemos que :

$$\lim_{h \rightarrow 0} y_0(t, h) = \lambda(t) \quad y \quad \lim_{h \rightarrow 0} y_1(t, h) = \lambda(t)$$

Además por derivabilidad $y_1(t, h) - y_0(t, h) = O(h)$, donde $O(h)$ indica una función de h tal que $\lim_{h \rightarrow 0} \frac{O(h)}{h}$ está acotado superior e inferiormente. Tenemos entonces que:

$$P(M([t, t+h] \times [0, y_1]) - M([t, t+h] \times [0, y_0]) > 1) = o(h)$$

De esta manera

$$\begin{aligned} & P(M(F([t, t+h])) = 1) = \\ & P(M(F([t, t+h])) = 1, M([t, t+h] \times [0, y_1]) - M([t, t+h] \times [0, y_0]) = 0) + \\ & P(M(F([t, t+h])) = 1, M([t, t+h] \times [0, y_1]) - M([t, t+h] \times [0, y_0]) \geq 1) \end{aligned} \quad (2.50)$$

El primer término de 2.50 es igual a :

$$P(M([t, t+h] \times [0, y_0]) = 1, M([t, t+h] \times [0, y_1]) - M([t, t+h] \times [0, y_0]) = 0)$$

y como el proceso $M(\cdot)$ tiene incrementos independientes eso nos da:

$$\begin{aligned} & P(M([t, t+h] \times [0, y_0]) = 1) P(M([t, t+h] \times [0, y_1]) - M([t, t+h] \times [0, y_0]) = 0) = \\ & h y_0(t, h) e^{-h y_0(t, h)} (1 - o(h)) = \end{aligned}$$

$$hy_0(t, h) e^{-hy_0(t, h)} - hy_0(t, h) e^{-hy_0(t, h)} o(h)$$

El segundo término de 2.50 está acotado por :

$$P(M([t, t+h] \times [0, y_1]) - M([t, t+h] \times [0, y_0]) \geq 1) = o(h) + h(y_1 - y_0)$$

Entonces:

$$P(M(F(t, t+h) = 1)) = hy_0(t, h) e^{-hy_0(t, h)} - hy_0(t, h) e^{-hy_0(t, h)} o(h) + h(y_1 - y_0) = h\lambda(t) + o(h)$$

□

Observemos ahora que el proceso de Poisson construido anteriormente tiene la propiedad de que el número de eventos en un intervalo cualquiera es una variable de Poisson con parámetro igual al área por debajo de la función $\lambda(t)$ en ese intervalo, es decir que

$$P(N((0, t]) = n) = e^{-\mu(t)} \frac{(\mu(t))^n}{n!}$$

donde $\mu(t) = \int_0^t \lambda(s) ds$. Esto es fácil de ver ya que

$$P(N((0, t]) = n) = P(M(F([0, t])) = n)$$

y sabemos que $M(F([0, t]))$ es una variable de Poisson con parámetro el área de la región $F([0, t])$ lo cual no es otra cosa que $\int_0^t \lambda(s) ds$.

Análogamente a lo establecido en la observación 2.9.3, la definición 2.49 tiene la ventaja que nos permite construir conjuntamente dos o mas procesos con tasas diferentes y la siguiente propiedad:

Proposición 2.11.2. Sean $\lambda_1(t)$ y $\lambda_2(t)$ dos funciones continuas a trozos que satisfacen la condición

$$\lambda_1(t) \leq \lambda_2(t) \quad \forall t \in \mathbb{R}$$

Entonces es posible construir conjuntamente procesos de Poisson no homogéneos $N_1(\cdot)$ y $N_2(\cdot)$ de tasas $\lambda_1(t)$ y $\lambda_2(t)$ respectivamente tales que para cada intervalo $I \subset \mathbb{R}$ se cumpla que

$$N_1(I) \leq N_2(I)$$

Demostración. Se deduce directamente de la definición de proceso no homogéneo. □

Capítulo 3

Cadenas de Markov de tiempo continuo

En este capítulo consideraremos cadenas de Markov de tiempo continuo. Al igual que en el caso de las cadenas de tiempo discreto, éstas están caracterizadas por la propiedad Markoviana de que dado el estado presente, el futuro es independiente del pasado.

En las primeras secciones definiremos cadenas de Markov de tiempo continuo y veremos como construir un proceso con dichas propiedades. En la cuarta sección estableceremos dos conjuntos de ecuaciones diferenciales -backward y forward equations- que describen la ley de probabilidades del sistema. El resto del capítulo está destinado a mostrar resultados y ejemplos interesantes de la teoría.

3.1. Procesos de Markov de salto puro

Definición 3.1.1. Sea $\{X_t\}_{t \geq 0}$ un proceso estocástico de tiempo continuo (es decir $t \in [0, T]$ con $T \in \mathbb{R}$ fijo), que toma valores en un conjunto numerable E . Decimos que el proceso $\{X_t\}_{t \geq 0}$ es una cadena de Markov de tiempo continuo si $\forall s, t \geq 0$ y $\forall i, j, x_u \in E$ con $0 \leq u \leq s$ se cumple que

$$P(X_{t+s} = j | X_s = i, X_u = x_u \forall 0 \leq u < s) = P(X_{t+s} = j | X_s = i) \quad (3.1)$$

En otras palabras, una cadena de Markov de tiempo continuo (CMTC) es un proceso estocástico que verifica la propiedad markoviana, es decir, que la probabilidad condicional de un futuro estado en el tiempo $t + s$, dado el estado presente en el tiempo s y todos los estados pasados, solo depende del presente estado y es independiente del pasado.

Definición 3.1.2. Si $\{X_t\}_{t \geq 0}$ es una CMTC y además verifica que

$$P(X_{t+s} = j | X_s = i) \quad (3.2)$$

es independiente de s entonces se dice que la CMTC tiene probabilidades de transición estacionarias u homogéneas.

A continuación construiremos un proceso $\{X_t\}$ con $t \in \mathbb{R}$ a valores en un espacio E numerable y veremos luego que dicha construcción resulta ser una CMTC.. Supongamos entonces que el espacio de estados E es un conjunto numerable. Sea $q : E \times E \rightarrow \mathbb{R}$ una función tal que $q(x, y) \geq 0$ para todos $x, y \in E$, $x \neq y$. Llamaremos a $q(x, y)$ velocidad de transición del estado x al estado y . Queremos construir un proceso con la propiedad de que la velocidad de salto de un estado x a otro y sea $q(x, y)$. En otras palabras el proceso debe verificar

$$P(X_{t+h} = y | X_t = x) = hq(x, y) + o(h) \quad \forall x \neq y \quad (3.3)$$

Sea $q : E \rightarrow \mathbb{R}$ tal que $q(x) := \sum_{y \in E} q(x, y)$ la velocidad de salida del estado x . Supondremos que las tasas de salida están uniformemente acotadas, es decir que

$$\lambda := \sup_{x \in E} \sum_{y \in E} q(x, y) = \sup_{x \in E} q(x) < +\infty \quad (3.4)$$

(Observar que si el espacio de estados E es finito, entonces esta condición se satisface automáticamente). Para construir el proceso, introducimos un proceso de Poisson bidimensional $M(\cdot) : \Omega \times \mathcal{P}(\mathbb{R}^2) \rightarrow \mathbb{R}$ con parámetro 1. Para cada $x \in E$ particionamos el intervalo $I^x := [0, q(x)]$ en intervalos $I(x, y)$ tales que $|I(x, y)| = q(x, y)$. De esta forma, tenemos que $\{I(x, y)\}_{y \in E} \subset \mathbb{R}$ verifica que:

1. $|I(x, y)| = q(x, y)$
2. $\biguplus_{y \in E} I(x, y) = I^x = [0, q(x)]$

Sea $x_0 \in E$ y supongamos que $X_0 = x_0$. Sea τ_1 el primer momento en que un evento del proceso $M(\cdot)$ aparece en el intervalo I^{x_0} . Entonces :

$$\tau_1 := \inf\{t > 0 : M([0, t] \times I^{x_0}) > 0\} \quad (3.5)$$

Definimos $x_1 = y$ donde y es el único estado que satisface que

$$\inf\{t > 0 : M([0, t] \times I^{x_0}) > 0\} = \inf\{t > 0 : M([0, t] \times I(x_0, y)) > 0\} \quad (3.6)$$

(observar que x_1 está bien definido, ya que como los $\{I(x, y)\}_{y \in E}$ son una partición del intervalo $[0, q(x)]$ entonces el estado y que verifica 3.6 es único).

Supongamos ahora que tenemos determinados τ_{n-1} y x_{n-1} . Definimos entonces de manera inductiva:

$$\tau_n := \inf\{t > \tau_{n-1} : M([\tau_{n-1}, t] \times I^{x_{n-1}}) > 0\} \quad (3.7)$$

y $x_n = y$ donde y es el único estado que verifica que

$$\inf\{t > \tau_{n-1} : M([\tau_{n-1}, t] \times I^{x_{n-1}}) > 0\} = \inf\{t > \tau_{n-1} : M([\tau_{n-1}, t] \times I(x_{n-1}, y)) > 0\} \quad (3.8)$$

Definición 3.1.3. *Definimos*

$$\tau_\infty := \sup_n \tau_n \quad (3.9)$$

y el proceso $X_t : \Omega \rightarrow E$ con $t \in [0, \tau_\infty)$ tal que

$$X_t := x_n \text{ siempre que } t \in [\tau_n, \tau_{n+1}) \quad (3.10)$$

De esta forma τ_n es el tiempo en el que ocurre el n -ésimo salto del proceso, y x_n es el n -ésimo estado visitado por el proceso $\{X_t\}_{t \in [0, \tau_\infty)}$.

Ejemplo 3.1.4. *Veamos como hacer la construcción anterior para un proceso con 3 estados. Supongamos que tenemos un sistema con 2 servidores, que no tiene lugar de espera para clientes no servidos, es decir, los clientes que llegan cuando los dos servidores están ocupados se van del sistema automáticamente. Supongamos también que los clientes llegan al sistema de acuerdo con un proceso de Poisson de tasa λ y que los servidores atienden con una distribución exponencial de parámetro β . Es claro que este sistema tiene 3 estados posibles: no hay clientes en el sistema; hay un cliente en el sistema; hay 2 clientes en el sistema. Sea entonces $E = \{0, 1, 2\}$ nuestro espacio de estados. Calculemos ahora cuales son las velocidades de transición entre los estados.*

1. *Si el sistema esta en el estado 0, es decir que no hay clientes, solo puede haber transferencia al estado 1 cuando llega un cliente, lo cual se produce con tasa λ .*
2. *Si en el sistema hay un cliente, entonces el estado siguiente puede ser que hayan 2 clientes (en el caso en que llegue un nuevo cliente al sistema, lo cual se produce con tasa λ), o puede ser que el sistema vuelva al estado 0 (en el caso de que el cliente que esta siendo atendido se retire antes de que llegue el nuevo, lo cual se produce con tasa β).*
3. *Si en el sistema hay 2 clientes entonces solo se puede pasar al estado 1 cuando un cliente termina de ser atendido, lo cual sucede con tasa 2β ya que hay dos servidores atendiendo, cada uno con tasa β .*

Entonces tenemos que

1. $q(0, 1) = \lambda$
2. $q(1, 2) = \lambda$
3. $q(1, 0) = \beta$
4. $q(2, 1) = 2\beta$

y las otras velocidades de transición son 0. Observar que las velocidades de salida de los estados están acotadas por $\max\{\lambda + \beta, 2\beta\}$.

El siguiente paso es particionar los intervalos $[0, q(x)]$ para cada $x \in E$ en intervalos

$I(x, y)$ de longitud $q(x, y)$. Hagamos uno como ejemplo, los otros son análogos. Sabemos que

$$q(1) = \sum_{y=0}^2 q(1, y) = q(1, 0) + q(1, 1) + q(1, 2) = \beta + \lambda$$

Entonces, una posible elección de los intervalos podría ser la siguiente: como $|I(1, 2)| = q(1, 2) = \lambda$ entonces elegimos $I(1, 2) = [0, \lambda]$. Por otro lado sabemos que $|I(1, 0)| = q(1, 0) = \beta$ y $|I(1, 1)| = 0$ entonces definimos $I(1, 0) = [\lambda, \lambda + \beta]$ y $I(1, 1) = \emptyset$. Siguiendo con este razonamiento los intervalos elegidos pueden ser:

$$I(0, 1) = I(1, 2) = [0, \lambda]$$

$$I(1, 0) = [\lambda, \lambda + \beta]$$

$$I(2, 1) = [0, 2\beta]$$

$$I(0, 0) + I(0, 2) = I(1, 1) = I(2, 0) + I(2, 2) = \emptyset$$

Luego, de acuerdo con un proceso de Poisson $M(\cdot)$ bidimensional y los intervalos anteriores se construye el proceso $\{X_t\}_{t \in [0, \tau_\infty)}$

Veamos a continuación que el proceso definido en 3.10 verifica la condición 3.3.

Proposición 3.1.5. *El proceso $\{X_t\}_{t \in [0, \tau_\infty)}$ definido por 3.10 verifica la propiedad*

$$P(X_{t+h} = y | X_t = x) = hq(x, y) + o(h) \quad \forall x \neq y$$

Demostración. Por la definición del proceso, es claro que:

$$\{X_{t+h} = y | X_t = x\} \subset \{M((t, t+h] \times I(x, y)) = 1\} \cup \{M((t, t+h] \times [0, q(x)]) \geq 2\}$$

Luego, se cumple que:

$$P(X_{t+h} = y | X_t = x) = P(M((t, t+h] \times I(x, y)) = 1) + P(A) \quad (3.11)$$

donde $A \subset \{M((t, t+h] \times [0, q(x)]) \geq 2\}$.

Por otro lado, como $q(x)$ está acotado, aplicando la proposición 2.8.1 tenemos que:

$$\begin{aligned} P(M((t, t+h] \times I(x, y)) = 1) &= h|I(x, y)|e^{-h|I(x, y)|} \\ &= hq(x, y)e^{-hq(x, y)} \\ &= hq(x, y) + o(h) \end{aligned}$$

y

$$\begin{aligned} P(M((t, t+h] \times [0, q(x)]) \geq 2) &= 1 - P(M((t, t+h] \times [0, q(x)]) = 0) \\ &\quad - P(M((t, t+h] \times [0, q(x)]) = 1) \\ &= 1 - e^{-hq(x)} - hq(x)e^{-hq(x)} \\ &= o(h) \end{aligned}$$

donde la última igualdad en cada ecuación se deduce usando el desarrollo en serie de la función exponencial. Utilizando el resultado anterior, para terminar solo hay que observar que como $A \subset \{M((t, t+h] \times [0, q(x)]) \geq 2\}$ entonces

$$P(A) \leq P(M((t, t+h] \times [0, q(x)]) \geq 2)$$

Luego

$$\lim_{h \rightarrow 0} \frac{P(A)}{h} \leq \lim_{h \rightarrow 0} \frac{P(M((t, t+h] \times [0, q(x)]) \geq 2)}{h} = 0$$

por lo cual $P(A)$ es un $o(h)$. Volviendo a 3.11 concluimos entonces que

$$P(X_{t+h} = y | X_t = x) = hq(x, y) + o(h)$$

para todos $x, y \in E$ diferentes, como queríamos probar. \square

3.2. Propiedades

En esta sección veremos que el proceso $\{X_t\}$ construido por 3.10 es una cadena de Markov de tiempo continuo, es decir que, dado el presente, el futuro y el pasado son independientes.

Proposición 3.2.1. *El proceso $\{X_t\}_{t \in [0, \tau_\infty)}$ definido en 3.10 satisface la siguiente condición:*

$$P(X_{t+s} = y | X_t = x, X_u = x_u \forall 0 \leq u \leq t) = P(X_{t+s} = y | X_t = x) \quad (3.12)$$

donde $x, y, x_u \in E, \forall 0 \leq u \leq t$.

Demostración. Por la construcción hecha, el proceso después del tiempo t_0 solo depende del proceso bidimensional $M(\cdot)$ en la región $(t_0, +\infty) \times \mathbb{R}$ y del valor que toma X_{t_0} en el tiempo t_0 . Luego, dado el valor que asume la cadena en el tiempo t_0 , el futuro es independiente del pasado. \square

En los próximos dos resultados veremos que:

1. dado que el proceso X_t se encuentra en el estado $x \in E$ en el tiempo τ_n , el tiempo transcurrido hasta el siguiente salto es una v.a. con distribución exponencial con media $\frac{1}{q(x)}$.
2. cuando el proceso decide saltar de un estado $x \in E$ a un estado diferente $y \in E$ lo hace con probabilidad $p(x, y) := \frac{q(x, y)}{q(x)}$.

Teorema 3.2.2. *Sea un proceso de Markov de tiempo continuo tal que las velocidades de salida de los estados $q(x)$ están uniformemente acotadas por $\lambda > 0$. Entonces vale que:*

$$P(\tau_{n+1} - \tau_n > t | X_{\tau_n} = x) = e^{-tq(x)} \quad (3.13)$$

Demostración. Sea $M(\cdot) : \Omega \times \mathcal{P}(\mathbb{R}^2) \rightarrow \mathbb{N}$ el proceso de Poisson bidimensional de parámetro uno. Sean S_n con $n \in \mathbb{N}$ los tiempos ordenados de ocurrencia de los eventos en la banda $[0, \lambda]$, y sean W_n con $n \in \mathbb{N}$ las segundas coordenadas de esos eventos (como en el teorema 2.10.2). Por el teorema 2.10.2 sabemos que

1. $\{S_{i+1} - S_i\}_{i \geq 1}$ son v.a.i. distribuídas exponencialmente con parámetro λ
2. $\{W_i\}_{i \geq 1}$ son v.a.i. distribuídas uniformemente en el intervalo $[0, \lambda]$
3. $\{S_{i+1} - S_i\}_{i \geq 1}$ y $\{W_i\}_{i \geq 1}$ son independientes.

Es claro que el conjunto $\{\tau_n\}_{n \geq 0} \subset \{S_n\}_{n \geq 0}$. De hecho, dado $x_0 \in E$ podemos definir:

$$\tau_n := \min\{S_k > \tau_{n-1} : W_k < q(x_{n-1})\} \quad (3.14)$$

$$K_n := \{k : S_k = \tau_n\} \quad (3.15)$$

$$x_n := \{y \in E : W_{K_n} \in I(x_{n-1}, y)\} \quad (3.16)$$

La distribución de $\tau_{n+1} - \tau_n$ condicionada a que $X_{\tau_n} = x$ es

$$P(\tau_{n+1} - \tau_n > t | X_{\tau_n} = x) = \frac{P(\tau_{n+1} - \tau_n > t, X_{\tau_n} = x)}{P(X_{\tau_n} = x)} \quad (3.17)$$

Condicionando a los posibles valores que K_n puede tomar, tenemos que

$$P(\tau_{n+1} - \tau_n > t, X_{\tau_n} = x) = \sum_{k \in \mathbb{N}} P(\tau_{n+1} - \tau_n > t, X_{\tau_n} = x, K_n = k)$$

lo cual, utilizando la definición de K_n es lo mismo que

$$\sum_{k \in \mathbb{N}} P(\tau_{n+1} - S_k > t, X_{S_k} = x, K_n = k)$$

Utilizando ahora la definición 3.14 y los resultados (1), (2) y (3) acerca de las variables $\{S_{i+1} - S_i\}_{i \geq 1}$ y $\{W_i\}_{i \geq 1}$ enunciados al comienzo de la demostración, tenemos que lo anterior es:

$$\begin{aligned} & \sum_{k \in \mathbb{N}} \left(\sum_{l \in \mathbb{N}} P(S_{k+l} - S_k > t, W_{k+1} > q(x), \dots, W_{k+l-1} > q(x), W_{k+l} < q(x)) \right) P(X_{S_k} = x, K_n = k) = \\ & = e^{-tq(x)} \sum_{k \in \mathbb{N}} P(X_{S_k} = x, K_n = k) \end{aligned}$$

Entonces tenemos que

$$\begin{aligned}
 P(\tau_{n+1} - \tau_n > t | X_{\tau_n} = x) &= \frac{e^{-tq(x)}}{P(X_{\tau_n} = x)} \sum_{k \in \mathbb{N}} P(X_{S_k} = x, K_n = k) \\
 &= \frac{e^{-tq(x)}}{P(X_{\tau_n} = x)} \sum_{k \in \mathbb{N}} P(X_{\tau_n} = x, K_n = k) \\
 &= \frac{e^{-tq(x)}}{P(X_{\tau_n} = x)} P(X_{\tau_n} = x) \\
 &= e^{-tq(x)}
 \end{aligned}$$

como queríamos probar. □

Definición 3.2.3. Definimos el proceso de tiempo discreto $Y_n : \Omega \rightarrow \mathbb{N}$ con $n \in \mathbb{N}$ de manera que $Y_n = X_{\tau_n} \forall n \in \mathbb{N}$. A este proceso lo llamaremos el esqueleto del proceso de tiempo continuo $\{X_t\}_{t \geq 0}$.

Teorema 3.2.4. El esqueleto $\{Y_n\}_{n \in \mathbb{N}}$ de un proceso de tiempo continuo $\{X_t\}_{t \geq 0}$ es una cadena de Markov con probabilidades de transición $\{p(x, y) : x, y \in E\}$ dadas por

$$p(x, y) := \frac{q(x, y)}{q(x)}$$

Demostración. Queremos probar que $P(Y_{n+1} = y | Y_n = x) = p(x, y)$ lo cual es equivalente a probar que $P(X_{\tau_{n+1}} = y | X_{\tau_n} = x) = p(x, y)$. Usaremos nuevamente las construcciones 3.14, 3.15 y 3.16. Partiendo el espacio según los valores que puede tomar K_n tenemos que:

$$P(X_{\tau_{n+1}} = y, X_{\tau_n} = x) = \sum_{k \in \mathbb{N}} P(X_{\tau_{n+1}} = y, X_{\tau_n} = x, K_n = k) \quad (3.18)$$

Observemos ahora que como $K_n = k$ entonces se cumple que el conjunto $\{X_{\tau_{n+1}} = y, X_{\tau_n} = x, K_n = k\}$ es

$$\bigcup_{l \geq 1} \{W_{k+1} > q(x), \dots, W_{k+l-1} > q(x), W_{k+l} \in I(x, y), X_{S_k} = x, K_n = k\}$$

Utilizando lo anterior junto con la independencia entre $\{W_i\}_{i \in \mathbb{N}}$ y $\{S_i\}_{i \in \mathbb{N}}$ tenemos que 3.18 es

$$\sum_{k \in \mathbb{N}} \sum_{l \geq 1} P(W_{k+1} > q(x), \dots, W_{k+l-1} > q(x), W_{k+l} \in I(x, y)) P(X_{S_k} = x, K_n = k)$$

Como $\{W_i\}_{i \in \mathbb{N}}$ son independientes tenemos que

$$\sum_{l \geq 1} P(W_{k+1} > q(x), \dots, W_{k+l-1} > q(x), W_{k+l} \in I(x, y)) =$$

$$\begin{aligned} \sum_{l \geq 1} P(W_{k+1} > q(x)) \dots P(W_{k+l-1} > q(x)) P(W_{k+l} \in I(x, y)) &= \\ \sum_{l \geq 1} \left(\frac{\lambda - q(x)}{\lambda} \right)^{l-1} \frac{|I(x, y)|}{\lambda} &= \frac{q(x, y)}{q(x)} \end{aligned}$$

donde las dos últimas igualdades se deben a que $\{W_i\}_{i \geq 1}$ son v.a.i. distribuídas uniformemente en el intervalo $[0, \lambda]$, $q(x) < \lambda$ y $|I(x, y)| = q(x, y)$. Por otro lado tenemos que

$$\begin{aligned} P(Y_n = x) &= P(X_{\tau_n} = x) \\ &= \sum_{k \geq 0} P(X_{\tau_n} = x, K_n = k) \\ &= \sum_{k \geq 0} P(X_{S_k} = x, K_n = k) \end{aligned}$$

Entonces tenemos que:

$$\begin{aligned} P(Y_{n+1} = y | Y_n = x) &= \frac{P(X_{\tau_{n+1}} = y, X_{\tau_n} = x)}{P(X_{\tau_n} = x)} \\ &= \frac{q(x, y)}{q(x)} \sum_{k \geq 0} P(X_{S_k} = x, K_n = k) \frac{1}{P(Y_n = x)} \\ &= \frac{q(x, y)}{q(x)} \\ &= p(x, y) \end{aligned}$$

□

3.3. Explosiones

La construcción hecha en 3.10 define el proceso $\{X_t\}$ hasta el tiempo $\tau_\infty := \sup_{n \in \mathbb{N}} \tau_n$. Si $\tau_\infty = \infty$ entonces el proceso está definido para todo tiempo $t \in [0, \infty)$. Sin embargo, si $\tau_\infty < \infty$ no tenemos en principio definido el proceso para tiempos mayores que τ_∞ . En esta sección veremos como tratar estos casos y daremos una condición necesaria y suficiente para que $\tau_\infty = \infty$.

Definición 3.3.1. Sea $\{X_t\}_{t \in [0, \tau_\infty)}$ definido por 3.10. Decimos que el proceso explota si

$$P \left(\lim_{n \rightarrow +\infty} \tau_n < \infty \right) > 0$$

es decir que el proceso explota cuando un número infinito de transiciones ocurren en tiempo finito. De esta manera, después de un tiempo finito τ_∞ el proceso no está formalmente definido. Definimos el proceso explosivo agregando un nuevo estado que llamaremos ∞ con velocidades de transición $q(\infty, x) = 0$ para todo $x \in E$.

La siguiente proposición nos da información acerca de la distribución exponencial que nos será útil mas adelante.

Proposición 3.3.2. Sean X e Y v.a.i. con distribución exponencial con parámetros λ y β respectivamente donde $\lambda > 0$ y $\beta > 0$. Entonces:

1. Propiedad de pérdida de memoria: Para $s, t > 0$:

$$P(X > t + s | X > s) = P(X > t) = e^{-\lambda t}$$

2. Notemos el mínimo entre dos números x e y mediante $x \wedge y$. Entonces la variable aleatoria $X \wedge Y$ tiene distribución exponencial con parámetro $\lambda + \beta$, es decir

$$P(X \wedge Y > x) = e^{-(\lambda+\beta)x}, \quad x \geq 0$$

3. Se cumple que

$$P(X > Y) = \frac{\beta}{\lambda + \beta}$$

4. Supongamos que $\{X_n\}_{n \in \mathbb{N}}$ son v.a.i. exponencialmente distribuídas tal que $X_n \sim \text{Exp}(\lambda_n)$, es decir que

$$P(X_n > x) = e^{-\lambda_n x}, \quad x \geq 0, \quad \lambda_n > 0, \quad n \in \mathbb{N}$$

Entonces se cumple que

$$\sum_{n \in \mathbb{N}} X_n < \infty \text{ c.s} \iff \sum_{n \in \mathbb{N}} \frac{1}{\lambda_n} < \infty$$

Demostración. 1. Calculando tenemos que:

$$\begin{aligned} P(X > t + s | X > s) &= \frac{P(X > t + s, X > s)}{P(X > s)} \\ &= \frac{P(X > t + s)}{P(X > s)} \\ &= \frac{e^{-\lambda(t+s)}}{e^{-\lambda s}} \\ &= e^{-\lambda t} \\ &= P(X > t) \end{aligned}$$

2. Debido a que $\{X \wedge Y > x\} = \{X > x, Y > x\}$ tenemos que

$$\begin{aligned} P(X \wedge Y > x) &= P(X > x, Y > x) \\ &= e^{-\lambda x} e^{-\beta x} \\ &= e^{-(\lambda+\beta)x} \end{aligned}$$

3. Integrando tenemos que

$$\begin{aligned}
 P(X > Y) &= \int \int_{\{(u,v):u>v>0\}} \lambda e^{-\lambda u} \beta e^{-\beta v} du dv \\
 &= \int_0^{+\infty} \beta e^{-\beta v} \left(\int_v^{+\infty} \lambda e^{-\lambda u} du \right) dv \\
 &= \int_0^{+\infty} \beta e^{-\beta v} e^{-\lambda v} dv \\
 &= \frac{\beta}{\lambda + \beta}
 \end{aligned}$$

4. Observar que si $\sum_{n \in \mathbb{N}} X_n < \infty$ entonces para todo $s > 0$ se cumple que

$$0 < \mathbf{E}(e^{-s \sum_{n \in \mathbb{N}} X_n}) = \prod_{n \in \mathbb{N}} \mathbf{E}(e^{-s X_n}) = \prod_{n \in \mathbb{N}} \frac{\lambda_n}{\lambda_n + s}$$

Luego, aplicando el lema 2.4.1 tenemos que:

$$\sum_{n \in \mathbb{N}} \left(1 - \frac{\lambda_n}{\lambda_n + s} \right) = \sum_{n \in \mathbb{N}} \frac{s}{\lambda_n + s} < \infty$$

Esto implica que $\lim_{n \rightarrow \infty} \frac{s}{\lambda_n + s} = 0$ y entonces $\lim_{n \rightarrow \infty} \lambda_n = \infty$. Luego si $n \rightarrow \infty$

$$\frac{s}{\lambda_n + s} \sim \frac{s}{\lambda_n}$$

por lo tanto

$$\sum_{n \in \mathbb{N}} \frac{s}{\lambda_n + s} < \infty \iff \sum_{n \in \mathbb{N}} \frac{1}{\lambda_n} < \infty$$

como queríamos probar.

Recíprocamente, si $\sum_{n \in \mathbb{N}} \frac{1}{\lambda_n} < \infty$ entonces $\sum_{n \in \mathbb{N}} \mathbf{E}(X_n) = \mathbf{E}(\sum_{n \in \mathbb{N}} X_n) < \infty$. Luego como $\sum_{n \in \mathbb{N}} X_n \geq 0$ tenemos que $\sum_{n \in \mathbb{N}} X_n < \infty$ □

El siguiente resultado nos permite enunciar una condición necesaria y suficiente para que un proceso no tenga explosiones.

Teorema 3.3.3. *Dado el proceso $\{X_t\}_{t \in [0, \tau_\infty)}$ se cumple que*

$$P(\tau_\infty < \infty) = P\left(\sum_{n \in \mathbb{N}} \frac{1}{q(X_n)} < \infty\right) \tag{3.19}$$

y entonces el proceso no tiene explosiones si y solo si

$$P\left(\sum_{n \in \mathbb{N}} \frac{1}{q(X_n)} < \infty\right) = 0$$

En particular se cumple que

1. Si existe $c > 0$ tal que $q(x) \leq c \forall x \in E$ entonces la cadena no tiene explosiones.
2. Si el espacio de estados E es finito entonces la cadena no tiene explosiones.

Demostración. Condicionado a la secuencia de estados $\{X_n\}$ tenemos que τ_∞ es suma de variables aleatorias independientes exponenciales de parámetros $q(X_n)$ es decir

$$\tau_\infty = \sum_{n \in \mathbb{N}} (\tau_{n+1} - \tau_n) = \sum_{n \in \mathbb{N}} S_n$$

donde las S_n son v.a.i. con distribución exponencial de parámetro $q(X_n)$. Luego por la proposición 3.3.2 tenemos que, condicionado a $\{X_n\}$

$$P(\tau_\infty < \infty) = \begin{cases} 1 & \text{si } \sum_{n \in \mathbb{N}} \frac{1}{q(X_n)} < \infty \\ 0 & \text{si } \sum_{n \in \mathbb{N}} \frac{1}{q(X_n)} = \infty \end{cases}$$

Tomando esperanzas concluimos que

$$P(\tau_\infty < \infty) = P\left(\sum_{n \in \mathbb{N}} \frac{1}{q(X_n)} < \infty\right)$$

Para (1) observemos que si $q(x) \leq c$ para todo $x \in E$ entonces

$$\sum_{n \in \mathbb{N}} \frac{1}{q(X_n)} \geq \sum_{n \in \mathbb{N}} \frac{1}{c} = \infty$$

entonces el proceso no tiene explosiones.

Por otro lado si el espacio E es finito entonces $q(x) \leq \max_{y \in E} q(y) = \hat{c} < \infty$ entonces aplicando (1) obtenemos (2). \square

3.4. Ecuaciones de Kolmogorov

De ahora en adelante utilizaremos las siguientes notaciones:

1. Q denotará la matriz real cuyas entradas son

$$\begin{cases} q(x, y) = \frac{p(x, y)}{q(x)} & \text{si } x \neq y \\ q(x, x) := -q(x) = -\sum_{y \neq x} q(x, y) & \text{en otro caso} \end{cases}$$

2. P_t denotará la matriz con entradas

$$p_t(x, y) := P(X_t = y | X_0 = x)$$

Teorema 3.4.1 (Ecuaciones de Chapman-Kolmogorov). *Dados $t, s \geq 0$ se cumple que*

$$P_{t+s} = P_s P_t \tag{3.20}$$

Demostración. Utilizando las definiciones correspondientes y las propiedades de una cadena de Markov de tiempo continuo (homogénea) tenemos que:

$$\begin{aligned}
 p_{t+s}(x, y) &= P(X_{t+s} = y | X_0 = x) \\
 &= \sum_{z \in E} P(X_{t+s} = y, X_s = z | X_0 = x) \\
 &= \sum_{z \in E} \frac{P(X_{t+s} = y, X_s = z, X_0 = x)}{P(X_0 = x)} \\
 &= \sum_{z \in E} P(X_{t+s} = y | X_s = z, X_0 = x) \frac{P(X_s = z, X_0 = x)}{P(X_0 = x)} \\
 &= \sum_{z \in E} P(X_{t+s} = y | X_s = z) P(X_s = z | X_0 = x) \\
 &= \sum_{z \in E} p_s(x, z) p_t(z, y)
 \end{aligned}$$

Esto muestra que el elemento de la matriz P_{t+s} que esta en la fila x y columna y se obtiene haciendo el producto de la fila x de la matriz P_s con la columna y de la matriz P_t , con lo cual queda probado que

$$P_{t+s} = P_s P_t$$

□

Observación 3.4.2. Observemos que fijados x e y en E , podemos considerar la función

$$p_{(\cdot)}(x, y) : \mathbb{R}^+ \cup \{0\} \rightarrow \mathbb{R}$$

tal que

$$t \mapsto p_t(x, y) = P(X_t = y | X_0 = x)$$

Notemos entonces por:

$$p'_t(x, y) = \frac{d}{dt}(p_t(x, y))$$

a su derivada (respecto de t). En el siguiente teorema daremos dos ecuaciones que vinculan P'_t con P_t donde P'_t indica la matriz de entradas $p'_t(x, y)$. Antes del teorema veamos un lema que nos será de utilidad.

Lema 3.4.3. Se cumplen:

1. $\lim_{t \rightarrow 0} \frac{1 - p_t(x, x)}{t} = q(x)$
2. $\lim_{t \rightarrow 0} \frac{p_t(x, y)}{t} = q(x, y)$ si $x \neq y$

Demostración. 1.

$$p_t(x, x) = P(X_t = x | X_0 = x) \quad (3.21)$$

Como $\sum_{y \in E} P(X_t = y | X_0 = x) = 1$ entonces $P(X_t = x | X_0 = x) = 1 - \sum_{y \neq x} P(X_t = y | X_0 = x)$.

Luego utilizando 3.3 tenemos que

$$\begin{aligned} p_t(x, x) &= 1 - \sum_{y \neq x} (tq(x, y) + o(t)) \\ &= 1 - t \sum_{y \neq x} q(x, y) - o(t) \\ &= 1 - tq(x) - o(t) \end{aligned}$$

Entonces

$$\lim_{t \rightarrow 0} \left(\frac{1 - p_t(x, x)}{t} \right) = \lim_{t \rightarrow 0} \left(q(x) - \frac{o(t)}{t} \right) = q(x)$$

2. Utilizando nuevamente 3.3 tenemos que si $x \neq y$ se cumple:

$$\begin{aligned} \lim_{t \rightarrow 0} \frac{p_t(x, y)}{t} &= \lim_{t \rightarrow 0} \frac{P(X_t = y | X_0 = x)}{t} \\ &= \lim_{t \rightarrow 0} \frac{tq(x, y) + o(t)}{t} \\ &= \lim_{t \rightarrow 0} \left(q(x, y) + \frac{o(t)}{t} \right) \\ &= q(x, y) \end{aligned}$$

□

Teorema 3.4.4 (Ecuaciones de Kolmogorov). Para todo $t \geq 0$ se cumplen

$$P'_t = Q P_t \quad \text{Backward equations} \quad (3.22)$$

$$P'_t = P_t Q \quad \text{Forward equations} \quad (3.23)$$

Demostración. Probaremos 3.22 ya que 3.23 se demuestra razonando de manera análoga. Utilizando las ecuaciones de Kolmogorov-Chapman:

$$p_{t+h}(x, y) = \sum_{z \in E} p_h(x, z) p_t(z, y)$$

Entonces

$$\begin{aligned} p_{t+h}(x, y) - p_t(x, y) &= \sum_{z \neq x} p_h(x, z) p_t(z, y) + p_h(x, x) p_t(x, y) - p_t(x, y) \\ &= \sum_{z \neq x} p_h(x, z) p_t(z, y) - p_t(x, y) (1 - p_h(x, x)) \end{aligned}$$

Entonces, si vemos que en las condiciones que estamos podemos intercambiar límite con sumatoria, tenemos que:

$$\begin{aligned}
 \lim_{h \rightarrow 0} \frac{p_{t+h}(x, y) - p_t(x, y)}{h} &= \lim_{h \rightarrow 0} \left(\sum_{z \neq x} \frac{p_h(x, z)}{h} p_t(z, y) - p_t(x, y) \frac{(1 - p_h(x, x))}{h} \right) \\
 &= \sum_{z \neq x} \lim_{h \rightarrow 0} \left(\frac{p_h(x, z)}{h} p_t(z, y) - p_t(x, y) \frac{(1 - p_h(x, x))}{h} \right) \\
 &= \sum_{z \neq x} q(x, z) p_t(z, y) - p_t(x, y) q(x) \\
 &= \sum_{z \in E} q(x, z) p_t(z, y)
 \end{aligned}$$

donde la penúltima igualdad se deduce del lema 3.4.3. De esta manera, para completar la prueba solo queda ver que se pueden intercambiar el límite y la sumatoria.

Fijado N tenemos que

$$\begin{aligned}
 \liminf_{h \rightarrow 0} \sum_{z \neq x} \frac{p_h(x, z)}{h} p_t(z, y) &\geq \liminf_{h \rightarrow 0} \sum_{z \neq x, z < N} \frac{p_h(x, z)}{h} p_t(z, y) \\
 &= \sum_{z \neq x, z < N} q(x, z) p_t(z, y)
 \end{aligned}$$

Como N es arbitrario entonces

$$\liminf_{h \rightarrow 0} \sum_{z \neq x} \frac{p_h(x, z)}{h} p_t(z, y) \geq \sum_{z \neq x} q(x, z) p_t(z, y) \quad (3.24)$$

Recíprocamente si $N > x$, como $p_t(z, y) \leq 1$:

$$\limsup_{h \rightarrow 0} \sum_{z \neq x} \frac{p_h(x, z)}{h} p_t(z, y) \leq \limsup_{h \rightarrow 0} \left(\sum_{z \neq x, z < N} \frac{p_h(x, z)}{h} p_t(z, y) + \sum_{z \geq N} \frac{p_h(x, z)}{h} \right) \quad (3.25)$$

Observar que

$$\begin{aligned}
 \sum_{z \geq N} p_h(x, z) &= 1 - \sum_{z < N} p_h(x, z) \\
 &= 1 - \left(\sum_{z < N, z \neq x} p_h(x, z) \right) - p_h(x, x)
 \end{aligned}$$

Entonces 3.25 resulta:

$$\limsup_{h \rightarrow 0} \left(\sum_{z \neq x, z < N} \frac{p_h(x, z)}{h} p_t(z, y) + \frac{1 - p_h(x, x)}{h} - \sum_{z \neq x, z < N} \frac{p_h(x, z)}{h} \right) =$$

$$\sum_{z \neq x, z < N} q(x, z)p_t(z, y) + q(x) - \sum_{z \neq x, z < N} q(x, z)$$

Esto es verdad para todo $N > x$ entonces haciendo $N \rightarrow \infty$ tenemos que

$$\lim_{N \rightarrow \infty} \sum_{z \neq x, z < N} q(x, z) = \sum_{z \neq x} q(x, z) = q(x)$$

Entonces:

$$\limsup_{h \rightarrow 0} \sum_{z \neq x} \frac{p_h(x, z)}{h} p_t(z, y) \leq \sum_{z \neq x} q(x, z) p_t(z, y) \quad (3.26)$$

De esta manera, 3.24 y 3.26 muestran que

$$\lim_{h \rightarrow 0} \frac{p_h(x, z)}{h} p_t(z, y) = \sum_{z \neq x} q(x, z) p_t(z, y)$$

completandose así la prueba del teorema. □

3.5. Clasificación de estados y medidas invariantes

En esta sección daremos algunas maneras de clasificar los estados de una cadena de Markov de tiempo continuo y definiremos cadena irreducible y medida invariante para una CMTC.

Definición 3.5.1. *Sea*

$$T^{x \rightarrow y} := \inf\{t > \tau_1 : X_t^x = y\} \quad (3.27)$$

el primer momento en que el proceso que comienza en el estado x alcanza al estado y .

Observar que en la definición pedimos $t > \tau_1$ para evitar que $T^{x \rightarrow x} = 0$

Definición 3.5.2. *Decimos que el estado $x \in E$ es*

1. *transitorio: si $P(T^{x \rightarrow x} = \infty) > 0$*
2. *recurrente nulo: si $P(T^{x \rightarrow x} = \infty) = 0$ y $\mathbf{E}(T^{x \rightarrow x}) = \infty$*
3. *recurrente positivo: si $\mathbf{E}(T^{x \rightarrow x}) < \infty$*
4. *recurrente: si es recurrente nulo o recurrente positivo.*

Dicho en palabras, un estado es transitorio si la probabilidad de que la cadena no vuelva a visitarlo en el futuro es positiva, es recurrente si el estado es visitado infinitas veces con probabilidad 1, y es recurrente positivo si el tiempo esperado de retorno al estado es finito.

Definición 3.5.3. Decimos que el proceso $\{X_t\}_{t \geq 0}$ es irreducible si para todo $x, y \in E$ la probabilidad de alcanzar el estado y saliendo del estado x en un tiempo finito es positiva, es decir si

$$P(T^{x \rightarrow y} < \infty) > 0 \quad (3.28)$$

A continuación daremos las definiciones de medida invariante para cadenas de tiempo discreto y de tiempo continuo.

Definición 3.5.4. Sea $\{X_n\}_{n \in \mathbb{N}}$ una cadena de Markov de tiempo discreto con espacio de estados E numerable y matriz de transición Q es decir $Q(x, y) = P(X_1 = y | X_0 = x)$ ($Q : E \times E \rightarrow [0, 1]$ tal que $\sum_{y \in E} Q(x, y) = 1$). Una medida de probabilidad μ definida en E se llama invariante respecto de Q si verifica que:

$$\mu(x) = \sum_{z \in E} \mu(z)Q(z, x) \quad (3.29)$$

Definición 3.5.5. Decimos que una medida de probabilidad en E que llamaremos π es invariante o estacionaria para el proceso continuo $\{X_t\}_{t \geq 0}$ si verifica que:

$$\pi Q = 0 \quad (3.30)$$

es decir

$$\sum_{x \in E} \pi(x)q(x, y) = 0 \quad \forall y \in E \quad (3.31)$$

y

$$\sum_{x \in E} \pi(x) = 1 \quad (3.32)$$

donde Q es la matriz definida al comienzo de la sección 3.4.

3.6. Esqueletos

Asumamos en esta sección que el proceso $\{X_t\}_{t \geq 0}$ y su esqueleto $\{Y_n\}_{n \in \mathbb{N}}$ definido por 3.2.3 son irreducibles y recurrentes positivos. Sabemos que una medida ν es invariante para la cadena discreta $\{Y_n\}_{n \in \mathbb{N}}$ si satisface

$$\sum_{x \in E} \nu(x)p(x, y) = \nu(y)$$

Por otro lado, una medida π es invariante para el proceso continuo $\{X_t\}_{t \geq 0}$ si satisface

$$\sum_{x \in E} \pi(x)q(x, y) = \pi(y)q(y)$$

El siguiente resultado nos muestra la relación que existe entre una medida invariante para el proceso discreto y una medida invariante para el proceso continuo.

Proposición 3.6.1. *Sea ν una medida de probabilidad en E , y consideramos la medida π definida por*

$$\pi(x) := \frac{\nu(x)}{q(x)} \left(\sum_{z \in E} \frac{\nu(z)}{q(z)} \right)^{-1} \quad (3.33)$$

Entonces se cumple que ν es invariante para $\{Y_n\}_{n \in \mathbb{N}}$ si y solo si la medida π es invariante para el proceso $\{X_t\}_{t \geq 0}$.

Demostración. Supongamos que ν es invariante para $\{Y_n\}_{n \in \mathbb{N}}$, entonces

$$\begin{aligned} \nu(y) &= \sum_{x \in E} \nu(x) p(x, y) \\ &= \sum_{x \in E} \nu(x) \frac{q(x, y)}{q(x)} \end{aligned}$$

Luego

$$\sum_{x \in E} \frac{\nu(x)}{q(x)} q(x, y) \left(\sum_{z \in E} \frac{\nu(z)}{q(z)} \right)^{-1} = \nu(y) \left(\sum_{z \in E} \frac{\nu(z)}{q(z)} \right)^{-1}$$

lo cual, aplicando la definición de la medida π es equivalente a

$$\sum_{x \in E} \pi(x) q(x, y) = \pi(y) q(y)$$

y por lo tanto π es invariante para $\{X_t\}_{t \geq 0}$.

Recíprocamente, sabemos que

$$\sum_{x \in E} \pi(x) q(x, y) = \pi(y) q(y)$$

luego usando la definición de la medida π tenemos que

$$\sum_{x \in E} \frac{\nu(x)}{q(x)} \left(\sum_{z \in E} \frac{\nu(z)}{q(z)} \right)^{-1} q(x, y) = \frac{\nu(y)}{q(y)} \left(\sum_{z \in E} \frac{\nu(z)}{q(z)} \right)^{-1} q(y)$$

Simplificando obtenemos que

$$\sum_{x \in E} \nu(x) p(x, y) = \nu(y)$$

y por lo tanto ν es invariante para el proceso $\{Y_n\}_{n \in \mathbb{N}}$. □

Corolario 3.6.2. *Si $q(x) = q(y)$ para todo $x, y \in E$ entonces $\pi(x) = \nu(x) \forall x \in E$.*

Demostración. Sea $q(x) = q \in \mathbb{R}$ para todo $x \in E$, entonces

$$\begin{aligned}\pi(x) &= \frac{\nu(x)}{q} \left(\sum_{z \in E} \frac{\nu(z)}{q} \right)^{-1} \\ &= \nu(x) \left(\sum_{z \in E} \nu(z) \right)^{-1} \\ &= \nu(x)\end{aligned}$$

Entonces

$$\nu(x) = \pi(x) \quad \forall x \in E$$

□

3.7. Procesos de nacimiento y muerte

Los procesos de nacimiento y muerte representan el crecimiento o la extinción de la población de cierto sistema. El valor de la variable X_t representa el número de individuos que hay en el sistema en el tiempo t . Las velocidades de nacimiento y muerte dependen solamente de la cantidad de individuos que hay en ese momento, es decir:

$$q(x, x+1) = \lambda_x^+ \quad y \quad q(x, x-1) = \lambda_x^-$$

donde $\{\lambda_x^+\}_{x \in E}$ y $\{\lambda_x^-\}_{x \in E}$ son familias de parámetros no negativos. A continuación usaremos las ecuaciones de balance para encontrar condiciones bajo las cuales el proceso admite una medida invariante.

Como el espacio de estados E es numerable supondremos que es \mathbb{N} para facilitar la notación. Queremos encontrar un vector $\pi = \{\pi(x)\}_{x \in \mathbb{N}}$ que satisfaga las siguientes condiciones:

$$\pi(0) \sum_{z \in \mathbb{N}} q(0, z) = \sum_{y \in \mathbb{N}} \pi(y) q(y, 0) \tag{3.34}$$

$$\pi(x) \sum_{z \in \mathbb{N}} q(x, z) = \sum_{y \in \mathbb{N}} \pi(y) q(y, x) \tag{3.35}$$

Sustituyendo en cada caso las velocidades de transición correspondientes 3.34 y 3.35 implican que

$$\pi(0)q(0, 1) = \pi(1)q(1, 0) \tag{3.36}$$

$$\pi(x) [q(x, x+1) + q(x, x-1)] = \pi(x-1)q(x-1, x) + \pi(x+1)q(x+1, x) \tag{3.37}$$

donde la última ecuación es para todo $x \geq 1$.

Operando con las ecuaciones anteriores no es difícil concluir que

$$\pi(x+1) = \pi(x) \frac{\lambda_x^+}{\lambda_{x+1}^-} \quad \forall x \geq 0$$

entonces

$$\pi(x) = \prod_{i=0}^{x-1} \frac{\lambda_i^+}{\lambda_{i+1}^-} \pi(0) \quad \forall x \geq 1 \quad (3.38)$$

De esta manera, por construcción se cumple que $\pi = \{\pi(x)\}_{x \in \mathbb{N}}$ verifica las ecuaciones de balance. Falta pedirle que satisfaga $\sum_{x \in \mathbb{N}} \pi(x) = 1$. Para esto es necesario que $\pi(0) > 0$ en virtud de la ecuación 3.38, luego si

$$\sum_{x \geq 1} \frac{\lambda_0^+ \lambda_1^+ \dots \lambda_{x-1}^+}{\lambda_1^- \lambda_2^- \dots \lambda_x^-} < \infty$$

tenemos solución, definiendo

$$\pi(0) := \left(\sum_{x \geq 1} \frac{\lambda_0^+ \lambda_1^+ \dots \lambda_{x-1}^+}{\lambda_1^- \lambda_2^- \dots \lambda_x^-} \right)^{-1}$$

y $\pi(x)$ inductivamente por 3.38.

Capítulo 4

Aplicaciones en Biología

Todos los sistemas biológicos comparten la propiedad del crecimiento de sus individuos, por lo tanto a la hora de estudiar poblaciones (por ejemplo de plantas o animales) la existencia de una teoría acerca de procesos de crecimiento es extremadamente deseada y necesaria. En este capítulo veremos algunas de las aplicaciones al crecimiento de poblaciones que los procesos estocásticos estudiados en los capítulos anteriores de este trabajo poseen.

4.1. Modelos estocásticos simples para el crecimiento de poblaciones.

4.1.1. Procesos de Nacimiento y Muerte.

En esta sección comenzaremos considerando procesos de nacimiento y muerte, que son procesos $\{X_t : t \geq 0\}$ donde la variable aleatoria X_t que indica el tamaño de cierta población en tiempo t puede experimentar saltos positivos como negativos. Las características de estos tipos de procesos son:

1. Si en el tiempo t el sistema se encuentra en el estado x ($x = 1, 2, \dots$) la probabilidad de transición al estado $x + 1$ en el intervalo de tiempo $(t, t + \Delta t)$ es $\lambda_x \Delta t + o(\Delta t)$, es decir

$$P(X_{t+\Delta t} = x + 1 | X_t = x) = \lambda_x \Delta t + o(\Delta t) \quad (4.1)$$

2. Si en el tiempo t el sistema se encuentra en el estado x ($x = 1, 2, \dots$) la probabilidad de transición al estado $x - 1$ en el intervalo de tiempo $(t, t + \Delta t)$ es $\mu_x \Delta t + o(\Delta t)$, es decir:

$$P(X_{t+\Delta t} = x - 1 | X_t = x) = \mu_x \Delta t + o(\Delta t) \quad (4.2)$$

3. La probabilidad de transición de un estado a otro que no sea vecino es $o(\Delta t)$ es decir:

$$P(X_{t+\Delta t} = y | X_t = x) = o(\Delta t) \quad \text{si } y \notin \{x - 1, x + 1\} \quad (4.3)$$

4. La probabilidad de que no se produzca cambio en el estado del proceso es $1 - (\lambda_x + \mu_x)\Delta t + o(\Delta t)$.
5. El estado $x = 0$ es absorbente

donde los coeficientes λ_x, μ_x son funciones del estado x . Notemos $P_x(t) := P(X_t = x)$, entonces las propiedades anteriores implican que:

$$P_x(t + \Delta t) = \lambda_{x-1}P_{x-1}(t)\Delta t + \mu_{x+1}P_{x+1}(t)\Delta t + (1 - (\lambda_x + \mu_x)\Delta t)P_x(t) + o(\Delta t) \quad (4.4)$$

De esta relación se deduce fácilmente que para todo $x = 1, 2, \dots$:

$$\frac{\partial P_x(t)}{\partial t} = \lambda_{x-1}P_{x-1}(t) + \mu_{x+1}P_{x+1}(t) - (\lambda_x + \mu_x)P_x(t) \quad (4.5)$$

y para $x = 0$:

$$\frac{\partial P_0(t)}{\partial t} = \mu_1P_1(t) \quad (4.6)$$

puesto que $\lambda_{-1} = \lambda_0 = \mu_0 = 0$ ya que el estado 0 es absorbente. Si asumimos que en tiempo cero el sistema esta en el estado $x = x_0$ con $0 < x_0 < \infty$ entonces las condiciones iniciales del sistema estan dadas por:

$$P_x(0) = \delta_{x,x_0} = \begin{cases} 1 & \text{si } x = x_0 \\ 0 & \text{si } x \neq x_0 \end{cases}$$

A continuación hallaremos la distribución $P_x(t)$ para el caso particular de los procesos de nacimiento y muerte simples o lineales en los que $\lambda_x = \lambda x$ y $\mu_x = \mu x$ para todo $x \in \mathbb{N}$ y donde $\lambda, \mu > 0$. Con este objetivo, introducimos la función generatriz(ver Apéndice B)

$$F(s, t) = \sum_{x=0}^{\infty} P_x(t)s^x \quad (4.7)$$

De las ecuaciones 4.5 y 4.6 es fácil ver que la función generatriz anterior verifica la siguiente ecuación diferencial:

$$\frac{\partial F(s, t)}{\partial t} = (\lambda s^2 - (\lambda + \mu)s + \mu) \frac{\partial F(s, t)}{\partial s} \quad (4.8)$$

La solución general de 4.8 es

$$F(s, t) = f\left(\frac{\mu - \lambda s}{1 - s}e^{-(\lambda - \mu)t}\right) \quad (4.9)$$

donde $f(\cdot)$ es una función arbitraria. Si asumimos que $X_0 = x_0 = 1$ entonces tenemos que $F(s, 0) = s$ o sea

$$s = f\left(\frac{\mu - \lambda s}{1 - s}\right)$$

y de esta manera

$$f(\zeta) = \frac{\mu - \zeta}{\lambda - \zeta} \quad (4.10)$$

Luego obtenemos que

$$F(s, t) = \frac{\mu(1 - e^{(\lambda-\mu)t}) - (\lambda - \mu e^{(\lambda-\mu)t})s}{\mu - \lambda e^{(\lambda-\mu)t} - \lambda(1 - e^{(\lambda-\mu)t})s} \quad (4.11)$$

Si expandimos 4.11 como serie en potencias de s , los coeficientes de s^x dan $P_x(t)$ que es lo que queremos encontrar. Entonces:

$$P_x(t) = (1 - \alpha(t))(1 - \beta(t))[\beta(t)]^{x-1} \quad (4.12)$$

para $x = 1, 2, \dots$ y

$$P_0(t) = \alpha(t) \quad (4.13)$$

donde

$$\alpha(t) = \frac{\mu(e^{(\lambda-\mu)t} - 1)}{\lambda e^{(\lambda-\mu)t} - \mu}$$

$$\beta(t) = \frac{\lambda(e^{(\lambda-\mu)t} - 1)}{\lambda e^{(\lambda-\mu)t} - \mu}$$

Generalmente en las aplicaciones de los procesos de nacimiento y muerte no son las probabilidades $P_x(t)$ las de mayor interés sino los momentos de X_t y la probabilidad de extinción de la población. Obtenemos entonces diferenciando la función generatriz y procediendo según B.1.1 y B.1.2 las siguientes expresiones para la esperanza y la varianza de la variable X_t :

$$\mathbf{E}(X_t) = e^{(\lambda-\mu)t} \quad (4.14)$$

$$\mathbf{var}(X_t) = \frac{\lambda + \mu}{\lambda - \mu} e^{(\lambda-\mu)t} (e^{(\lambda-\mu)t} - 1) \quad (4.15)$$

De esta manera vemos que:

$$\lim_{t \rightarrow +\infty} \mathbf{E}(X_t) = \begin{cases} 0 & \text{si } \lambda < \mu \\ 1 & \text{si } \lambda = \mu \\ \infty & \text{si } \lambda > \mu \end{cases}$$

lo cual nos dice que si las velocidades de muerte y nacimiento λ, μ son iguales, entonces la velocidad esperada de crecimiento de la población es cero y el tamaño de la población es estacionario.

Otro dato importante que se deduce de los resultados anteriores es la probabilidad de extinción de la población $P_0(t)$, es decir la probabilidad de que la población "desaparezca" en tiempo t :

$$P_0(t) = \alpha(t) = \frac{\mu(e^{(\lambda-\mu)t} - 1)}{\lambda e^{(\lambda-\mu)t} - \mu} \quad (4.16)$$

Entonces la probabilidad de que la población se extinga en algún momento es :

$$\lim_{t \rightarrow +\infty} P_0(t) = \begin{cases} 1 & \text{si } \lambda < \mu \\ \frac{\mu}{\lambda} & \text{si } \lambda > \mu \end{cases}$$

y entonces vemos que como es natural de pensar, la población se extinguirá con probabilidad 1 si la velocidad de muertes es mayor que la velocidad de nacimientos, pero cuando la velocidad de nacimiento es mayor que la de muertes la probabilidad de extinción está dada por $\frac{\mu}{\lambda}$.

4.1.2. Procesos de Nacimiento y Muerte no homogéneos.

Procesos de nacimiento y muerte más generales fueron estudiados por Kendall [7], en los cuales las velocidades de nacimientos y muertes son funciones arbitrarias del tiempo y no funciones lineales de la variable estados como hemos asumido hasta ahora. Estos procesos están caracterizados por la siguiente ecuación diferencial:

$$\frac{\partial P_x(t)}{\partial t} = \lambda(t)(x-1)P_{x-1}(t) + \mu(t)(x+1)P_{x+1}(t) - (\lambda(t) + \mu(t))xP_x(t) \quad (4.17)$$

para $x = 1, 2, \dots$ y

$$\frac{\partial P_0(t)}{\partial t} = \mu(t)P_1(t) \quad (4.18)$$

donde $P_x(t) = P(X_t = x)$ con $x \in \mathbb{N}$ y X_t denota el tamaño de la población en tiempo t . Análogamente al razonamiento hecho en la sección anterior se ve que la solución de este sistema con condición inicial

$$P_x(0) = \begin{cases} 1 & \text{si } x = 1 \\ 0 & \text{si } x \neq 1 \end{cases}$$

es

$$P_x(t) = (1 - \alpha(t))(1 - \beta(t))[\beta(t)]^{x-1} \quad (4.19)$$

para $x = 1, 2, \dots$ y

$$P_0(t) = \alpha(t) \quad (4.20)$$

donde

$$\alpha(t) = 1 - \frac{e^{-\gamma(t)}}{\omega(t)} \quad (4.21)$$

$$\beta(t) = 1 - \frac{1}{\omega(t)} \quad (4.22)$$

con

$$\gamma(t) = \int_0^t (\mu(\tau) - \lambda(\tau)) d\tau \quad (4.23)$$

y

$$\omega(t) = e^{-\gamma(t)} \left(1 + \int_0^t \mu(\tau) e^{\gamma(\tau)} d\tau \right) \quad (4.24)$$

Como se dijo anteriormente, lo interesante de estos procesos es el comportamiento de los momentos y la probabilidad de extinción de la población. De la misma manera que se hizo en la sección anterior, no es difícil de ver que:

$$\mathbf{E}(X_t) = e^{-\gamma(t)} \quad (4.25)$$

$$\mathbf{var}(X_t) = e^{-2\gamma(t)} \int_0^t (\lambda(\tau) + \mu(\tau)) e^{\gamma(\tau)} d\tau \quad (4.26)$$

De esta manera dados $\lambda(t)$ y $\mu(t)$ podemos saber explícitamente las expresiones para la esperanza y la varianza de la variable X_t y estudiar su comportamiento, como por ejemplo se estudio el caso lineal en la sección anterior. Por otro lado, debido a 4.19 y a la definición de $\alpha(t)$ vemos que en este caso general la probabilidad de que la población se extinga en tiempo t es:

$$P_0(t) = \frac{\int_0^t \mu(\tau) e^{\gamma(\tau)} d\tau}{1 + \int_0^t \mu(\tau) e^{\gamma(\tau)} d\tau} \quad (4.27)$$

Observar que en el caso de que $X_0 = x_0 > 1$, la independencia entre los individuos de la población implica que la probabilidad de extinción de la población se obtiene elevando 4.27 a la x_0 -ésima potencia.

Consideremos ahora el “proceso acumulativo” asociado con cierto proceso de nacimiento y muerte. De esta manera, además de la v.a. X_t que representa el número de individuos que hay en la población en tiempo t , introducimos una nueva variable Y_t que representa en número total de nacimientos que se produjeron en la población hasta tiempo t . Entonces, Y_t es un proceso de “nacimiento puro” en el sentido de que un cambio en Y_t es inducido por un nacimiento en la población, es decir por un incremento en una unidad en la variable aleatoria X_t .

Sea entonces $\{X_t, Y_t : t \geq 0\}$ el proceso acumulativo, y notemos

$$P_{x,y}(t) := P(X_t = x, Y_t = y) \quad \forall x = 0, 1, \dots \quad (4.28)$$

Para poder estudiar el proceso bi-variable anterior introducimos la función generatriz

$$G(r, s, t) = \sum_{x=0}^{\infty} \sum_{y=0}^{\infty} P_{x,y}(t) r^x s^y \quad |r| \leq 1, |s| \leq 1 \quad (4.29)$$

y sea $F(r, t)$ la función generatriz asociada al proceso $\{X_t\}_{t \geq 0}$. Sabemos por la sección anterior que se verifica

$$\frac{\partial F(r, t)}{\partial t} = (\lambda r^2 - (\lambda + \mu)r + \mu) \frac{\partial F(r, t)}{\partial r} \quad (4.30)$$

y observando que Y_t solo comparte con X_t los saltos “positivos”, se prueba que:

$$\frac{\partial G(r, s, t)}{\partial t} = (\lambda r^2 s - (\lambda + \mu)r + \mu) \frac{\partial G(r, s, t)}{\partial r} \quad (4.31)$$

En virtud de que el objetivo de este capítulo es presentar algunas aplicaciones de la teoría expuesta en los capítulos anteriores, analizaremos a continuación la solución de la ecuación diferencial 4.31 y algunos otros resultados para el caso particular en que las velocidades de nacimientos y muertes son constantes, y cuando las condiciones iniciales son $X_0 = Y_0 = 1$. Nos interesa entonces hallar los momentos de las variables X_t e Y_t , y para ello introducimos la función generatriz “acumulativa” $K(u, v, t) = \log G(e^u, e^v, t)$. De la ecuación 4.31 se deduce que:

$$\frac{\partial K(u, v, t)}{\partial t} = (\lambda(e^{u+v} - 1) - \mu(1 - e^{-u})) \frac{\partial K(u, v, t)}{\partial u} \quad (4.32)$$

y sabemos por otro lado que:

$$K(u, v, t) = u\mathbf{E}(X_t) + v\mathbf{E}(Y_t) + \frac{u^2}{2}\mathbf{var}(X_t) + uv\mathbf{Cov}\{X_t, Y_t\} + \frac{v^2}{2}\mathbf{var}(Y_t) + \dots \quad (4.33)$$

Expandiendo ambos lados de la ecuación 4.32 en potencias de u y v e igualando coeficientes se obtienen expresiones que relacionan los momentos de las variables, y resolviendo dichas ecuaciones obtenemos que:

$$\mathbf{E}(X_t) = e^{-\gamma(t)} \quad (4.34)$$

$$\mathbf{var}(X_t) = e^{-2\gamma(t)} \int_0^t (\lambda(\tau) + \mu(\tau))e^{\gamma(\tau)} d\tau \quad (4.35)$$

$$\mathbf{E}(Y_t) = 1 + \int_0^t e^{-\gamma(\tau)} \lambda(\tau) d\tau \quad (4.36)$$

$$\mathbf{Cov}\{X_t, Y_t\} = e^{-\gamma(t)} \int_0^t \left(1 + \frac{\mathbf{var}(X_\tau)}{\mathbf{E}(X_\tau)}\right) \lambda(\tau) d\tau \quad (4.37)$$

$$\mathbf{var}(Y_t) = \int_0^t [\mathbf{E}(X_\tau) + 2\mathbf{Cov}\{X_\tau, Y_\tau\}] \lambda(\tau) d\tau \quad (4.38)$$

Para el caso especial en que $\lambda(t) = \lambda$ y $\mu(t) = \mu$ sabemos por los resultados obtenidos en la primera sección de este capítulo como es el comportamiento límite de los primeros momentos de la variable X_t . De las ecuaciones anteriores podemos también estudiar dichos momentos para la nueva variable Y_t obteniendo los siguientes resultados:

$$\lim_{t \rightarrow \infty} \mathbf{E}(Y_t) = \frac{\mu}{\lambda - \mu} \quad (4.39)$$

$$\lim_{t \rightarrow \infty} \mathbf{var}(Y_t) = \frac{\lambda\mu(\lambda + \mu)}{(\mu - \lambda)^3} \quad (4.40)$$

$$\lim_{t \rightarrow \infty} \mathbf{Cov}\{X_t, Y_t\} = 0 \quad (4.41)$$

4.1.3. Procesos simples de Nacimiento, Muerte e Inmigración.

El tipo de proceso de nacimiento y muerte que es de interés en diversas aplicaciones es el proceso de nacimiento, muerte e inmigración. Las características de estos procesos son las mismas que las de los procesos simples de nacimiento y muerte, con la propiedad adicional de que en el intervalo de tiempo $(t, t + \Delta t)$ la probabilidad de que el tamaño de la población aumente una unidad debido a la inmigración desde fuera de la población es $\nu\Delta t + o(\Delta t)$, donde $\nu > 0$ es la velocidad de inmigración. Observar que aunque la inmigración aumenta el tamaño de la población, el efecto de la velocidad de inmigración ν difiere del de la velocidad de nacimientos λ en el sentido de que es independiente del tamaño de la población.

Sea $F_I(s, t)$ la función generatriz asociada a este proceso. Entonces se cumple:

$$\frac{\partial F_I(s, t)}{\partial t} = (\lambda s^2 - (\lambda + \mu)s + \mu) \frac{\partial F_I(s, t)}{\partial s} + \nu(s - 1)F_I(s, t) \quad (4.42)$$

Se puede ver [8] que la ecuación 4.42 se puede escribir de la forma:

$$\frac{\partial F_I(s, t)}{\partial t} = \nu F_I(s, t) (F(s, t) - 1) \quad (4.43)$$

donde $F(s, t)$ es la función generatriz asociada al proceso de nacimiento y muerte. Si $X_0 = 0$ entonces 4.42 o 4.43 serán resueltas con la condición inicial $F_I(s, 0) = 1$. La solución de 4.42 es:

$$F_I(s, t) = \left(\frac{\lambda - \mu}{\lambda e^{(\lambda - \mu)t} - \mu} \right)^{\nu/\lambda} \left[1 - s \left(\frac{\lambda e^{(\lambda - \mu)t} - 1}{\lambda e^{(\lambda - \mu)t} - \mu} \right) \right]^{-\nu/\lambda} \quad \text{si } \lambda \neq \mu \quad (4.44)$$

y

$$F_I(s, t) = (1 + \lambda t - \lambda t s)^{-\nu/\lambda} \quad \text{si } \lambda = \mu \quad (4.45)$$

Derivando las ecuaciones de arriba obtenemos la siguiente expresión para la esperanza del tamaño de la población en tiempo t :

$$\left. \frac{\partial F_I(s, t)}{\partial t} \right]_{s=1} = \mathbf{E}(X_t) = \begin{cases} \frac{\nu}{\lambda - \mu} (e^{(\lambda - \mu)t} - 1) & \text{si } \lambda \neq \mu \\ \nu t & \text{si } \lambda = \mu \end{cases}$$

Entonces:

$$\lim_{t \rightarrow +\infty} \mathbf{E}(X_t) = \begin{cases} \frac{\nu}{\mu - \lambda} & \text{si } \lambda < \mu \\ \infty & \text{si } \lambda \geq \mu \end{cases}$$

4.2. Modelos estocásticos para el crecimiento de población por sexos.

En las secciones anteriores hemos restringido nuestra atención al crecimiento de poblaciones constituidas por una sola clase de individuos. En muchos problemas que son de

interés en el estudio del crecimiento de la población humana por ejemplo, no es el número total de individuos lo que es de principal interés, sino el crecimiento de cada sexo en la población general. En esta sección presentaremos modelos estocásticos para el crecimiento de poblaciones según sexos.

Sean las variables aleatorias X_t e Y_t que representan respectivamente el número de hembras y machos que hay en cierta población en tiempo t , y sea $P_{x,y}(t) := P(X_t = x, Y_t = y)$ con $x, y \geq 0$ la distribución conjunta de X_t e Y_t . Asumimos que la población se desarrolla de acuerdo al siguiente mecanismo:

1. Las poblaciones generadas por dos individuos son independientes.
2. Una hembra viva en tiempo t tiene probabilidad $\lambda p \Delta t + o(\Delta t)$ de reproducir otra hembra y probabilidad $\lambda q \Delta t + o(\Delta t)$ de reproducir un macho en el intervalo de tiempo $(t, t + \Delta t)$, donde $\lambda > 0$ y $p + q = 1$.
3. Una hembra viva en tiempo t tiene probabilidad $\mu \Delta t + o(\Delta t)$ de morir en el intervalo $(t, t + \Delta t)$ y un macho vivo en tiempo t tiene probabilidad $\mu' \Delta t + o(\Delta t)$ de morir en dicho intervalo de tiempo, donde $\mu, \mu' \geq 0$.

De las condiciones anteriores se ve que el proceso estocástico de crecimiento de población según sexos es un proceso bivariable del tipo de nacimiento y muerte. Considerando el caso en que en tiempo $t = 0$ la población está compuesta solo por una hembra y trabajando de la misma manera que en secciones anteriores (utilizando funciones generatrices) se obtienen ecuaciones diferenciales que vinculan los diferentes momentos, y resolviendo dichas ecuaciones se puede estudiar fácilmente cual es el comportamiento asintótico de dichos momentos [9]. Ahora lo interesante en estos casos es también estudiar el comportamiento asintótico del cociente entre el número esperado de machos y hembras de la población. Teniendo en cuenta entonces que las expresiones que se obtienen para los primeros momentos de las variables son:

$$\mathbf{E}(X_t) = e^{(\lambda p - \mu)t} \tag{4.46}$$

$$\mathbf{E}(Y_t) = \frac{\lambda q}{\lambda p - \mu + \mu'} \left(e^{(\lambda p - \mu)t} - e^{-\mu' t} \right) \tag{4.47}$$

se tiene entonces que en el caso en que $(\lambda p - \mu + \mu') > 0$:

$$\lim_{t \rightarrow +\infty} \frac{\mathbf{E}(X_t)}{\mathbf{E}(Y_t)} = \frac{(\lambda p - \mu + \mu')}{\lambda q} \tag{4.48}$$

De esta manera si la probabilidad de muerte de las hembras es igual a la de los machos, es decir $\mu = \mu'$ entonces:

$$\lim_{t \rightarrow +\infty} \frac{\mathbf{E}(X_t)}{\mathbf{E}(Y_t)} = \frac{p}{q} \tag{4.49}$$

Otro punto que puede resultar interesante en este modelo es considerar la distribución de las hembras en la población. De las características de la definición de este proceso y de

la expresión del número esperado de las hembras en la población se ve que el crecimiento de las hembras sigue un proceso simple de nacimiento y muerte.

Sea

$$P_x(t) := P(X_t = x) = \sum_{y=0}^{\infty} P_{x,y}(t) \quad x = 0, 1, \dots \quad (4.50)$$

De las ecuaciones diferenciales válidas para $P_{x,y}(t)$ (ver ref. Bharucha-Reid)se obtiene que:

$$\frac{\partial P_x(t)}{\partial t} = -(\lambda p + \mu)xP_x(t) + \lambda p(x-1)P_{x-1}(t) + \mu(x+1)P_{x+1}(t) \quad x = 1, 2, \dots \quad (4.51)$$

y

$$\frac{\partial P_0(t)}{\partial t} = \mu P_1(t) \quad (4.52)$$

Las condiciones iniciales son

$$P_x(0) = \begin{cases} 1 & \text{si } x = 1 \\ 0 & \text{si } x \neq 1 \end{cases}$$

De la teoría expuesta en las primeras dos secciones de este capítulo sabemos que:

$$P_x(t) = (1 - \alpha(t)) (1 - \beta(t)) [\beta(t)]^{x-1} \quad x = 1, 2, \dots \quad (4.53)$$

y

$$P_0(t) = \alpha(t) \quad (4.54)$$

donde

$$\alpha(t) = \frac{\mu(e^{(\lambda p - \mu)t} - 1)}{\lambda p e^{(\lambda p - \mu)t} - \mu} \quad (4.55)$$

$$\beta(t) = \frac{\lambda p (e^{(\lambda p - \mu)t} - 1)}{\lambda p e^{(\lambda p - \mu)t} - \mu} \quad (4.56)$$

Se tiene entonces que la probabilidad de extinción de la población es:

$$\lim_{t \rightarrow +\infty} P_0(t) = \begin{cases} 1 & \text{si } \lambda p \leq \mu \\ \frac{\mu}{\lambda p} & \text{si } \lambda p > \mu \end{cases}$$

Los modelos que hemos estado describiendo a lo largo de las secciones de este capítulo son algunos modelos importantes a la hora de estudiar crecimiento de poblaciones, pero existen muchos mas. Además de éstos, modelos de crecimiento de población de gran importancia en ecología por ejemplo son aquellos que describen el crecimiento de poblaciones constituídas por 2 o mas clases de individuos que interactúan entre si (ver por ejemplo [10] o [11]), o modelos de migración de población en los cuales individuos de una región pueden migrar a otra. También hay procesos que modelan el desarrollo de cierta epidemia en una población, o modelos para el crecimiento de poblaciones heterogéneas donde un tipo de organismo puede “mutar” en un organismo diferente. En general son muchas las aplicaciones que los procesos estocásticos de tiempo continuo tienen en la biología pero estas aplicaciones no se restringen solo a dicho campo sino que estos procesos tambien son de gran ayuda en física, química, teoría de colas, etc. Por mas aplicaciones citamos por ejemplo [9].

Apéndice A

Proceso Casa de Cartas

En esta sección daremos la definición y veremos algunas propiedades de un proceso estocástico que ha sido utilizado en el transcurso de estas notas. Este proceso, al cual hemos llamado Proceso Casa de Cartas, posee espacio de estados el conjunto de los números naturales \mathbb{N} y se define de la siguiente manera. Consideremos una sucesión creciente $\{a_k\}_{k \in \mathbb{N}} \subset [0, 1]$, tal que verifica que

$$\lim_{k \rightarrow \infty} a_k = 1$$

Definimos la familia de cadenas $\{W_n^m : n \geq m\} : m \in \mathbb{Z}\}$ de manera que $W_m^m := 0$ y para $n > m$:

$$W_n^m := (W_{n-1}^m + 1) \mathbf{1}_{\{U_n < a_{W_{n-1}^m}\}} \quad (\text{A.1})$$

donde $\{U_n\}_{n \in \mathbb{N}}$ es una sucesión de variables aleatorias independientes uniformemente distribuidas en el intervalo $[0, 1]$. De esta manera W_n^m denota la posición en el tiempo n de la cadena que empieza en tiempo m en el origen.

Observación A.0.1. *El nombre Casa de Cartas dado al proceso refleja el hecho de que una vez que la cadena esta en el estado n puede en el siguiente instante saltar al estado $n + 1$ o puede volver al estado inicial 0 , como cuando se arma una casa con cartas, agregando una a una una nueva carta en cada paso.*

Dados $x, y \in \mathbb{N}$ dos estados posibles del proceso, veamos a continuación como son las probabilidades de transición de este proceso. Dividimos el estudio en tres casos:

1. Sabemos que el proceso puede saltar de un estado x al estado $x + 1$ o volver al origen, por lo tanto es claro que si $y \notin \{0, x + 1\}$ entonces:

$$p(x, y) = P(W_n^m = y | W_{n-1}^m = x) = 0 \quad (\text{A.2})$$

2. Supongamos ahora que $y = x + 1$, entonces en este caso, debido a la definición del proceso y a la distribución de las variables aleatorias U_n tenemos que:

$$p(x, x + 1) = P(W_n^m = x + 1 | W_{n-1}^m = x) = P(U_n < a_x) = a_x \quad (\text{A.3})$$

3. Por último

$$p(x, 0) = P(W_n^m = 0 | W_{n-1}^m = x) = P(U_n > a_x) = 1 - a_x \quad (\text{A.4})$$

Entonces concluimos que:

$$p(x, y) = \begin{cases} a_x & \text{si } y = x + 1 \\ 1 - a_x & \text{si } y = 0 \\ 0 & \text{en otro caso} \end{cases}$$

Observación A.0.2. *Observemos que el hecho de que $a_k \nearrow 1$ implica que:*

$$W_n^m \geq W_n^k \quad \forall m < k \leq n \quad (\text{A.5})$$

por lo tanto $W_n^m = 0$ implica que $W_n^k = 0$ para todo $m < k \leq n$. Se cumple entonces que:

$$W_n^m = 0 \implies W_t^m = W_t^k \quad \forall m \leq k \leq n \leq t \quad (\text{A.6})$$

El siguiente resultado describe el comportamiento del proceso en función de sus probabilidades de transición.

Lema A.0.3. *Sea $\{W_n^m : n \geq m\}$ una cadena de Markov con valores en \mathbb{N} con matriz de transición $Q = (Q(x, y))_{x, y \in \mathbb{N}}$ tal que:*

$$Q(x, y) = \begin{cases} a_x & \text{si } y = x + 1 \\ 1 - a_x & \text{si } y = 0 \\ 0 & \text{en otro caso} \end{cases}$$

Entonces:

1. *La cadena es recurrente no positiva $\iff \sum_{n \geq 0} \prod_{k=0}^n a_k = \infty$*
2. *La cadena es transitoria $\iff \prod_{k=0}^{\infty} a_k > 0$*

Apéndice B

Funciones Generatrices

El método de funciones generatrices es una herramienta importante en el estudio de procesos estocásticos con espacio de estados discreto. En este apéndice se reúnen algunos resultados concernientes a las funciones generatrices que son útiles para el desarrollo del capítulo 4 presentado en este trabajo.

Definición B.0.4. *Consideremos la sucesión de números reales $\varphi_0, \varphi_1, \dots$. La serie de potencias:*

$$F(s) = \sum_{i=1}^{\infty} \varphi_i s^i \quad (\text{B.1})$$

puede ser considerada como la transformación que lleva la sucesión $\{\varphi_i\}_{i \in \mathbb{N}}$ en la función $F(s)$. Si dicha serie de potencias converge en cierto intervalo $-k < s < k$ entonces la función $F(s)$ es llamada función generatriz de la sucesión $\{\varphi_i\}_{i \in \mathbb{N}}$.

Consideremos ahora el proceso estocástico $\{X_t\}_{t \geq 0}$ y la sucesión de probabilidades $P_x(t) := P(X_t = x)$ donde el parámetro de tiempo t puede ser discreto o continuo. Sea

$$F(s, t) = \sum_{x=0}^{\infty} P_x(t) s^x \quad (\text{B.2})$$

la función generatriz de las probabilidades $P_x(t)$ con $x = 0, 1, \dots$. Observar que como las probabilidades $P_x(t)$ están acotadas para todo x y para todo t y $\sum_{x=0}^{\infty} P_x(t) = 1$ para todo $t \geq 0$, comparando la serie B.2 con la serie geométrica vemos que B.2 converge uniformemente con t , al menos para $|s| < 1$.

B.1. Algunos teoremas y propiedades.

Teorema B.1.1. *El valor esperado de la variable X_t está dado por:*

$$\mathbf{E}(X_t) = \left. \frac{\partial F(s, t)}{\partial s} \right]_{s=1} \quad (\text{B.3})$$

Teorema B.1.2. Si $\mathbf{E}(X_t^2) = \sum_{x=0}^{\infty} x^2 P_x(t) < \infty$, entonces:

$$\mathbf{E}(X_t^2) = \left[\frac{\partial^2 F(s, t)}{\partial s^2} + \frac{\partial F(s, t)}{\partial s} \right]_{s=1} \quad (\text{B.4})$$

y por definición de varianza se tiene entonces que:

$$\mathbf{var}(X_t) = \left[\frac{\partial^2 F(s, t)}{\partial s^2} + \frac{\partial F(s, t)}{\partial s} - \left(\frac{\partial F(s, t)}{\partial s} \right)^2 \right]_{s=1} \quad (\text{B.5})$$

Teorema B.1.3. Sean $X_1(t)$ y $X_2(t)$ dos variables aleatorias independientes con distribuciones $P_x(t)$ y $Q_x(t)$ y funciones generatrices $F_1(s, t)$ y $F_2(s, t)$ respectivamente. Sea $Y(t) = X_1(t) + X_2(t)$ y sea $G(s, t)$ la función generatriz de las probabilidades $R_y(t) := P(Y(t) = y)$. Entonces $G(s, t)$ está dada por:

$$G(s, t) = F_1(s, t)F_2(s, t) \quad (\text{B.6})$$

Para terminar presentamos a continuación un resultado que es de utilidad en muchos de los problemas que utilizan funciones generatrices.

Si $F(s, t)$ es analítica en $s = 0$, entonces se puede escribir F como la serie de Maclaurin:

$$F(s, t) = F(0, t) + \sum_{x=1}^{\infty} \frac{F^{(x)}(0, t)}{x!} s^x \quad (\text{B.7})$$

Entonces por comparación con las series $F(s, t) = \sum_{x=0}^{\infty} P_x(t)s^x$ se tiene que:

$$P_0(t) = F(0, t) \quad (\text{B.8})$$

y

$$P_x(t) = \frac{F^{(x)}(0, t)}{x!} \quad x = 1, 2, \dots \quad (\text{B.9})$$

Bibliografía

- [1] Pablo Ferrari, Antonio Galvez, *Coupling and Regeneration for stochastic processes*, Notes for a minicourse presented in XIII Escuela Venezolana de Matematicas, 2000.
- [2] Pablo Ferrari, Antonio Galvez, *Acoplamiento em processos estocasticos*, SBM, IMPA, Rio de Janeiro, 1997.
- [3] Francis Comets, Roberto Fernandez, Pablo A. Ferrari, *Processes with Long Memory: Regenerative Construction and Perfect Simulation*, Ann. Appl. Probab. vol. 12 3:921-943, 2002.
- [4] Sheldon M. Ross, *Stochastic Processes*, Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. Lectures in Mathematics, 14. John Wiley and Sons, Inc., New York , 1983.
- [5] H. Berbee, *Chains with complete connections: Uniqueness and Markov representation*, Prob. Th. Rel. Fields, 76:243-253.
- [6] Sidney I. Resnick, *Adventures in Stochastic Processes*, Birkhauser Boston, 1992.
- [7] D. G. Kendall, *Les Processus Stachistiques de croissance en biologie*, Ann. inst. H. Poincaré, vol.13, pp.43-108, 1952.
- [8] D. G. Kendall, *Stochastic Processes and Population Growth*, J. Roy. Statist. Soc., ser. B, vol. 11, pp. 230-264, 1949.
- [9] A. T. Bharucha-Reid, *Elements of the Theory of Markov Processes and Their applications*, McGraw-Hill Book Company, Inc, New York, 1960.
- [10] A. J. Lotka, *Elements of Physical Biology*, The williams and Wilkins Company, Baltimore, 1925.
- [11] V. Volterra, *Leçons sur la théorie mathématique de la lutte pour la vie*, Gauthier-Villars, Paris, 1931.