UNIVERSIDAD DE BUENOS AIRES
Facultad de Ciencias Exactas y Naturales
Departamento de Matemática

Tesis de Licenciatura

Cadenas de Markov y ecuaciones diferenciales ordinarias: un acoplamiento para probar convergencia

Mateo Astelarra

Director: Pablo Groisman

21 de Diciembre de 2010

Agradecimientos

Quiero agradecerle especialmente a Patu por haber sido tan buen director. Por la constante buena onda, ayuda y motivación que me dió. Quería hacer extensivo este agradecimiento a todo el grupo de probabilidad por generar un espacio en donde dan ganas de trabajar y seguir estudiando.

Fuera de eso, creo que a lo largo de la carrera hubo muchas personas que me ayudaron, apoyaron, guiaron y enseñaron algo de alguna manera u otra. Entre ellos docentes, compañeros de trabajo, familiares, amigos y alumnos. Creo que nombrarlos acá no tiene mucho sentido pero espero haber sido lo suficientemente agradecido en el día a día y seguir siendo agradecido con esas personas, y con las nuevas personas que me cruce, de acá en más.

Índice general

1. Preliminares **9**

1.1. Procesos estocásticos y cadenas de Markov 9

1.1.1. Cadenas de Markov a tiempo continuo 11

1.2. Procesos de Poisson 15

1.3. Desigualdad de Gronwall 17

1.4. Método de Euler 18

1.4.1. Análisis del error y convergencia 18

2. Familias que dependen de una densidad **23**

2.1. Modelos poblacionales 23

2.2. Modelos epidémicos 29

2.2.1. Modelo SIS 29

2.2.2. Modelo SIR 30

2.3. Teoría de Colas 31

2.3.1. Cola M/M/1 33

2.3.2. Cola M/M/∞ 34

2.4. Algunos resultados conocidos 35

3. Convergencia **37**

3.1. Idea de la prueba 37

3.2. Construcción gráfica de los procesos 39

3.3. Aproximación de Z_t^V a X_t^V 40

3.4. Aproximación de Z_t^V a Y_t^V 44

3.5. Aproximación de Y_t^V a $x(t)$ 46

3.6. Prueba del Teorema 49

3.7. Aplicaciones 50

4. Apéndice **53**
4.1. Prueba de Darling y Norris 53
4.1.1. Aplicación del Teorema a un modelo epidémico 56
4.2. Las simulaciones 57

Introducción

Consideremos el problema de modelar la proporción de la población infectada por cierta enfermedad para estudiar si puede transformarse en una epidemia. Buscamos un modelo para analizar cómo cambia la proporción de individuos enfermos y sanos. Una forma de hacerlo es plantear una ecuación diferencial ordinaria (ODE) que relacione ambas clases de individuos mediante ciertas tasas de infección, de muerte y de nacimiento. Así, planteamos dos variables y sus derivadas en función de esas dos variables y las tasas mencionadas. Una vez planteado el sistema de ecuaciones, podemos estudiar su comportamiento para que las trayectorias que son solución nos den alguna información de lo que podría pasar con la población que queremos estudiar.

Otra forma de hacerlo es asumiendo que la propagación de la enfermedad tiene un carácter aleatorio. Habiendo asumido esto, tenemos la necesidad de plantear un modelo que no sea determinístico. Así, podemos pensar que tenemos una población de tamanõ fijo y que con alguna ley van cambiando la cantidad de individuos sanos e infectados a través del tiempo. Parece adecuado pensar que la ley de cambio de los individuos depende solo de la cantidad de individuos infectados que hay en ese momento y de algunos parámetros. Necesitamos saber también los tiempos en los cuales va a haber cambios, y observamos que, al haber finitos estados no puede haber más de numerables cambios. Teniendo esto en cuenta, un modelo aleatorio adecuado son las cadenas de Markov a tiempo continuo. Necesitamos especificar entonces la ley del tiempo de espera entre cambios y la ley de decisión para saber cómo va a cambiar la población. Todo esto se presenta adecuadamente en el Capítulo 1 pero adelantamos acá que para tener esa información sólo necesitamos una matriz que llamaremos matriz de tasas.

Una pregunta que surge naturalmente es: ¿obtendremos con los dos modelos los mismos resultados? Para responderla Thomas Kurtz introdujo en una publicación de 1970 intitulada “Ordinary differential equations as limits of pure jump Markov processes” (ver [4]), el concepto de familias de cadenas de Markov que

dependen de una densidad. ¿Por qué es importante este concepto? Porque nos permite escribir tanto la ecuación diferencial como la matriz de tasas de la cadena en términos de una única función f. Más aun, a partir de esta función podemos definir también una función F que nos sirva para escribir a la ecuación diferencial como $\dot{x}(t) = F(x(t))$ y que además nos sirva para describir la dirección en la cual va en media la cadena de Markov. Este hecho, y el hecho de que podamos describir los dos modelos con una misma función nos hace pensar que la respuesta a nuestra pregunta es afirmativa. Vamos a desarrollar entonces en el Capítulo 2 el concepto de familias de cadenas de Markov que dependen de una densidad e introducir diferentes problemas que caen dentro de esta categoría. Entre ellos, el ya mencionado problema de epidemias, modelos poblacionales de tipo predador-presa y algunos problemas que aparecen en teoría de colas. En ese capítulo también presentamos algunas simulaciones comparando los modelos determinísticos con los estocásticos.

La respuesta a la pregunta antes mencionada es afirmativa. Fue contestada por Kurtz en la publicación que mencionamos en el párrafo anterior. El objetivo de esta tesis es dar una prueba alternativa a la de Kurtz intentando usar herramientas más simples que permitan comprender mejor el fenómeno. La prueba que damos esta basada en acoplamientos, ¿qué es un acoplamiento? Es la construcción de dos o más cadenas de Markov en un mismo espacio de probabilidad de forma tal de poder probar propiedades que involucran a las distribuciones de estas cadenas a partir de propiedades sobre las trayectorias de las mismas. El principal resultado de esta tesis se encuentra en el Capítulo 3, donde se prueba que dada una familia de cadenas de Markov que depende de una densidad y su ecuación diferencial correspondiente, la familia de cadenas de Markov tiende en L^1 a la solución de la ecuación diferencial. Para eso, definimos dos nuevas familias de procesos estocásticos. La primera va saltando con la media de como salta la familia original. La segunda va saltando, en tiempos aleatorios, como el método de Euler para la ODE. Es de esperarse que la primera se parezca a la familia original y que la segunda se parezca a la solución de la ecuación. Probamos eso y que estas dos nuevas familias también se parecen entre sí para probar finalmente mediante la desigualdad triangular el resultado original.

En el apéndice contamos una prueba del resultado original que se basa en el uso de martingalas y la escritura de ambos modelos de una forma en particular que permite relacionarlos fácilmente. Esta prueba fue dada por Darling y Norris en [2]. La prueba de Kurtz puede hallarse en [4].

Capítulo 1

Preliminares

En este capítulo introducimos algunos conceptos que usaremos a lo largo de la tesis. Entre ellos, el de cadenas de Markov y el de procesos de Poisson. A su vez, enunciamos y demostramos algunos resultados que vamos a necesitar para los capítulos siguientes.

1.1. Procesos estocásticos y cadenas de Markov

Un proceso estocástico puede interpretarse como la evolución de algun fenómeno a través del tiempo y las cadenas de Markov son un proceso estocástico. En esta sección vamos a definir primero las cadenas de Markov pensando el tiempo como un conjunto discreto para poder después definirlas pensando el tiempo como un intervalo.

Definición 1.1.1 *Sea $(\Omega, \mathscr{F}, \mathbb{P})$ un espacio de probabilidad. Un proceso estocástico es un conjunto de variables aleatorias $(X_j)_{j\in J}$ definidas en ese espacio . Si el proceso representa la evolución temporal de un sistema aleatorio entonces vamos a tomar a $J = \mathbb{N}_0$ (caso discreto) o $J = [0, +\infty)$ (caso continuo).*

Definición 1.1.2 *Sea S un conjunto finito o numerable. Decimos que P es una matriz estocástica si $(P(x,y))_{x,y\in S}$ son números mayores o iguales a cero tal que $\Sigma_{y\in S} P(x,y) = 1$.*

Definición 1.1.3 *Sean $(X_n)_{n\in\mathbb{N}}$ variables aleatorias definidas en $(\Omega, \mathscr{F}, \mathbb{P})$ y P una matriz estocástica. Decimos que $(X_n)_{n\in\mathbb{N}}$ es una cadena de Markov homogénea*

a tiempo discreto con matriz de transiciones P sii

$$\mathbb{P}(X_n = y | X_{n-1} = x, ..., X_1 = x_1, X_0 = x_0) = P(x,y) \tag{1.1.1}$$

para todo $x,y \in S$ y $x_0, x_1, ..., x_{n-2}, x_{n-1} \in S$ para los cuales el evento al cual condicionamos tiene probabilidad positiva.

Observación 1.1.4 *Notar que la ecuación* (1.1.1) *determina la distribución del proceso $(X_n)_{n\in\mathbb{N}}$.*

Notar que la idea intuitiva de las cadenas de Markov es que para decidir a dónde vamos a ir lo único que nos importa es donde estamos. O sea, el futuro es independiente del pasado si condicionamos al presente.

A lo largo de este capítulo cuando hablamos de una cadena de Markov vamos a llamar S al espacio de estados y vamos a suponer que S es a lo sumo numerable.

Teorema 1.1.5 (Construcción y simulación de cadenas de Markov) *Sea $(X_n)_{n\in\mathbb{N}}$ un proceso estocástico. Entonces $(X_n)_{n\in\mathbb{N}}$ es una cadena de Markov con matriz de transiciones P sii existen variables $(\widetilde{X}_n)_{n\in\mathbb{N}}$ con la misma distribución conjunta que $(X_n)_{n\in\mathbb{N}}$, una función $F : S \times [0,1] \to S$ y $(U_i)_{i\in\mathbb{N}}$ variables aleatorias independientes con distribución uniforme en el $[0,1]$, tales que*

$$\widetilde{X}_{n+1} = F(\widetilde{X}_n, U_{n+1}) \text{ para todo } n \geq 1.$$

Prueba.

$\Rightarrow$) Para un x fijo consideremos una partición del intervalo $[0,1]$ en intervalos $I(x,y)$ para $x,y \in S$ tal que la medida de $I(x,y)$ es $P(x,y)$. Definimos

$$F(x,u) = y \text{ si } u \in I(x,y),$$

$$\widetilde{X}_1 = X_1 \text{ y}$$

$$\widetilde{X}_{n+1} = F(\widetilde{X}_n, U_{n+1}) \text{ para todo } n \geq 1.$$

Como F es medible y X_n, U_n son medibles para todo $n \geq 1$, $\widetilde{X}_n$ resulta medible para todo $n \geq 1$. Tenemos entonces que

$$\begin{aligned}
\mathbb{P}(\widetilde{X}_{n+1} = x_{n+1} \quad | \quad & \widetilde{X}_0 = x_0, ..., \widetilde{X}_n = x_n) \\
= \quad & \mathbb{P}(F(x_n, U_{n+1}) = x_{n+1} | \widetilde{X}_0 = x_0, ..., \widetilde{X}_n = x_n) \\
= \quad & \mathbb{P}(F(x_n, U_{n+1}) = x_{n+1} | \widetilde{X}_n = x_n) \\
= \quad & \mathbb{P}(U_{n+1} \in I(x_n, x_{n+1})) \\
= \quad & P(x_n, x_{n+1}).
\end{aligned}$$

donde en la tercera igualdad usamos que la variable U_{n+1} es independiente de $(U_i)_{i\leq n}$ y, por lo tanto, también lo es de $(\widetilde{X}_i)_{i\leq n}$ por construcción.

$\Leftarrow$) Sea U una variable aleatoria con distribución uniforme en el $[0,1]$, definimos para $x,y \in S$, $P(x,y) = \mathbb{P}(F(x,U) = y)$. Vale entonces que $P(x,y) \geq 0$ para todo $x,y \in S$. Vale también que

$$\sum_{y\in S} P(x,y) = \sum_{y\in S} \mathbb{P}(F(x,U) = y) = \sum_{y\in S} \mathbb{P}(U \in I(x,y)) = 1.$$

Ahora, con la misma cuenta que antes vemos que

$$\begin{aligned} \mathbb{P}(\widetilde{X}_{n+1} = x_{n+1} | \widetilde{X}_0 = x_0, ..., \widetilde{X}_n = x_n) &= \mathbb{P}(\widetilde{X}_{n+1} = x_{n+1} | \widetilde{X}_n = x_n) \\ &= P(x_n, x_{n+1}). \end{aligned}$$

Entonces $(\widetilde{X}_n)_{n\in\mathbb{N}}$ es una cadena de Markov con matriz de transiciones P y, por lo tanto, $(X_n)_{n\in\mathbb{N}}$ también.

■

1.1.1. Cadenas de Markov a tiempo continuo

La idea ahora es construir una cadena de Markov a tiempo continuo $(X_t)_{t\in[0,+\infty)}$ en un espacio de estados S finito o numerable. Como S es numerable la cadena se tiene que mover de a saltos, no se puede mover de manera continua. Con lo cual, el proceso tiene que quedarse un tiempo aleatorio en el lugar x y luego elegir aleatoriamente un lugar y al cual saltar. Una vez que está en y espera un tiempo aleatorio (que en principio puede depender del lugar en el que está), después elige un lugar z a cual saltar, y así sucesivamente. Entonces, dado un estado inicial tenemos que especificar la distribución de los tiempos de espera en cada estado y la regla de decisión para saltar de un estado al otro.

Ahora, la propiedad de Markov dice que lo único que importa es el estado en el que estoy, no importa cuanto tiempo ya estuve en este estado. Deducimos entonces que la distribución de los tiempos de espera tiene la propiedad de pérdida de memoria y, por lo tanto, es exponencial. Como el parámetro puede depender del lugar en donde estoy lo llamamos, para cada x, $c(x)$. En definitiva, si estamos en el lugar x en un momento determinado, el tiempo T que vamos a seguir en ese estado cumple que $\mathbb{P}(T > t) = e^{-c(x)t}$. A su vez, la propiedad de Markov nos dice que la regla de decisión para pasar de un estado al otro solo puede depender del estado en donde estoy. Teniendo eso en cuanta, lo único que nos falta es una matriz

estocástica P tal que la probabilidad de pasar de x a y sea $P(x,y)$. Notemos que si estamos en el estado x una forma de decidir a que estado saltar es dividiendo el intervalo $[0,1]$ en intervalos disjuntos $I(x,y)$ de longitud $P(x,y)$ y luego usar una variable uniforme de la siguiente manera, si la variable cae en el intervalo $I(x,y)$ entonces el proceso salta al estado y.

Esto sugiere que para construir una cadena de Markov a tiempo continuo $(X_t)_{t\in[0,+\infty)}$ podemos definir una cadena de Markov a tiempo discreto $(Y_n)_{n\in\mathbb{N}_0}$ con matriz de transiciones P, y ajustar los tiempos de espera para que sean exponenciales de parámetro $c(x)$. Llamamos a $(Y_n)_{n\in\mathbb{N}_0}$ el esqueleto de X_t. Más precisamente, la secuencia de estados que visita $(X_t)_{t\geq 0}$ es $Y_0, Y_1, Y_2, Y_3....$ y los tiempos de espera son $\tau_n = c(Y_n)^{-1}E_n$, con $(E_n)_{n\in\mathbb{N}}$ v.a.i.i.d con distribución exponencial y media 1. Para que se cumpla la propiedad de Markov tenemos que pedir que, dado Y_n, la variable τ_n sea independiente de $(\tau_k, Y_k)_{0\leq k\leq n-1}$. Entonces, si definimos $T_0 = 0$ y $T_n = \tau_0 + ... + \tau_n$ para $n \geq 1$ tenemos que

$$X_t = Y_n \text{ para } T_n \leq t < T_{n+1} \text{ para } n = 1,2,3,....$$

Dicho en palabras, X_t espera un tiempo τ_n en Y_n, y después salta a Y_{n+1}. X_t está definido para todo $t \in \mathbb{R}_{\geq 0}$ si $T \nearrow \infty$ cuando $n \nearrow \infty$. Esto pasa en casi todo punto si, por ejemplo, asumimos que existe C tal que $c(x) \leq C$ para todo $x \in S$. Notemos también que las trayectorias de X_t son continuas a derecha.

Por último, vamos a estudiar cómo es el comportamiento infinitesimal del proceso. Definimos el operador lineal L cuyo dominio es el espacio de funciones acotadas en S de la siguiente forma

$$L(f)(x) = c(x)\sum_{y\in S} P(x,y)(f(y) - f(x))$$

Este operador se llama el generador del proceso $(X_t)_{t\geq 0}$ y se puede interpretar de la siguiente manera: empezando en el estado x , el próximo salto ocurre con tasa $c(x)$ y, cuando ocurre, saltar al estado y ocurre con probabilidad $P(x,y)$. En ese caso, el valor de f cambia en $f(y) - f(x)$.

Otra forma de describir el proceso es mediante la tasas instantáneas. Estas tasas vienen dadas por una matriz $Q = (q(x,y))_{x,y\in S}$ definida por

$$q(x,y) = c(x)P(x,y) \text{ para } y \neq x,$$

$$q(x,x) = -\sum_{y\neq x} q(x,y).$$

Así, suponiendo que $P(x,x) = 0$ tenemos que

$$c(x) = \sum_{x \neq y} c(x)P(x,y) = \sum_{x \neq y} q(x,y).$$

Luego, dada la matriz Q queda determinada la distribución del proceso. Para simplificar la notación vamos a llamar $q(x)$ a $c(x)$.

Definición 1.1.6 (Cadenas de Markov a tiempo continuo) *Sea $(X_t)_{t\geq 0}$ un proceso estocástico con trayectorias continuas a derecha en un espacio de estados S. Decimos que $(X_t)_{t\geq 0}$ es una cadena de Markov a tiempo continuo con matriz de tasas Q, si su esqueleto $(Y_n)_{n\in\mathbb{N}}$ es una cadena de Markov a tiempo discreto con matriz de transiciones $P = (P(x,y), x,y \in S)$ dada por*

$$P(x,y) = \frac{q(x,y)}{q(x)} \text{ si } x \neq y,\ P(x,x) = 0.$$

Además se tiene que cumplir que condicional a $Y_0, Y_1 ...., Y_{n-1}$ los tiempos de salto $\tau_1, \tau_2, ..., \tau_n$ son variable aleatorias independientes con distribución exponencial de parámetro $q(Y_i)$ con $i = 0, 1, ..., n-1$ respectivamente.

Definición 1.1.7 *Sea $(X_t)_{t\geq 0}$ una cadena de Markov a tiempo continuo en espacio de estados $S \subseteq \mathbb{R}^d$ a lo sumo numerable con matriz de tasas Q. Definimos el drift de $(X_t)_{t\geq 0}$ como*

$$\beta(x) = \sum_{x' \in S:\ x' \neq x} (x' - x) q(x,x')$$

Observación 1.1.8 *Sea $(X_t)_{t\geq 0}$ una cadena de Markov a tiempo continuo en un espacio de estados $S \subseteq \mathbb{R}^d$ a lo sumo numerable con matriz de tasas Q. Notemos que si f_i con $i = 1, ..., d$ son las proyecciones en la i-ésima coordenada respectivamente entonces*

$$\beta(x) = (L(f_1)(x), ..., L(f_d)(x)).$$

Esto dice que el drift da el comportamiento infinitesimal de $(X_t)_{t\geq 0}$ coordenada a coordenada. Indica hacia donde va el proceso en media condicional a que estamos en el estado x.

A continuación probamos un resultado que será de utilidad más adelante.

Lema 1.1.9 *Sea Q una matriz de tasas y supongamos que $C = \sup_{x\in S}\{q(x)\} < \infty$ y sea $q \geq C$. Sea $(X_t)_{t\geq 0}$ construido de la siguiente forma: si X_t esta en lugar x espera un tiempo exponencial de parámetro q, salta al lugar y con probabilidad $\frac{q(x,y)}{q}$ y se queda en el lugar con probabilidad $1-\Sigma_{y\neq x}\frac{q(x,y)}{q}$. Entonces $(X_t)_{t\geq 0}$ es una cadena de Markov con matriz de tasas Q.*

Prueba. Sea $p(x,x) = 1-\Sigma_{y\neq x}\frac{q(x,y)}{q} = 1-\frac{q(x)}{q}$. Sean $(U_n)_{n\in\mathbb{N}}$ variables aleatorias independientes distribuidas uniformemente en el $[0,1]$ y supongamos que X_t esta en el lugar x. Una forma de elegir a donde saltar es dividir el intervalo $[0,1]$ en intervalos disjuntos $I(x,y)$ de longitud $\frac{q(x,y)}{q}$ si $x \neq y$ y $p(x,x)$ si $x = y$. Entonces X_t salta al estado y si la variable uniforme cae en $I(x,y)$ y se queda si la variable uniforme cae en $I(x,x)$. Una forma de dividir el $[0,1]$ es tomando el $[0,p(x,x))$ y después dividir $[p(x,x),1]$ adecuadamente. Así, para que X_t salga del lugar x necesitamos que $U_i \geq p(x,x)$. Luego si $N = \text{mín}\{i\colon U_i \geq p(x,x)\}$ lo que tarda X_t en salir del lugar de donde está es $E = \sum_{i=1}^{N} E_i$ con $(E_i)_{i\in\mathbb{N}}$ v.a.i.i.d con distribución exponencial de parámetro q . Usando que $E|_{N=k} \sim \Gamma(k,q)$ veamos que E es exponencial de parámetro $q(x)$

$$\begin{aligned}
\mathbb{P}(E \geq r) &= \sum_{k=1}^{\infty} \mathbb{P}(E \geq r|_{N=k})\mathbb{P}(N=k) \\
&= \sum_{k=1}^{\infty} \int_r^{\infty} \frac{q^k}{\Gamma(k)} s^{k-1} e^{-qs} ds p(x,x)^{k-1}(1-p(x,x))) \\
&= (1-p(x,x))q \int_r^{\infty} \sum_{i=0}^{\infty} \frac{(qsp(x,x))^i}{(i)!} e^{-qs} ds \\
&= (1-p(x,x))q \int_t^{\infty} e^{-(1-p(x,x))qs} ds \\
&= e^{-(1-p(x,x))qr} \\
&= e^{-q(x)r}.
\end{aligned}$$

Ahora como en ese momento, la variable uniforme correspondiente cae en el intervalo $(p(x,x),1]$ y para saltar al lugar y tiene que caer en un intervalo de longitud $\frac{q(x,y)}{q}$ entonces la probabilidad de saltar al lugar y es $\frac{q(x,y)}{q}\frac{q}{q(x)} = \frac{q(x,y)}{q(x)}$.

■

1.2. Procesos de Poisson

En esta sección definimos los Procesos de Poisson, enunciamos algunas equivalencias de la definición y demostramos algunas propiedades que nos serán de utilidad más adelante.

Definición 1.2.1 *Un proceso de Poisson con tasa $\lambda > 0$ es una cadena de Markov $(N_t)_{t\geq 0}$ con espacio de estados $\mathbb{N}_0$ y matriz de tasas Q dada por $q(j, j+1) = \lambda$ y $q(j, j) = -\lambda$.*

En palabras, un proceso de Poisson va saltando hacia arriba de a un paso y el tiempo de espera entre salto y salto es exponencial con parámetro λ. N_t es la cantidad de saltos que hubo hasta tiempo t. Nos van a ser útiles las siguientes equivalencias.

Teorema 1.2.2 *Sea $(X_t)_{t\geq 0}$ un proceso estocástico a valores enteros, creciente, con trayectorias continuas a derecha y tal que $X_0 = 0$. Sea $\lambda > 0$. Entonces las siguientes condiciones son equivalentes*

a) Los tiempos de espera $(\tau_i)_{i\in\mathbb{N}}$ son v.a.i.i.d con distribución exponencial de parámetro λ y el esqueleto de $(X_t)_{t\geq 0}$ viene dado por $Y_n = n$ para todo $n \in \mathbb{N}$.

b) $(X_t)_{t\geq 0}$ tiene incrementos independientes (es decir, dado $n \in \mathbb{N}$ si $s_1 < t_1 < \ldots < s_n < t_n$ entonces las variables aleatorias $X_{t_1} - X_{s_1}$,....,$X_{t_n} - X_{s_n}$ son independientes) y, además cuando $h \searrow 0$,

$$\mathbb{P}(X_{t+h} - X_t = 0) = 1 - \lambda h + \mathrm{o}(h), \tag{1.2.1}$$

$$\mathbb{P}(X_{t+h} - X_t = 1) = \lambda h + \mathrm{o}(h), \tag{1.2.2}$$

uniformemente en t.

c) $(X_t)_{t\geq 0}$ tiene incrementos estacionarios e independientes y, para cada $s, t \in \mathbb{R}_{\geq 0}$ con $s < t$, la variable aleatoria $N_t - N_s$ tiene distribución Poisson de parámetro $\lambda(t-s)$.

Si $(X_t)_{t\geq 0}$ satisface cualquiera de estas tres condiciones resulta un Proceso de Poisson con tasa λ.

Prueba. Ver [5, pag.76]. ■

Corolario 1.2.3 *Sea $(X_t)_{t\geq 0}$ un Proceso de Poisson con tasa λ. Entonces cuando $h \searrow 0$*

$$\mathbb{P}(X_{t+h} - X_t \geq 2) = \mathrm{o}(h),$$

uniformemente en t.

Prueba. Como $(X_t)_{t\geq 0}$ es un Proceso de Poisson con tasa λ valen (1.2.1) y (1.2.2) y, por lo tanto

$$\mathbb{P}(X_{t+h} - X_t \geq 2) = 1 - \mathbb{P}(X_{t+h} - X_t = 0) - \mathbb{P}(X_{t+h} - X_t = 1) = \mathrm{o}(h)$$

■

Lema 1.2.4 *Sea $(N^1(t))_{t\geq 0}$ un Proceso de Poisson de parámetro 1. Definimos para $t \geq 0$ y $V \in \mathbb{N}$, $N^V(t) = N^1(Vt)$. Entonces, $(N^V(t))_{t\geq 0}$ es un Proceso de Poisson de parámetro V y, para t fijo, $\frac{N^V(t)}{Vt} \to 1$ c.t.p cuando $V \to \infty$.*

Prueba. Tenemos

$$\begin{aligned}
\frac{N^V(t)}{Vt} &= \frac{N^1(Vt)}{Vt} \\
&= \sum_{i=0}^{[Vt]} \frac{N^1(i+1) - N^1(i)}{Vt} + \frac{N^1(Vt) - N^1([Vt])}{Vt} \\
&= \frac{[Vt]}{Vt} \sum_{i=0}^{[Vt]} \frac{N^1(i+1) - N^1(i)}{[Vt]} + \frac{N^1(Vt) - N^1([Vt])}{Vt}
\end{aligned}$$

El primer término tiende a uno c.t.p por la ley de los grandes números y porque $\frac{[Vt]}{Vt}$ tiende a uno cuando $V \to \infty$. Veamos que el segundo término tiende a cero c.t.p. Sea $\varepsilon > 0$ y $k \in \mathbb{N}$ tal que $k\varepsilon t \geq 1$. Entonces

$$\begin{aligned}
\mathbb{P}\left(\frac{N^1(Vt) - N^1([Vt])}{Vt} > \varepsilon\right) &\leq \mathbb{P}\left(\frac{N^1([Vt]+1) - N^1([Vt])}{Vt} > \varepsilon\right) \\
&= \mathbb{P}\left(N^1([Vt]+1) - N^1([Vt]) > \varepsilon Vt\right) \\
&= \mathbb{P}\left(N^1([Vt]+1) - N^1([Vt]) > [\varepsilon Vt] + 1\right).
\end{aligned}$$

Para cada V, $N^1([Vt]+1) - N^1([Vt])$ es una variable Poisson de parámetro 1. Si tomamos X con esa distribución sabemos que $E(X) = \sum_{n=1}^{\infty} \mathbb{P}(X \geq n)$ y, por lo tanto,

$$\sum_{V=1}^{\infty} \mathbb{P}\left(N^1([Vt]+1) - N^1([Vt]) \geq [\varepsilon Vt] + 1\right)$$

$$\leq k \sum_{n=1}^{\infty} \mathbb{P}\left(N^1([Vt]+1) - N^1([Vt]) \geq n\right) = kE(X),$$

ya que como mucho hay k valores de V para los cuales $[\varepsilon Vt]+1 \in (n, n+1]$.

Entonces $\left(\mathbb{P}(\frac{N^1(Vt)-N^1([Vt])}{Vt} > \varepsilon)\right)_{V\in\mathbb{N}}$ es sumable y, por lo tanto $\frac{N^1(Vt)-N^1([Vt])}{Vt}$ tiende a cero c.t.p.

■

Para más información acerca de Cadenas de Markov a tiempo continuo y procesos de Poisson ver [5, 6, 7]. En [7] se puede ver también una definición de acoplamiento y varios ejemplos.

1.3. Desigualdad de Gronwall

En esta sección enunciamos y demostramos la desigualdad de Gronwall. Vamos a usar esta desigualdad en el capítulo que sigue.

Teorema 1.3.1 (Desigualdad de gronwall) *Sea $I = [0,T]$, H una función continua e integrable en I tal que existen A,B constantes para las cuales se cumple que $H(t) \leq A + \int_0^t BH(s)ds \;\forall\, t \in I$. Entonces, $H(t) \leq Ae^{BT}$.*

Prueba. Definimos para $s \in [0,T]$,

$$G(s) = e^{-Bs} \int_0^s BH(r)dr.$$

Tenemos entonces que

$$G'(s) = -Be^{-Bs} \int_0^s BH(r)dr + e^{-Bs}BH(s) = Be^{-Bs}(H(s) - \int_0^s BH(r)dr) \leq ABe^{-Bs}.$$

Como $G(s) = 0$, integrando esta desigualdad llegamos a que

$$G(t) \leq A(1 - e^{-Bt}).$$

Usando la definición de $G(t)$ y la hipótesis tenemos que

$$H(t) - A \leq e^{Bt}G(t) \leq e^{Bt}A(1 - e^{-Bt}),$$

de donde se deduce la desigualdad que queremos probar. ■

1.4. Método de Euler

Consideremos la ecuación diferencial

$$\begin{cases} \dot{x}(t) = F(x(t)), \\ \\ x(0) = y_0, \end{cases} \tag{1.4.1}$$

con $x : \mathbb{R} \to \mathbb{R}^d$ y $F : \mathbb{R}^d \to \mathbb{R}^d$, denotemos,

$$x(t) = (x^1(t), ..., x^d(t)), \quad F(x) = (F^1(x), ..., F^d(x)).$$

Observemos que si conocemos el valor de $x(t)$ para algún t entonces conocemos también el valor de la pendiente de la recta tangente a $x^i(t)$ en el punto $(t, x^i(t))$ para cada $i = 1, ..., d$. Como $x(0)$ es conocido sabemos exactamente el valor de $\dot{x}(0) = F(x(0))$. Esto sugiere que el valor de $x^i(h)$ va a ser aproximadamente

$$y_1^i = y_0^i + hF^i(y_0)$$

o, el valor de $x(h)$ va a ser aproximadamente

$$y_1 = y_0 + hF(y_0).$$

Si trabajamos en el intervalo $[0, T]$ y queremos dar una aproximación de $x(T)$, elegimos $V \in \mathbb{N}$ y puntos equiespaciados $t_1, ..., t_n$ con paso $h = \frac{T}{V}$, de manera tal que $t_k = kh$, $k = 0, ..., n$. Ahora dada y_k una aproximación de $x(t_k)$, definimos una aproximación de $x(t_{k+1}) = x(t_k + h)$ como

$$y_{i+1} = y_i + hF(y_i).$$

Definición 1.4.1 *El método de Euler con paso h para la ecuación* (1.4.1) *en el intervalo* $[0, T]$ *es la sucesión definida por* y_0, $y_1 = y_0 + hF(y_0)$ *e, inductivamente* $y_k = y_{k-1} + hF(y_{k-1})$.

1.4.1. Análisis del error y convergencia

El objetivo de esta sección es analizar cuál es el error que cometemos al utilizar el método de Euler para aproximar la solución de una ecuación de la forma (1.4.1). Para esto introducimos dos conceptos, el de error de truncamiento local y el error global. El primero es el error que cometemos al calcular $x(t + h)$ conociendo $x(t)$, el segundo es el error que acumulamos al aplicar V pasos del método, es decir $x(t_V) - y_V$.

Definición 1.4.2 *Llamamos π_k al error de truncamiento local que se comete en el k-ésimo paso. Este error está dado por la expresión*

$$x(t_k+h) = x(t_k) + hF(x(t_k)) + h\pi_k.$$

Es decir,

$$\pi_i = \frac{x(t_k+h)-x(t_k)}{h} - F(x(t_k)) = h\frac{\ddot{x}(\xi)}{2},$$

para algún $\xi \in (t_k, t_k+h)$.

Definición 1.4.3 *Dado $V \in \mathbb{N}$, llamamos ρ_V al error global que cometemos después de aplicar V pasos del método de Euler a una ecuación de la forma* (1.4.1)*. Es decir,*

$$\rho_V = x(t_V) - y_V$$

A continuación enunciamos un teorema que nos permite controlar el error global que cometemos al aproximar $x(T)$ con el método de Euler. Elegimos $V \in \mathbb{N}$ y $t_1, ..., t_V$ puntos equiespaciados, $t_i = hi$, $i = 1, ..., V$, con $h = \frac{T}{V}$ de manera tal que $t_V = T$.

Teorema 1.4.4 *Dado $V \in \mathbb{N}$ y $h = \frac{T}{V}$. Sean y_V el V-ésimo término del método de Euler con paso h y valor inicial y_0 para la ecuación* (1.4.1)*. Sea $x(t)$ la solución de dicha ecuación. Supongamos que F es una función Lipschitz con constante K, o sea, $|F(x)-F(y)| \leq K|x-y|$ con K independiente de x, y. Entonces, se tiene*

$$\max_{1\leq i\leq V} \{|x(t_i) - y_i|\} \leq (|x(0) - y_0| + KhT\|F\|_\infty)e^{KT}.$$

En particular si $x(0) = y_0$ entonces

$$\max_{1\leq i\leq V} \{|x(t_i) - y_i|\} \leq KhT\|F\|_\infty e^{KT}.$$

Prueba. Escribimos

$$x(t) = x(0) + \int_0^t F(x(s))ds.$$

Sea $\bar{y}$ la función que vale y_i en el intervalo $[ih, (i+1)h)$ para $i = 0, ..., V$. Definimos

$$y(t) = y_0 + \int_0^t F(\bar{y}(s))ds.$$

Sea $t_i \leq t < t_{i+1}$ entonces

$$\begin{aligned} y(t) &= y_0 + \sum_{k=0}^{k} F(y_k)[(k+1)h - kh] + (t-ih)F(y_i) \\ &= y_0 + \sum_{k=0}^{i-1} F(y_k)h + (t-ih)F(y_i) \\ &= y_i + (t-ih)F(y_i). \end{aligned}$$

En particular $y(t_i) = y_0 + \sum_{k=0}^{i-1} F(y_k)h = y_i$, e $y(t)$ es una interpolación lineal de los primeros V términos del método de Euler con paso h. Además, si $t_i \leq t < t_{i+1}$ entonces

$$\begin{aligned} |y(t) - \overline{y}(t)| &= |y_i + (t-ih)F(y_i) - y_i| \\ &= |(t-ih)F(y_i)| \\ &\leq h|F(y_i)|. \end{aligned}$$

Comparemos $y(t)$ con $x(t)$

$$\begin{aligned} |x(t) - y(t)| &= |x(0) - y_0 + \int_0^t F(x(s)) - F(\overline{y}(s))ds| \\ &\leq |x(0) - y_0| + \int_0^t |F(x(s)) - F(\overline{y}(s))|ds \\ &\leq |x(0) - y_0| + \int_0^t K|x(s) - \overline{y}(s)|ds \\ &\leq |x(0) - y_0| + \int_0^t K|x(s) - y(s)|ds + \int_0^t K|y(s) - \overline{y}(s)|ds \\ &\leq |x(0) - y_0| + K\int_0^t |x(s) - y(s)|ds + KhT\|F\|_\infty. \end{aligned}$$

Aplicando la desigualdad de Gronwall tenemos que

$$|x(t) - y(t)| \leq (|x(0) - y_0| + KhT\|F\|_\infty)e^{Kt}.$$

En particular

$$\max_{1 \leq i \leq V} \{|x(t_i) - y_i|\} \leq (|x(0) - y_0| + KhT\|F\|_\infty)e^{KT}.$$

■

Corolario 1.4.5 *Dado $V \in \mathbb{N}$, sean y_i^V el i-ésimo término del método de Euler con paso $\frac{T}{V}$ y valor inicial y_0^V para la ecuación* (1.4.1)*. Supongamos que F es una función acotada, Lipschitz con constante K y que $y_0^V \to x(0)$ cuando $V \to \infty$. Entonces*

$$\max_{1\leq i\leq V}\{|x(t_i)-y_i|\}\to 0 \quad \textit{cuando} \quad V\to\infty.$$

Prueba. Por el Teorema 1.4.4 tenemos que

$$\max_{1\leq i\leq V}\{|x(t_i)-y_i|\}\leq (|x(0)-y_0^V|+K\frac{T}{V}T\|F\|_\infty)e^{KT},$$

de donde se deduce el corolario. ■

Capítulo 2

Familias que dependen de una densidad

En este capítulo vamos a definir el concepto de familias de cadenas de Markov que depende de una densidad. Para eso, vamos a presentar distintos ejemplos en donde, mediante este concepto, podamos relacionar ecuaciones diferenciales ordinarias con cadenas de Markov. Este análisis esta acompanãdo por simulaciones que comparan los modelos determinísticos con los estocásticos.

2.1. Modelos poblacionales

Consideremos la ecuación diferencial

$$\begin{cases} \dot{x}_1(t) = ax_1(t) - bx_1(t)x_2(t), \\ \\ \dot{x}_2(t) = cx_1(t)x_2(t) - dx_2(t). \end{cases} \tag{2.1.1}$$

Con a,b,c y d parámetros positivos. Esta ecuación sirve de modelo para una situación de predador-presa, por ejemplo, zorros y conejos. Podemos interpretar $x(t) = (x_1(t), x_2(t))$ como la densidad de la población de cada especie respectivamente en una región infinita.

Veamos cómo puede pensarse esta ecuación,

- Asumimos que ante la ausencia de zorros la tasa de crecimiento de la densidad de la población de conejos es proporcional a $x_1(t)$. Asumimos también que la tasa con la cual los zorros comen a los conejos es proporcional a

$x_1(t)x_2(t)$. De ahí que la tasa de cambio en la densidad de la población de conejos viene dada por $\dot{x}_1(t) = ax_1(t) - bx_1(t)x_2(t)$.

- Asumimos que ante la ausencia de conejos la tasa con la cual cambia la densidad de la población de zorros es proporcional a $x_2(t)$ con constante negativa. Asumimos también que la tasa con la cual crece la densidad de población de zorros ante la presencia de conejos es proporcional a $x_1(t)x_2(t)$. De ahí que la tasa de cambio en la densidad de la población de zorros viene dada por $\dot{x}_2(t) = cx_1(t)x_2(t) - dx_2(t)$.

Podemos pensar que $a = \lambda_1 - \mu_1$ donde λ_1 es la tasa con la cual nacen conejos ante la ausencia de zorros y μ_1 la tasa con la cual mueren. Podemos pensar también que $d = \mu_2 - \lambda_2$ donde λ_2 es la tasa con la cual nacen los zorros ante la ausencia de conejos y μ_2 la tasa con la cual mueren.

Definimos una cadena de Markov que sirva de modelo para la misma situación de predador-presa en una superficie de tamañ̃o V. O sea,

$$X_t^V = (X_t^{V,1}, X_t^{V,2})$$

con $X_t^{V,1}$ la cantidad de conejos a tiempo t y $X_t^{V,2}$ la cantidad de zorros a tiempo t. Sean $k = (k_1, k_2)$ con $k_1, k_2 \in \mathbb{N}_0$, e_1, e_2 los vectores canónicos de $\mathbb{R}^2$ y $q(k, k \pm e_i)$ con $i = 1, 2$ las tasas de la cadena de Markov.

- Como cada conejo tiene crias con tasa λ_1, la tasa con la que la cantidad de conejos crece en uno es $q(k, k + e_1) = \lambda_1 k_1$.

- La tasa de muerte de los conejos viene separada en dos partes. Ante la ausencia de zorros, cada conejo muere con tasa μ_1. Tenemos entonces que, la tasa con la que la cantidad de conejos decrece en uno es $\mu_1 k_1$. Por otro lado, ante la presencia de zorros asumimos que cada vez que un zorro se encuentra con un conejo decrece en uno la cantidad de conejos. Teniendo en cuenta el tamaño de la superficie en la cual viven las poblaciones, tenemos que la tasa de encuentro es $\frac{b}{V}$. Luego, como la cantidad de encuentros posibles es $k_1 k_2$ la tasa con la cual decrece en uno la cantidad de conejos es $\frac{bk_1k_2}{V}$. Juntando estos dos casos tenemos que $q(k, k - e_1) = \mu_1 k_1 + b\frac{k_1k_2}{V}$.

- La tasa de crecimiento de los zorros también va a estar separada en dos partes. Ante la ausencia de conejos, cada zorro nace con tasa λ_2. Tenemos entonces que la tasa con la que la cantidad de zorros crece es $\lambda_2 k_2$. Por otro lado, ante la presencia de conejos, la tasa con la que la cantidad de zorros

crece en uno es $c\frac{k_1k_2}{V}$. Juntando estos dos casos tenemos que $q(k,k+e_2) = \lambda_2 k_2 + c\frac{k_1k_2}{V}$

- Por último, cada zorro muere a tasa μ_2. Por lo tanto, la tasa con la que la cantidad de zorros decrece en uno es $q(k,k-e_2) = \mu_2 k_2$

Resumiendo, las tasas vienen dadas por

$$\begin{array}{ll} q(k,k+e_1) = \lambda_1 k_1 & \text{nacimiento de un conejo,} \\ q(k,k-e_1) = \mu_1 k_1 + b\frac{k_1k_2}{V} & \text{muerte de un conejo,} \\ q(k,k+e_2) = \lambda_2 k_2 + c\frac{k_1k_2}{V} & \text{nacimiento de un zorro,} \\ q(k,k-e_2) = \mu_2 k_2 & \text{muerte de un zorro.} \end{array}$$

Vamos a pensarlo de la siguiente manera. Sea $f : \mathbb{R}^2 \times \mathbb{Z}^2 \to \mathbb{R}$ definida por

$$\begin{array}{l} f(x,e_1) = \lambda_1 x_1, \\ f(x,-e_1) = \mu_1 x_1 + b x_1 x_2, \\ f(x,e_2) = \lambda_2 x_2 + c x_1 x_2, \\ f(x,-e_1) = \mu_2 x_2, \\ f(x,e) = 0 \quad \text{para todo } e \neq \pm e_1, \pm e_2. \end{array}$$

Entonces

$$\begin{array}{l} q(k,k+e_1) = V f(\frac{k}{V}, e_1), \\ q(k,k-e_1) = V f(\frac{k}{V}, -e_1), \\ q(k,k+e_2) = V f(\frac{k}{V}, e_2), \\ q(k,k-e_2) = V f(\frac{k}{V}, -e_2). \end{array}$$

Calculemos ahora el drift del proceso $\frac{1}{V}(X_1^V(t), X_2^V(t))$. Como los únicos saltos posibles vienen dados por $\frac{1}{V}L$, para $L = \{(1,0),(-1,0),(0,1),(0,-1)\}$ tenemos que

$$\begin{array}{rcl} \Sigma_{l\in L} \dfrac{l}{V} q(Vx, Vx+l) & = & \Sigma_{l\in L} l f(x,l) \\ & = & (1,0)\lambda_1 x_1 + (0,1)(\lambda_2 x_2 + c x_1 x_2) \\ & - & (1,0)(\mu_1 x_1 + b x_1 x_2) - (0,1)\mu_2 x_2 \\ & = & (\lambda_1 x_1 - (\mu_1 x_1 + b x_1 x_2), \lambda_2 x_2 + c x_1 x_2 - \mu_2 x_2) \\ & = & (a x_1 - b x_1 x_2, c x_1 x_2 - d x_2). \end{array}$$

Luego si definimos

$$F(x_1,x_2)=\sum_{l\in L} lf(x,l)=(ax_1-bx_1x_2,cx_1x_2-dx_2),$$

entonces $F(x)$ es el drift del proceso $\frac{1}{V}(X_1^V(t),X_2^V(t))$ en el punto (x_1,x_2) y además la ecuación original se puede escribir como

$$(\dot{x_1}(t),\dot{x_2}(t))=F(x_1(t),x_2(t)).$$

El drift del proceso de una cadena de Markov nos dice hacia donde va el proceso en media (condicional al pasado) y la derivada de la solución de la ecuación nos dice hacia donde van las trayectorias. Es de esperarse que las trayectorias que son solución de la ecuación y las de los procesos se parezcan cuando V es grande.

Teniendo en cuenta que el supuesto de que las poblaciones pueden crecer indefinidamente es poco realista podemos agregarle un factor social a la ecuación diferencial. Este factor obliga a que cuando la población exceda cierto valor η, al cual llamamos población limite, la tasa de crecimiento se vuelva negativa. Esto es, ante la ausencia de interaccción entre las especies, la ecuación para los conejos es

$$\dot{x_1}(t)=a(\eta-x_1(t))x_1(t).$$

Para los zorros agregamos un factor proporcional a $x_2^2(t)$. En definitiva,

$$\begin{cases} \dot{x}_1(t)=ax_1(t)-bx_1(t)x_2(t)-\gamma x_1^2(t), \\ \\ \dot{x}_2(t)=cx_1(t)x_2(t)-dx_2(t)-\delta x_2^2(t). \end{cases} \tag{2.1.2}$$

Pensando de nuevo que $a=\lambda_1-\mu_1$ y $d=\mu_2-\lambda_2$. Definimos un proceso $X_t^V=(X_1^V(t),X_2^V(t))$ que modele el comportamiento de esta situación con las siguientes tasas

$q(k,k+e_1)=\lambda_1 k_1$	nacimiento de un conejo,
$q(k,k-e_1)=\mu_1 k_1+b\frac{k_1k_2}{V}+\gamma\frac{k_1^2}{V}$	muerte de un conejo,
$q(k,k+e_2)=\lambda_2 k_2+c\frac{k_1k_2}{V}$	nacimiento de un zorro,
$q(k,k-e_2)=\mu_2 k_2+\delta\frac{k_2^2}{V}$	muerte de un zorro.

Es razonable dividir por V a los términos que tienen que ver con el factor social ya que estamos teniendo en cuenta el tamanõ de la superficie en donde las

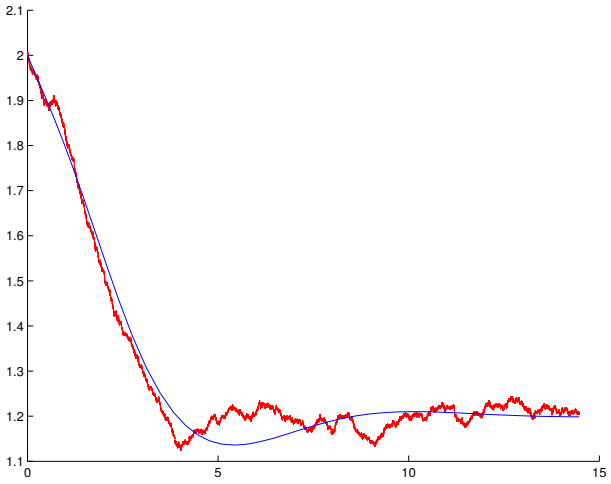

Figura 2.1: En azul la proporción de conejos del modelo determinístico dado por la ecuación (2.1.2) y en rojo la proporción de conejos para el modelo estocástico con $V = 5000$. En ambos casos el punto inicial es 2 y los parámetros son $c = b = \mu_1 = \lambda_2 = 1$, $\mu_2 = \lambda_1 = 2$ y $\lambda = \mu = 0{,}5$.

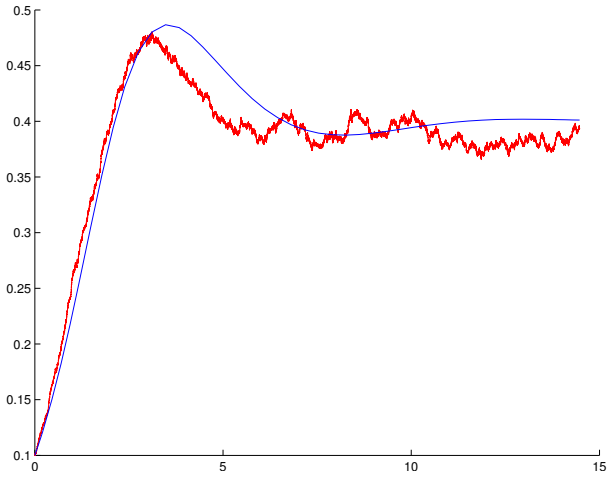

Figura 2.2: En azul la proporción de zorros del modelo determinístico dado por la ecuación (2.1.2) y en rojo la proporción de zorros para el modelo estocástico con $V = 5000$. En ambos casos el punto inicial es 0,1 y los parámetros son $c = b = \mu_1 = \lambda_2 = 1$, $\mu_2 = \lambda_1 = 2$ y $\gamma = \delta = 0{,}5$.

poblaciones viven. Nuevamente, podemos pensar que las tasas vienen dadas por $q(k,k+l) = Vf(\frac{k}{V},l)$ para la siguiente f

$$\begin{aligned}
&f(x,e_1) = \lambda_1 x_1,\\
&f(x,-e_1) = \mu_1 x_1 + bx_1x_2 + \gamma x_1^2,\\
&f(x,e_2) = \lambda_2 x_2 + cx_1x_2,\\
&f(x,-e_2) = \mu_2 x_2 + \delta x_2^2,\\
&f(x,e) = 0 \quad \text{para todo } e \neq \pm e_1, \pm e_2.
\end{aligned}$$

Nuevamente,

$$F(x) = \sum_{l\in L} lf(x,l) = (ax_1 - bx_1x_2 - \gamma x_1^2, cx_1x_2 - dx_2 - \delta x_2^2),$$

es el drift del proceso $\frac{1}{V}(X_1^V(t), X_2^V(t))$ evaluado en x y también sirve para escribir la ecuación diferencial como $\dot{x}(t) = F(x(t))$. Podemos ver en las figuras 2.1 y 2.2 una comparación entre el modelo determinístico y el estocástico. A este tipo de familias de cadenas de Markov se las llama familias que dependen de una densidad (ver [4]).

Definición 2.1.1 *Dada una familia de cadenas de Markov $(X_t^V)_{V\in\mathbb{N}}$, con espacio de estados en $\mathbb{Z}^d$. Decimos que $(X_t^V)_{V\in\mathbb{N}}$ es una familia de cadenas de Markov que depende de una densidad si existe una función continua $f : \mathbb{R}^d \times \mathbb{Z}^d \to \mathbb{R}$ tal que las tasas de X_t^V vienen dadas por*

$$q(k,k+l) = Vf(\frac{1}{V}k,l), \text{ para } l \neq 0 \text{ y } k \in \mathbb{Z}^d,$$

para todo $V \in \mathbb{N}$. En ese caso definimos

$$L = \{l \in \mathbb{Z}^d : \text{ existe } x \in \mathbb{R}^d \text{ con } f(x,l) \neq 0\} \text{ y}$$

$$F(x) = \sum_{l\in L} lf(x,l).$$

Veamos otros ejemplos de aplicaciones que caen dentro de este tipo de familias de cadenas de Markov. De aquí en adelante vamos a definir las densidades f solo en los puntos en los que no valgan cero. Lo mismo hacemos para las tasas de las Cadenas de Markov.

2.2. Modelos epidémicos

Ante la presencia de una enfermedad contagiosa definimos tres tipos de individuos: susceptibles, infectados y recuperados o inmunes. Los primeros son los que pueden contagiarse, los segundos los que tienen la enfermedad y pueden contagiarla, y los terceros son los que se curaron y no pueden volver a contagiarse. Teniendo esto en cuenta hay dos grandes tipos de modelos: el modelo SIS y el modelo SIR. Para mas información sobre modelos epidémicos ver [1].

2.2.1. Modelo SIS

En este modelo un individuo susceptible una vez que se infecta no se vuelve inmune. O sea, cuando se recupera vuelve al grupo de los susceptibles. Este modelo fue usado para el estudio de enfermedades de transmisión sexual. Para simplificar asumimos también que todos los individuos nacen susceptibles (nadie puede nacer con la enfermedad) y que no hay muertes causadas por la enfermedad. La ecuación diferencial que sirve para este modelo viene dada por

$$\begin{cases} \dot{S}(t) = -\beta S(t)I(t) + (b+\gamma)I(t), \\ \\ \dot{I}(t) = \beta S(t)I(t) - (b+\gamma)I(t), \end{cases} \tag{2.2.1}$$

donde $S(t), I(t)$ representa la proporción de susceptibles e infectados respectivamente, $\beta > 0$ es la tasa de contacto, $\gamma > 0$ es la tasa con la cual se recuperan los individuos y $b \geq 0$ es la tasa de nacimiento. Asumimos que la tasa de muerte es igual a la de nacimiento, con lo cual, el sistema es conservativo (notar que $\dot{S}(t) + \dot{I}(t) = 0$).

Definimos una familia de cadenas de Markov (S_t^V, I_t^V) que modele esta misma situación. Aquí S_t^V es la cantidad de individuos susceptibles a tiempo t e I_t^V es la cantidad de individuos infectados a tiempo t. Las tasas correspondientes son, para $k \in \mathbb{N}^2$

$$\begin{array}{ll} q(k, k+e_1-e_2) = (b+\gamma)k_2 & \text{se muere o se recupera un infectado,} \\ q(k, k+e_2-e_1) = \beta \frac{k_1 k_2}{V} & \text{se infecta un susceptible.} \end{array}$$

Donde V es el tamaño total de la población. Como asumimos que las tasas de nacimiento y muerte son las mismas, cuando muere un individuo infectado nace uno susceptible. Al estudiar estas Cadenas de Markov estamos estudiando

una versión microscópica del fenómeno que modela la ecuación (2.2.1), ya que en la ecuación estamos estudiando la proporción de individuos infectados mientras que en la cadena de Markov estamos estudiando solo la cantidad de individuos infectados en una parte de la población de tamaño V.

En este caso podemos definir, para $x \in \mathbb{R}^2$

$$\begin{aligned} f(x, e_1 - e_2) &= (b+\gamma)x_2, \\ f(x, e_2 - e_1) &= \beta x_1 x_2. \end{aligned}$$

Así, vale que $q(k, k+l) = V f(\frac{k}{V}, l)$. Además,

$$F(x) = \sum_{l \in L} l f(x,l) = (-\beta x_1 x_2 + (b+\gamma)x_2, \beta x_1 x_2 - (b+\gamma)x_2)$$

es el drift del proceso $\frac{1}{V}(S_t^V, I_t^V)$ evaluado en x. Se verifica también que

$$(\dot{S}(t), \dot{I}(t)) = F(S(t), I(t)).$$

2.2.2. Modelo SIR

En este modelo una vez que un individuo infectado se recupera, se vuelve inmune. Es decir, una vez que se recupera pasa a la clase R de los individuos que no están infectados pero no son susceptibles. Este modelo se aplicó para enfermedades infantiles como la varicela, el sarampión y las paperas.

La ecuación diferencial que sirve para este modelo viene dada por

$$\begin{cases} \dot{S}(t) = -\beta S(t) I(t) + b(I(t) + R(t)), \\ \dot{I}(t) = \beta S(t) I(t) - (b+\gamma) I(t), \\ \dot{R}(t) = \gamma I(t) - b R(t). \end{cases} \tag{2.2.2}$$

Donde $S(t), I(t), R(t)$ representa la proporción de individuos susceptibles, infectados y recuperados respectivamente. El parámetro $\beta > 0$ es la tasa de contacto, $\gamma > 0$ es la tasa con la cual se recuperan los individuos y $b \geq 0$ es la tasa de nacimiento. Asumimos de nuevo que la tasa de muerte es la misma que la de nacimiento con lo cual el sistema es conservativo (notar que $\dot{S}(t) + \dot{I}(t) + \dot{R}(t) = 0$). Entonces, si $S(0) + I(0) + R(0) = 1$ tenemos que $R(t) = 1 - S(t) - I(t)$ y podemos

escribir el sistema (2.2.2) como

$$\begin{cases} \dot{S}(t) = -\beta S(t)I(t) + b(1-S(t)) \\ \\ \dot{I}(t) = \beta S(t)I(t) - (b+\gamma)I(t) \end{cases} \tag{2.2.3}$$

Definimos una familia de cadenas de Markov (S_t^V, I_t^V) que modele esta misma situación. Las tasas correspondientes son, para $k \in \mathbb{N}^2$

$q(k, k+e_2-e_1) = \beta \frac{k_1 k_2}{V}$	Se infecta un susceptible,
$q(k, k-e_2) = \gamma k_2$	Se recupera un infectado,
$q(k, k+e_1-e_2) = bk_2$	Se muere un infectado,
$q(k, k+e_1) = b(V-k_1-k_2)$	Se muere un recuperado.

Donde V es el tamaño total de la población, y como asumimos que las tasas de nacimiento y muerte son iguales cuando muere un individuo también nace uno. Los individuos nacen siempre susceptibles.

Para $x \in \mathbb{R}^2$, definimos

$$\begin{aligned} f(x, e_2-e_1) &= \beta x_1 x_2, \\ f(x, -e_2) &= \gamma x_2, \\ f(x, e_1-e_2) &= b x_2, \\ f(x, e_1) &= b(1-x_1-x_2). \end{aligned}$$

Así, vale que $q(k, k+l) = V f(\frac{k}{V}, l)$. Además,

$$F(x) = \sum_{l \in L} l f(x,l) = (-\beta x_1 x_2 + b(1-x_1), \beta x_1 x_2 - (b+\gamma)x_2),$$

es el drift del proceso $\frac{1}{V}(S_t^V, I_t^V)$. Se verifica también que $(\dot{S}(t), \dot{I}(t)) = F(S(t), I(t))$. Podemos ver en las figuras 2.3 y 2.4 la similitud entre el modelo determinístico y el estocástico.

2.3. Teoría de Colas

En estos modelos se estudia una cola de clientes. Hay una determinada cantidad de servidores que van atendiendo a estos clientes. El objetivo de estos modelos es determinar la cantidad de clientes que va a haber en la cola después de determinado tiempo. La teoría de colas se aplica, por ejemplo, para la distribución de la senãl de internet. Para leer más sobre este tema ver [3].

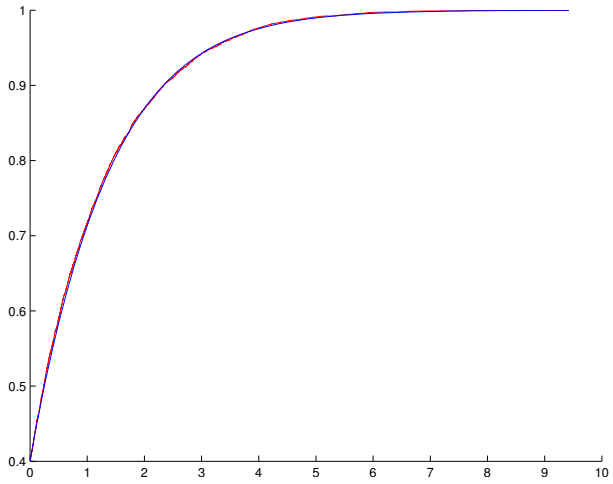

Figura 2.3: En azul la proporción de individuos susceptibles del modelo determinístico dado por la ecuación (2.2.2) y en rojo la proporción de individuos susceptibles para el modelo estocástico con $V = 10000$. En ambos casos el punto inicial es 0,4 y los parámetros son $b = \beta = \gamma = 1$.

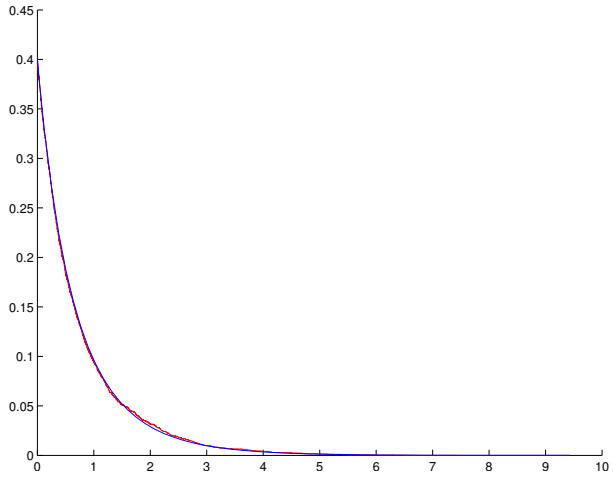

Figura 2.4: En azul la proporción de individuos infectados del modelo determinístico dado por la ecuación (2.2.2) y en rojo la proporción de individuos infectados para el modelo estocástico con $V = 10000$. En ambos casos el punto inicial es 0,4 y los parámetros son $b = \beta = \gamma = 1$.

2.3.1. Cola M/M/1

Consideremos una cola en la que los clientes llegan con tasa λ y son atendidos a tasa μ por un servidor. Llamamos X_t a la cantidad de clientes que hay en la cola a tiempo t. En estas colas el servidor atiende por orden de llegada. Si asumimos que las llegadas y los servicios tienen distribución exponencial entonces $(X_t)_{t\geq 0}$ es Markov con tasas

$$\begin{array}{ll} q(k,k+1)=\lambda & \text{Llega un cliente a la cola,} \\ q(k,k-1)=\mu \mathbf{1}\{k>0\} & \text{El servidor atiende a un cliente,} \end{array} \tag{2.3.1}$$

para $k \in \mathbb{N}$.

En este caso, tenemos que, para $x \in \mathbb{R}_{\geq 0}$

$$\begin{array}{ll} f(x,1)=\lambda, & \\ f(x,-1)=\mu & \text{para } x>0, \\ f(0,-1)=0. & \end{array}$$

Pero el problema es que f no es continua. Lo que podemos hacer es definir, para $x \in \mathbb{R}$

$$\begin{array}{l} f(x,1)=\lambda, \\ f(x,-1)=\mu. \end{array}$$

Entonces tenemos que $F(x)=\lambda-\mu$, con lo cual tenemos la ecuación

$$\dot{x}(t)=\lambda-\mu$$

que tiene como solución a

$$x(t)=(\lambda-\mu)t+x_0.$$

Como la cantidad de personas en la cola es siempre mayor o igual que cero tomamos como trayectoria límite a la parte positiva de $x(t)$. En ese caso las cadenas de Markov X_t^V tienen tasas

$$\begin{array}{ll} q(k,k+1)=V\lambda, & \text{Llega un cliente a la cola,} \\ q(k,k-1)=V\mu. & \text{el servidor atiende a un cliente.} \end{array} \tag{2.3.2}$$

Así, el drift de $\frac{1}{V}X_t^V$ evaluado en un x de la forma $\frac{k}{V}$ es $\frac{1}{V}(V\lambda-V\mu)=\lambda-\mu$.

En definitiva, una cola M/M/1 es una cadena de Markov con tasas dadas por (2.3.1). El M en el nombre de la cola viene de memory loss, que en inglés significa perdida de memoria. Esto se debe a que la distribución de llegada y atención de los clientes es exponencial. El uno es porque tenemos un solo servidor. Lo que estamos haciendo en (2.3.2) es reescalar el proceso en el tiempo y el espacio. Al multiplicar las tasas por V aceleramos los tiempos y cuando estudiamos $\frac{1}{V}X_t^V$ estamos estudiando, a nivel macroscópico, el comportamiento de X_t^V. Esto es útil justamente para poder estudiar la ecuación limite.

En este caso, si bien las cadenas de Markov que tomamos no respetan el esquema de la cola M/M/1 sabemos que hasta que los procesos se vuelvan negativos las dinámicas son las mismas.

2.3.2. Cola M/M/∞

En esta cola tenemos una cantidad infinita de servidores y cada cliente que llega a la cola es atendido inmediatamente. Asumimos que las llegadas tienen distribución exponencial de parámetro λ y lo que tarda cada servidor en atender a un cliente tiene distribución exponencial de parámetro μ. Las tasas del proceso son, para $k \in \mathbb{N}$

$$\begin{array}{ll} q(k,k+1) = \lambda & \text{Llega un cliente a la cola,} \\ q(k,k-1) = \mu k & \text{algun servidor atiende a un cliente.} \end{array} \tag{2.3.3}$$

En este caso, tenemos que, para $x \in \mathbb{R}_{\geq 0}$

$$\begin{aligned} f(x,1) &= \lambda, \\ f(x,-1) &= \mu x. \end{aligned}$$

Entonces tenemos que

$$F(x) = \lambda - \mu x$$

y la ecuación $\dot{x}(t) = \lambda - \mu x(t)$. Esta ecuación tiene como solución a

$$x(t) = \frac{\lambda}{\mu} + (x_0 - \frac{\lambda}{\mu})e^{-\mu t}.$$

Las cadenas de Markov X_t^V tienen tasas

$$\begin{array}{ll} q(k,k+1) = V\lambda, & \text{Llega un cliente a la cola,} \\ q(k,k-1) = \mu k. & \text{Un servidor termina de atender a un cliente.} \end{array} \tag{2.3.4}$$

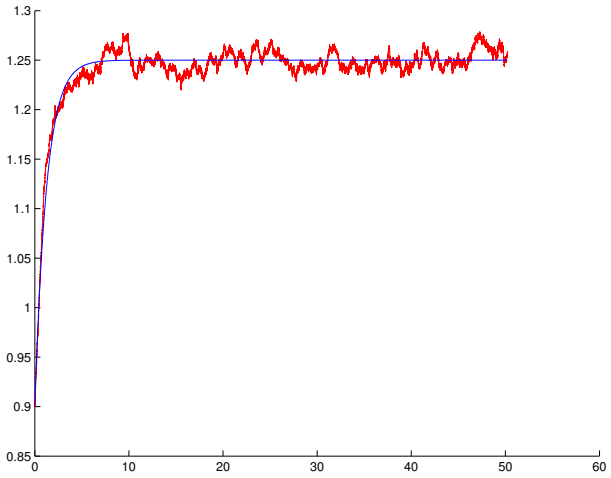

Figura 2.5: En azul la proporción de individuos en la cola dada por el modelo determinístico dado con ecuación 2.3.2 y en rojo la proporción de individuos en la cola dada por el modelo estocástico con $V = 10000$. En ambos casos el punto inicial es 0,9 y los parámetros son $\lambda = 1$ y $\mu = 0$.

Así, el drift de $\frac{1}{V}X_t^V$ evaluado en un x de la forma $\frac{k}{V}$ es $\frac{1}{V}(V\lambda - \mu k) = \lambda - \mu x$.

En definitiva una cola M/M/∞ es una cadena de Markov con tasas dadas por (2.3.3). Nuevamente, lo que estamos haciendo en (2.3.4) es reescalar el proceso en el tiempo y el espacio para poder estudiar la ecuación límite.

2.4. Algunos resultados conocidos

Thomas G. Kurtz probó en una publicación de 1970 intitulada "Ordinary differential equations as limits of pure jump Markov processes" [4] que las familias de cadenas de Markov que dependen de una densidad tienden en probabilidad a las trayectorias que son solución de la ecuación $\dot{x}(t) = F(x(t))$ cuando la F, definida como antes, es Lipschitz. El objetivo de esta tesis es probar un resultado similar al de Kurtz pero usando otra técnica basada en acoplamientos.

Capítulo 3

Convergencia

Sea $(X_t^V)_{V\in\mathbb{N}}$ una familia de cadenas de Markov que depende de una densidad con densidad f y la ecuación diferencial correspondiente $\dot{x}(t) = F(x(t))$. Con f, F y L como en la Definición 2.1.1. Probamos en este capítulo que

$$E(|X_t^V - x(t)|) \to 0 \quad \text{cuando } V \to \infty.$$

Este es el principal resultado de esta tesis y lo enunciamos en el Teorema 3.0.3.

Notación 3.0.1 *Dado $l \in \mathbb{Z}^d$ notamos con l^j a la j-ésima coordenada de l.*

Observación 3.0.2 *Observemos que* $\text{máx}_{l\in L}|l^j| < \infty$ *si y solo si L es un subconjunto finito de $\mathbb{Z}^d$.*

Teorema 3.0.3 *Sea $(\widetilde{X}_t^V)_{V\in\mathbb{N}}$ una familia de cadenas de Markov que depende de una densidad f y $X_t^V = \frac{1}{V}\widetilde{X}_t^V$. Sea $U \subseteq \mathbb{R}^d$ acotado tal que $X_t^V \in U$ para todo $V \in \mathbb{N}$ y $t \geq 0$. Supongamos también que $\bar{q} = \sup_{x\in U}\{\sum_{l\in L} f(x,l)\} < \infty$,* $\text{máx}_{l\in L}|l^j| < \infty$*. Sea $x(t)$ la solución de la ecuación $\dot{x}(t) = F(x(t))$ con valor inicial x_0 y F una función localmente Lipchitz relacionada con f como en la Definición 2.1.1. Supongamos que $x(t) \in U$ para todo $t \geq 0$ y que $X_0^V \to x_0$ cuando $V \to \infty$. Entonces para cada $t \geq 0$ tenemos que*

$$E(|X_t^V - x(t)|) \to 0 \quad \textit{cuando } V \to \infty.$$

3.1. Idea de la prueba

Tenemos una familia de cadenas de Markov que depende de una densidad $\widetilde{X}_t^V$, es decir, existe una f, a la cual llamamos densidad, tal que si $\widetilde{X}_t^V$ está en el lugar

x, entonces espera un tiempo exponencial de parámetro $q(x) = \sum_{l\in L} V f(\frac{x}{V}, l)$ y salta al lugar $x+l$ con probabilidad $\frac{q(x,x+l)}{q(x)} = V\frac{f(\frac{x}{V},l)}{q(x)}$. Definimos también $F(x) = \sum_{l\in L} l f(x,l)$ y tenemos entonces la ecuación diferencial $\dot{x}(t) = F(x(t))$.

A lo largo de esta sección vamos a suponer que $\widetilde{X}_t^V$ cumple las hipótesis del Teorema 3.0.3. Vamos a trabajar con $t \in [0,T]$ para $T \geq 0$ arbitrario, y de esta manera vamos a obtener resultados para cualquier $t \geq 0$ fijo.

Tomamos $q_1 > \bar{q}$ tal que $q = \frac{q_1 V}{T} \geq \bar{q}V$ para todo $V \in \mathbb{N}$ y definimos la siguiente familia de procesos, $(\widetilde{Y}_t^V)_{V\in\mathbb{N}}$

- Si $\widetilde{Y}_t^V$ está en el lugar y, espera un tiempo exponencial de parámetro $\frac{q_1 V}{T}$ y salta al lugar

$$\sum_{l\in L}(y+l)V\frac{f(\frac{y}{V},l)}{\frac{V}{q_1 T}} + y(1 - \sum_{l\in L} V\frac{f(\frac{y}{V},l)}{\frac{q_1 V}{T}}) = y + \frac{T}{q_1}F(\frac{y}{V}).$$

  Lo que hace $\widetilde{Y}_t^V$ en definitiva es saltar a la esperanza de donde saltaría $\widetilde{X}_t^V$ si estuviese en y.

Definimos ahora $X_t^V = \frac{\widetilde{X}_t^V}{V}$ e $Y_t^V = \frac{\widetilde{Y}_t^V}{V}$. Tenemos entonces que

- Si X_t^V está en el lugar x, espera un tiempo exponencial de parámetro $q(x)$, y salta al lugar $x + \frac{l}{V}$ con probabilidad $V\frac{f(x,l)}{q(x)}$.

- Si Y_t^V está en el lugar y espera un tiempo exponencial de parámetro $\frac{q_1 V}{T}$ y salta al lugar $y + \frac{1}{q}F(y)$.

Observación 3.1.1 *Por el Lema 1.1.9, usando que $q \geq q(x)$ para todo x podemos construir X_t^V de la siguiente forma,*

- *Si X_t^V está en el lugar x entonces espera un tiempo exponencial de parámetro q, salta al lugar $x + \frac{l}{V}$ con probabilidad $V\frac{f(x,l)}{q}$ y se queda en el lugar con probabilidad $1 - \sum_{l\in L} V\frac{f(x,l)}{q}$. Es decir, espera menos tiempo pero le damos la posibilidad de quedarse en el lugar y la distibución del tiempo de espera termina siendo la misma.*

Por último definimos una familia de procesos más

- Z_t^V avanza con la media de como avanza X_t^V. Esto es, si X_t^V está en el lugar x entonces en media avanza

$$\sum_{l\in L}(x+\frac{l}{V})V\frac{f(\frac{x}{V},l)}{q}+x(1-\sum_{l}V\frac{f(\frac{x}{V},l)}{q})-x=\frac{F(x)}{q}.$$

  Luego, si Z_t^V está en el lugar z espera un tiempo exponencial de parámetro q y , si en ese momento X_t^V esta en el lugar x, entonces Z_t^V salta al lugar $z+\frac{F(x)}{q}$.

La idea ahora es probar que X_t^V se parece a Z_t^V, Z_t^V se parece a Y_t^V y, como Y_t^V es en cierta forma el método de Euler para la ecuación $\dot{x}(t)=F(x(t))$, Y_t^V se parece a $x(t)$. Podemos usar después la desigualdad triangular para probar el Teorema 3.0.3. Antes de eso vamos a hacer una construcción gráfica de estos procesos.

3.2. Construcción gráfica de los procesos

En esta sección vamos a construir los tres procesos mencionados anteriormente para todo $V\geq 1$ utilizando un único Proceso de Poisson $N=(N(t))_{t\geq 0}$ de parámetro $\frac{q_1}{T}$ y una sucesión $(U_i)_{i\geq 1}$ de variables aleatorias independientes (entre sí y con N) con distribución uniforme en $[0,1]$.

Consideramos, para $i\in\mathbb{N}$, los tiempos de salto del Proceso de Poisson

$$\tau_i=\text{ínf}\{t\geq 0\colon N(t)\geq i\}.$$

Observemos que $N^V(t)=N(Vt)$ es un Proceso de Poisson de parámetro $\frac{Vq_1}{T}=q$ con tiempos de salto $(\frac{\tau_i}{V})_{i\in\mathbb{N}}$. Notar que por el Lema 1.2.4 $\frac{N^V(t)T}{q_1V}\to t$ cuando $V\to\infty$.

Teniendo en cuenta la Observación 3.1.1, construimos la función $G\colon \mathbb{Z}^d\times[0,1]\to\mathbb{Z}^d$ dada por

$$G(k,u)=k+l,\quad \text{si } u\in I(k,l),$$

donde $(I(k,l))_l$ es una familia de intervalos disjuntos, tal que la medida de Lebesgue de $I(k,l)$ es $\frac{q(k,k+l)}{q}$. Si $u\notin\cup_l I(k,l)$, definimos

$$G(k,u)=k.$$

Finalmente definimos $\widetilde{X}_t^V$ como

$$\widetilde{X}_t^V = V x_0^V, \quad \text{si } 0 \leq t < \frac{\tau_1}{V},$$

y para $\frac{\tau_i}{V} \leq t < \frac{\tau_{i+1}}{V}$,

$$\widetilde{X}_t^V = G(\widetilde{X}^V_{\frac{\tau_{i-1}}{V}}, U_i).$$

Por último, notamos $X_t^V = \frac{1}{V}\widetilde{X}_t^V$. El proceso Z_t^V viene dado por

$$Z_t^V = x_0^V, \quad \text{si } 0 \leq t < \frac{\tau_1}{V},$$

y para $\frac{\tau_i}{V} \leq t < \frac{\tau_{i+1}}{V}$,

$$Z_t^V = Z^V_{\frac{\tau_{i-1}}{V}} + \frac{1}{q} F\left(X^V_{\frac{\tau_{i-1}}{V}}\right).$$

Construimos también Y_t^V como

$$Y_t^V = x_0^V, \quad \text{si } 0 \leq t < \frac{\tau_1}{V},$$

y para $\frac{\tau_i}{V} \leq t < \frac{\tau_{i+1}}{V}$,

$$Y_t^V = Y^V_{\frac{\tau_{i-1}}{V}} + \frac{1}{q} F\left(Y^V_{\frac{\tau_{i-1}}{V}}\right).$$

De esta manera construimos X_t^V como en la Observación 3.1.1 y, por lo tanto, va a tener la distribución correcta.

3.3. Aproximación de Z_t^V a X_t^V

Sea $\mathscr{F}_t = \sigma((N^V(s))_{s\leq t})$. Entonces dado $\mathscr{F}_t$ conocemos los tiempos de salto τ_i tal que $\tau_i \leq t$. Llamamos $X^{V,j}_{\tau_i}$ y $F^j(X^V_{\tau_{i-1}})$ a las j-ésimas coordenadas de $X^V_{\tau_i}$ y $F(X^V_{\tau_{i-1}})$ respectivamente.

Definición 3.3.1 *Dados $n, j \in \mathbb{N}$ con $j \leq d$ definimos*

$$A^j(n) = X^{V,j}_{\tau_n} - X^{V,j}_{\tau_{n-1}} - \frac{F^j(X^V_{\tau_{n-1}})}{q}. \tag{3.3.1}$$

Lema 3.3.2 *Sean $i,h \in \mathbb{N}$ tales que $i \neq h$ entonces*

$$E\left(A^j(i)A^j(h)\mathbf{1}\{\tau_i,\tau_h \leq t\}\Big|\mathscr{F}_t\right) = 0$$

para todo $j \in \mathbb{N}$ con $j \leq d$.

Prueba. Podemos suponer $i < h$ y observemos que, dado $\tau_n \leq t$, si conocemos $(U_i)_{i\leq n}$ entonces conocemos también los valores de $(X_{\tau_i})_{i\leq n-1}$. Veamos que vale el resultado del lema

$$\begin{aligned}
&E\left(A^j(i)A^j(h)\mathbf{1}\{\tau_i,\tau_h \leq t\}\Big|\mathscr{F}_t\right)\\
&= E\left(E\left(\left(A^j(i)A^j(h)\mathbf{1}\{\tau_i,\tau_h \leq t\}\Big|\sigma(\mathscr{F}_t,(U_i)_{i\leq h})\right)\right)\Big|\mathscr{F}_t\right)\\
&= E\left(A^j(i)\mathbf{1}\{\tau_i,\tau_h \leq t\}E\left(X^{V,j}_{\tau_n} - X^{V,j}_{\tau_{n-1}} - \frac{F^j(X^V_{\tau_{n-1}})}{q}\Big|\sigma(\mathscr{F}_t,(U_i)_{i\leq h})\right)\Big|\mathscr{F}_t\right)\\
&= E\left(A^j(i)\mathbf{1}\{\tau_i,\tau_h \leq t\}\left[E\left(X^{V,j}_{\tau_h}\Big|\sigma(\mathscr{F}_t,(U_i)_{i\leq h})\right) - X^{V,j}_{\tau_{h-1}} - \frac{F^j(X^V_{\tau_{h-1}})}{q}\right]\Big|\mathscr{F}_t\right) = 0,
\end{aligned}$$

pues

$$\begin{aligned}
E\left(\left(X^{V,j}_{\tau_h}|\sigma(\mathscr{F}_t,(U_i)_{i\leq h}\right)\right) &= \sum_{l/l^j\neq 0}(X^{V,j}_{\tau_{h-1}} + \frac{l^j}{V})V\frac{f(X^{V,j}_{\tau_{h-1}},l)}{q}\\
&+ X^{V,j}_{\tau_{h-1}}(1 - \sum_{l/l^j\neq 0} V\frac{f(X^V_{\tau_{h-1}},l)}{q})\\
&= X^{V,j}_{\tau h-1} + \frac{1}{q}\sum_{l/l^j\neq 0} l^j f(X^V_{\tau_{h-1}},l)\\
&= X^{V,j}_{\tau h-1} + \frac{F^j(X^V_{\tau_{h-1}})}{q}.
\end{aligned}$$

■

Lema 3.3.3 *Supongamos que* $\max_{l\in L}|l^j| < \infty$ *entonces*

$$|X^{V,j}_{\tau_i} - X^{V,j}_{\tau_{i-1}} - \frac{F^j(X^V_{\tau_{i-1}})}{q}| \leq \frac{2}{V}\max_l |l^j|,$$

para todo $i \in \mathbb{N}$.

Prueba.

$$\begin{aligned}
\left|X^{V,j}_{\tau_i} - X^{V,j}_{\tau_{i-1}} - \frac{F^j(X^V_{\tau_{i-1}})}{q}\right| &= \left|X^{V,j}_{\tau_i} - X^{V,j}_{\tau_{i-1}} - \sum_{l/l^j\neq 0} l^j \frac{f(X^V_{\tau_{i-1},l})}{q}\right| \\
&\leq |X^{V,j}_{\tau_i} - X^{V,j}_{\tau_{i-1}}| + \sum_{l/l^j\neq 0} |l^j| \frac{f(X^V_{\tau_{i-1},l})}{q} \\
&= |X^{V,j}_{\tau_i} - X^{V,j}_{\tau_{i-1}}| + \frac{1}{V}\sum_{l/l^j\neq 0} |l^j| V \frac{f(X^V_{\tau_{i-1},l})}{q} \\
&\leq \max_{l\in L}|l^j|\frac{1}{V} + \max_{l\in L}|l^j|\frac{1}{V} \\
&= \frac{2}{V}\max_{l\in L}|l^j|.
\end{aligned}$$

Veamos porqué vale la última desigualdad. Llamamos $C(l^j)$ a la cantidad de vectores $l \in L$ con j-ésima coordenada l^j. Sea L^j una variable aleatoria que vale l^j con probabilidad $C(l^j)V\frac{f(X^V_{\tau_{i-1},l})}{q}$ y vale 0 con probabilidad $1 - \sum_{l/l^j\neq 0} V\frac{f(X^V_{\tau_{h-1},l})}{q}$. Entonces

$$\sum_{l/l^j\neq 0} |l^j| V \frac{f(X^V_{\tau_{h-1},l})}{q} = E(|L^j|) \leq \max_{l\in L}|l^j|.$$

■

Como consecuencia de los lemas 3.3.2 y 3.3.3 obtenemos la siguiente proposición

Proposición 3.3.4 *Supongamos que* $\max_{l\in L}|l^j| < \infty$. *Entonces existe* $\widetilde{C}$, *que depende de la dimension del espacio, de* f *y de* $max_{l\in L}|l^j|$ *tal que*

$$E(|X^V_t - Z^V_t|^2) \leq \frac{\widetilde{C}}{V}.$$

Prueba. Sea $K = N^V(t)$,

$$\begin{aligned}
E(|X_t^V - Z_t^V|^2) &= E\big(E(|X_t^V - Z_t^V|^2|\mathscr{F}_t)\big) \\
&= E\Big(E\Big(|\sum_{i=1}^{K} X_{\tau_i}^V - X_{\tau_{i-1}}^V - \frac{F(X_{\tau_{i-1}}^V)}{q}|^2|\mathscr{F}_t\Big)\Big) \\
&= E\Big(\sum_{j=1}^{d} E\big(\Big[\sum_{i=1}^{K} X_{\tau_i}^{V,j} - X_{\tau_{i-1}}^{V,j} - \frac{F^j(X_{\tau_{i-1}}^V)}{q}\Big]^2|\mathscr{F}_t\big)\Big) \\
&= \sum_{j=1}^{d} E\Big(E\big(\sum_{i=1}^{K}\Big[X_{\tau_i}^{V,j} - X_{\tau_{i-1}}^{V,j} - \frac{F^j(X_{\tau_{i-1}}^V)}{q}\Big]^2|\mathscr{F}_t\big)\Big) \\
&+ \sum_{j=1}^{d} E\Big(E\big(\sum_{i,h:\ i,h\le K, i\neq h} A^j(i)A^j(h)|\mathscr{F}_t\big)\Big) \\
&= \sum_{j=1}^{d} E\Big(E\big(\sum_{i=1}^{K}\Big[X_{\tau_i}^{V,j} - X_{\tau_{i-1}}^{V,j} - \frac{F^j(X_{\tau_{i-1}}^V)}{q}\Big]^2|\mathscr{F}_t\big)\Big) \\
&\le \sum_{j=1}^{n} E\big(C\frac{K}{V^2}\big) \\
&= dC\frac{Vq_1}{V^2}\frac{t}{T} \\
&\le \frac{\widetilde{C}}{V}.
\end{aligned}$$

■

Corolario 3.3.5 *Bajo las hipótesis de la proposición anterior vale que*

$$E(|X_t^V - Z_t^V|) \le (\frac{\widetilde{C}}{V})^{1/2}.$$

Prueba. Usando la desigualdad de Jensen tenemos que

$$(E(|X_t^V - Z_t^V|))^2 \le E(|X_t^V - Z_t^V|^2) \le \frac{\widetilde{C}}{V},$$

y aplicando raíz cuadrada a ambos lados de la igualdad obtenemos el resultado. ■

3.4. Aproximación de Z_t^V a Y_t^V

En esta sección vamos a probar que $E(|Z_t^V - Y_t^V|) \to 0$ cuando $V \to 0$. Para eso, acotamos $\frac{d}{dt}E(|Z_t^V - Y_t^V|)$.

Proposición 3.4.1 *Supongamos que existe C tal que* $E(|Z_t^V|^2) \leq C$ *e* $|Y_t^V| \leq C$ *para todo* $t \geq 0$*, con C independiente de V. Supongamos también que* $\max_{l\in L}|l^j| < \infty$*. Entonces* $E(|Z_t^V - Y_t^V|)$ *es derivable y vale*

$$\frac{d}{dt}E(|Z_t^V - Y_t^V|) \leq K(\frac{\widetilde{C}}{V})^{1/2} + KE(|Z_t^V - Y_t^V|)$$

Prueba. Sean $\mathscr{H}_t = \sigma((X_s^V)_{s\leq t}))$, $N^V = (N^V(t)_{t\geq 0})$ y $\mathscr{T}^V$ los tiempos de salto de N^V. Tomamos $t \geq 0$ y $\Delta t > 0$, tenemos

$$\begin{aligned}
E\left(|Z_{t+\Delta t}^V - Y_{t+\Delta t}^V|\mathscr{H}_t\right) &= E\left(|Z_{t+\Delta t}^V - Y_{t+\Delta t}^V|\mathbf{1}\{\mathscr{T}^V \cap (t,t+\Delta t] = 1\}|\mathscr{H}_t\right) \\
&+ E\left(|Z_{t+\Delta t}^V - Y_{t+\Delta t}^V|\mathbf{1}\{\mathscr{T}^V \cap (t,t+\Delta t] = 0\}|\mathscr{H}_t\right) \\
&+ E\left(|Z_{t+\Delta t}^V - Y_{t+\Delta t}^V|\mathbf{1}\{\mathscr{T}^V \cap (t,t+\Delta t] \geq 2\}|\mathscr{H}_t\right) \\
&= E\left(|Z_t^V + \frac{F(X_t)}{q} - Y_t^V - \frac{F(Y_t)}{q}|\mathbf{1}\{\mathscr{T}^V \cap (t,t+\Delta t] = 1\}\right) \\
&+ E\left(|Z_t^V - Y_t^V|\mathbf{1}\{\mathscr{T}^V \cap (t,t+\Delta t] = 0\}|\mathscr{H}_t\right) \\
&+ E\left(|Z_{t+\Delta t}^V - Y_{t+\Delta t}^V|\mathbf{1}\{\mathscr{T}^V \cap (t,t+\Delta t] \geq 2\}|\mathscr{H}_t\right).
\end{aligned}$$

Tomando esperanza de los dos lados de la igualdad queda

$$\begin{aligned}
E(|Z_{t+\Delta t}^V - Y_{t+\Delta t}^V|) &= E\left(|Z_t^V + \frac{F(X_t^V)}{q} - Y_t - \frac{F(Y_t^V)}{q}|\right)\mathbb{P}(\mathscr{T}^V \cap (t,t+\Delta t] = 1) \\
&+ E(|Z_t^V - Y_t^V|)\mathbb{P}(\mathscr{T}^V \cap (t,t+\Delta t] = 0) \\
&+ E(|Z_{t+\Delta t}^V - Y_{t+\Delta t}^V|\mathbf{1}\{\mathscr{T}^V \cap (t,t+\Delta t] \geq 2\}).
\end{aligned}$$

Observemos que por el Teorema 1.2.2, haciendo tender Δt a cero llegamos a que $E(|Z_t^V - Y_t^V|)$ es continua por derecha como función de t. Si Cambiamos t por $t - \Delta t$ llegamos a que $E(|Z_t^V - Y_t^V|)$ es continua como función de t. Ahora restando $E(|Z_t^V - Y_t^V|)$ y dividiendo por Δt,

$$\frac{E(|Z^V_{t+\Delta t}-Y^V_{t+\Delta t}|)-E(|Z^V_t-Y^V_t|)}{\Delta t}=$$
$$E\left(|Z^V_t+\frac{F(X^V_t)}{q}-Y^V_t-\frac{F(Y^V_t)}{q}|\right)\frac{\mathbb{P}(\mathscr{T}^V\cap(t,t+\Delta t]=1)}{\Delta t}$$
$$+E(|Z^V_t-Y^V_t|)\frac{\mathbb{P}(\mathscr{T}^V\cap(t,t+\Delta t]=0)-1}{\Delta t}$$
$$+\frac{E(|Z^V_{t+\Delta t}-Y^V_{t+\Delta t}|\mathbf{1}\{\mathscr{T}^V\cap(t,t+\Delta t]\geq 2\})}{\Delta t}.$$

Si hacemos tender Δt a 0, usando el Teorema 1.2.2, el hecho de que N^V es un Proceso de Poisson de parámetro q y que la cantidad de saltos de N^V en el intervalo $(t,t+\Delta t]$ es $N^V(t+\Delta t)-N^V(t)$ llegamos a

$$\frac{d}{dt}E(|Z^V_t-Y^V_t|)=E(|Z^V_t+\frac{F(X^V_t)}{q}-Y^V_t-\frac{F(Y^V_t)}{q}|)q-E(|Z^V_t-Y^V_t|)q.$$

Ya que, usando la desigualdad de Hölder y el Corolario 1.2.3 tenemos que

$$E(|Z^V_{t+\Delta t}-Y^V_{t+\Delta t}|\mathbf{1}\{\mathscr{T}^V\cap(t,t+\Delta t]\geq 2\})$$
$$\leq E(|Z^V_{t+\Delta t}-Y^V_{t+\Delta t}|^2)E(\mathbf{1}\{\mathscr{T}^V\cap(t,t+\Delta t]\geq 2\})$$
$$=2(E(|Z^V_{t+\Delta t}|^2)+E(|Y^V_{t+\Delta t}|^2))P(\mathbf{1}\{\mathscr{T}^V\cap(t,t+\Delta t]\geq 2\})=o(\Delta t).$$

A su vez, usando la desigualdad triangular, y el hecho de que F es una función Lipchitz con constante K obtenemos

$$\begin{aligned}E(|Z^V_t+\frac{F(X^V_t)}{q}-Y^V_t-\frac{F(Y^V_t)}{q}|)q &- E(|Z^V_t-Y^V_t|)q\\ &\leq E(|\frac{F(X^V_t)}{q}-\frac{F(Y^V_t)}{q}|)q\\ &= E(F(X^V_t)-F(Y^V_t))\\ &\leq KE(|X^V_t-Y^V_t|)\\ &\leq K(E(|X^V_t-Z^V_t|)+E(|Z^V_t-Y^V_t|)).\end{aligned}$$

Teniendo en cuenta el Corolario 3.3.5 obtenemos la siguiente desigualdad en el intervalo $[0,T]$

$$\frac{d}{dt}E(|Z_t^V - Y_t^V|) \leq K(\frac{\widetilde{C}}{V})^{1/2} + KE(|Z_t^V - Y_t^V|) \tag{3.4.1}$$

■

Corolario 3.4.2 *Bajo las hipótesis de la proposición anterior tenemos que*

$$E(|Z_t^V - Y_t^V|) \leq K(\frac{\widetilde{C}}{V})^{1/2} T e^{KT}.$$

Prueba. Integrando la desigualdad (3.4.1) en el intervalo $[0,t]$ llegamos a que

$$\begin{aligned} E(|Z_t^V - Y_t^V|) &\leq K(\frac{\widetilde{C}}{V})^{1/2} t + \int_0^t KE(|Z_s^V - Y_s^V|)ds \\ &\leq K(\frac{\widetilde{C}}{V})^{1/2} T + \int_0^t E(|Z_s^V - Y_s^V|)ds. \end{aligned}$$

Con lo cual $E(|Z_t^V - Y_t^V|)$ cumple las hipótesis del teorema de Gronwall en $[0,T]$ y, por lo tanto, vale la desigualdad que queremos probar.

■

3.5. Aproximación de Y_t^V a $x(t)$

En esta sección vamos a probar que $E(|Y_t^V - x(t)|) \to 0$ cuando $V \to \infty$. Para eso vamos a usar la siguiente proposición.

Proposición 3.5.1 *Supongamos que $|Y_t^V| \leq C$ para todo $V \in \mathbb{N}$ y $t \geq 0$, con $C \in \mathbb{R}$ independiente de V. Supongamos también que F es una función localmente Lipchitz y que $Y_0^V \to x_0$ cuando $V \to \infty$. Entonces*

$$E(|Y_t^V - x(t)|) \to 0 \text{ cuando } V \to \infty$$

Prueba. Tenemos que $(Y_t^V)_{t\geq 0}$ espera un tiempo exponencial con esperanza $\frac{Vq_1}{T}$ y salta a $Y_t^V + \frac{1}{V}\frac{T}{q_1}F(Y_t)$. Sea $(\mathscr{Y}_k^V)_{k\in\mathbb{N}}$ el esqueleto de $(Y_t^V)_{t\geq 0}$ entonces, $\mathscr{Y}_{k+1}^V = \mathscr{Y}_k^V + \frac{1}{V}\frac{T}{q_1}F(\mathscr{Y}_k^V)$. O sea, $\mathscr{Y}_k^V$ es el k-ésimo paso del método de Euler con paso $\frac{1}{V}$ para la ecuación $\dot{\underline{x}}(t) = \frac{T}{q_1}F(\underline{x}(t))$ en el intervalo $[0,1]$. Entonces por el Corolario 1.4.5 tenemos que

$$\max_{1\leq k\leq V}\{|\mathscr{Y}_k^V - \underline{x}\left(\frac{k}{V}\right)|\} \to 0 \text{ cuando } V\to\infty$$

Ahora si $x(t)$ es la solución de la ecuación $\dot{x}(t) = F(x(t))$ entonces, por unicidad de la solución, tenemos que $\underline{x}(t) = x(\frac{tT}{q_1})$. Por lo tanto

$$\max_{1\leq k\leq V}\{|\mathscr{Y}_k^V - x\left(\frac{kT}{q_1V}\right)|\} \to 0 \text{ cuando } V\to\infty$$

Luego si $s < \frac{T}{q_1}$ y $N^V(s)\leq V$, entonces

$$|Y_s^V - x(\frac{N^V(s)T}{q_1V})| \to 0 \text{ cuando } V\to\infty$$

Ya que si $N^V(s) = k$ entonces $\mathscr{Y}_k^V = Y_s$.
Por otro lado, por como construimos $(Y_t^V)_{t\geq 0}$ tenemos que

$$\frac{N^V(s)T}{q_1V} \to s \text{ c.t.p cuando } V\to\infty.$$

Sea $\delta > 0$ tal que $\frac{q_1(s+\delta)}{T}\leq 1$, entonces $\frac{N^V(s)T}{q_1V}\leq s+\delta$ implica que

$$N^V(s)\leq \frac{(s+\delta)q_1V}{T}\leq V.$$

Ahora, como

$$\frac{N^V(s)T}{q_1V} \to s \text{ c.t.p cuando } V\to\infty$$

vale en particular que

$$\mathbb{P}\left(\frac{N^V(s)T}{q_1V} > s+\delta\right)\to 0 \text{ cuando } V\to\infty.$$

Tenemos

$$E(|Y_s^V - x(s)|)\leq E(|Y_s^V - x(\frac{N^V(s)T}{q_1V})|) + E(|x(\frac{N^V(s)T}{q_1V}) - x(s)|).$$

Para terminar la prueba veamos que cada uno de esos términos tiende a cero

1) Por el Corolario 1.4.5 sabemos que si $N^V(s) \leq V$ entonces

$$|Y_s^V - x(\frac{N^V(s)T}{q_1 V})| \leq (|x_0 - y_0^V| + K\frac{T}{V}T\|F\|_\infty)e^{KT}.$$

Con K la constante Lipchitz de F en $\overline{B_C(0)} \cup U$. Usando esto, la desigualdad de Hölder y que la solución de la ecuación es una trayectoria acotada tenemos que

$$E\left(|Y_s^V - x\left(\frac{N^V(s)T}{q_1 V}\right)|\right)$$

$$= E\left(|Y_s^V - x\left(\frac{N^V(s)T}{q_{\ 1}}V\right)|\mathbf{1}\left\{\frac{N^V(s)T}{q_1 V} \leq s+\delta\right\}\right)$$

$$+E\left(|Y_s^V - x\left(\frac{N^V(s)T}{q_1 V}\right)|\mathbf{1}\left\{\frac{N^V(s)T}{q_1 V} > s+\delta\right\}\right)$$

$$\leq \frac{c}{VK}(e^{K\frac{T}{q_1}} - 1)\mathbb{P}\left(\frac{N^V(s)T}{q_1 V} \leq s+\delta\right)$$

$$+E\left(|Y_s^V - x\left(\frac{N^V(s)T}{q_1 V}\right)|^2\right)\mathbb{P}\left(\frac{N^V(s)T}{q_1 V} > s+\delta\right).$$

ambos términos tienden a 0 cuando $V \to \infty$ ya que

$$E\left(|Y_s^V - x\left(\frac{N^V(s)T}{q_1 V}\right)|^2\right) \leq 2\left(E\left(|Y_s^V|^2\right) + E\left(|x\left(\frac{N^V(s)T}{q_1 V}\right)|^2\right)\right) \leq C.$$

2) Como x es una función continua y $\frac{N^V(s)T}{q_1 V} \to s$ c.t.p cuando $V \to \infty$, se tiene que $x\left(\frac{N^V(s)T}{q_1 V}\right) \to x(s)$ c.t.p o, dicho de otra forma

$$\left|x\left(\frac{N^V(s)T}{q_1 V}\right) - x(s)\right| \to 0 \quad \text{c.t.p}$$

Luego, por el teorema de convergencia mayorada sigue que

$$E\left(\left|x\left(\frac{N^V(s)T}{q_1 V}\right) - x(s)\right|\right) \to 0,$$

■

3.6. Prueba del Teorema

Finalmente, vamos a probar el Teorema 3.0.3,

Prueba del Teorema 3.0.3. Para empezar podemos suponer que existe U_2 tal que $Y_t^V, Z_t^V \in U_2$ para todo $t \geq 0$ y $V \in \mathbb{N}$ ya que en caso de que esto no se cumpla, como sabemos que $X_t^V \in U$ para todo $t \geq 0$, $V \geq 0$ y $x(t) \in U$ para todo $t \geq 0$ podemos modificar f fuera de U para que se cumpla. Tomamos U_1 tal que $\overline{U} \subset U_1$ y $\widetilde{f}$ tal que $\widetilde{f} = f$ en U, $\widetilde{f} = 0$ fuera de U_1 y $\widetilde{f}$ tenga las mismas condiciones de regularidad que f.

Ahora, como $\max_l |l^j| < \infty$ vale, por el Corolario 3.3.5, que

$$E(|X_t^V - Z_t^V|) \to 0 \quad \text{cuando } V \to \infty.$$

Como $X_t^V \in U$ para todo $V \in \mathbb{N}$ sabemos que $E(|X_t^V|^2) \leq c$, con c independiente de V , entonces usando nuevamente el Corolario 3.3.5

$$E(|Z_t^V|^2) = E(|X_t^V - Z_t^V + X_t^V|^2) \leq 2\big(E(|X_t^V - Z_t^V|^2) + E(|X_t^V|^2)\big) \leq \bar{c}.$$

Luego, se cumplen las condiciones del Corolario 3.4.2 y, por lo tanto,

$$E(|Z_t^V - Y_t^V|) \to 0 \quad \text{cuando } V \to \infty.$$

Como $|Y_t^V| \leq C$ para todo $t \geq 0$ con C independiente de V, $Y_0^V \to x_0$ cuando $V \to \infty$ y F es una función localmente Lipchitz vale por la Proposicón 3.5.1 que

$$E(|Y_t^V - x(t)|) \to 0 \quad \text{cuando } V \to \infty.$$

En definitiva

$$E(|X_t^V - x(t)|) \leq E(|X_t^V - Z_t^V|) + E(|Z_t^V - Y_t^V|) + E(|Y_t^V - x(t)|) \to 0$$

cuando $V \to \infty$.

■

Observación 3.6.1 *Como convergencia en L^1 implica convergencia en probabilidad probamos también que, para cada $t \geq 0$, X_t^V tiende a $x(t)$ en probabilidad cuando V tiende a infinito.*

3.7. Aplicaciones

En esta sección vamos a relacionar el Teorema 3.0.3 con las aplicaciones que presentamos en el Capítulo 2. Vamos a ver cuáles de los ejemplos caen en las hipótesis del teorema y cuáles no. Empecemos con uno de los modelos epidémicos

Ejemplo 3.7.1 *Vamos a aplicar el Teorema 3.0.3 al modelo SIR mencionado en la subsección 2.2.2. Tomamos* $U = \{(s,i) \in \mathbb{R}^2 : s,i \geq 0, s+i \leq 1\}$*, para ver que* $(S(t),I(t)) \in U$ *para todo* $t \geq 0$ *vamos a ver que si tenemos la ecuación* (2.2.2) *entonces se cumple que* $I(t),R(t),I(t) \geq 0$ *y luego como* $S(t)+I(t)+R(t) = 1$ *tenemos que* $S(t)+I(t) \leq 1$*. Supongamos que* $I(t) = 0$ *para algun* $t > 0$ *entonces tenemos que* $I(t) = 0$ *para todo* $t > 0$ *ya que si no, estamos contradiciendo el teorema de unicidad de solución. Ahora si* $R(t) = 0$ *entonces* $\dot{R}(t) = \gamma I(t)$ *y aumentan los recuperados pues ya vimos que* $I(t) \geq 0$*. Por último, si* $S(t) = 0$ *tenemos que* $\dot{S}(t) = b(I(t)+R(t))$ *y aumentan los susceptibles ya que* $I(t),R(t) \geq 0$*.*

Por otro lado, dado $(s,i) \in \mathbb{Z}^2$*, si* $s = 0$ *entonces* $q((s,i),(s-1,i)) = 0$ *entonces si empezamos con* $s \geq 0$ *seguimos con* $s \geq 0$*. Ahora si* $i = 0$ *entonces* $q((s,i),(s,i-1)) = q((s,i),(s+1,i-1)) = 0$ *entonces si empezamos con* $i \geq 0$ *seguimos con* $i \geq 0$ *y si* $\frac{s}{V} + \frac{i}{V} = 1$ *entonces* $q((s,i),(s+1,i)) = 0$ *entonces se va a cumplir que* $\frac{s}{V} + \frac{i}{V} \leq 1$ *ya que los demás saltos no alteran esta condición. Luego,* $\frac{1}{V}(S_t^V, I_t^V) \in U$ *para todo* $V \in \mathbb{N}$ *y* $t \geq 0$*.*

Es claro que también se cumple que $\max_l |l^j| < \infty$ *y si* $x = (x_1,x_2) \in U$ *entonces* $0 \leq x_1,x_2 \leq 1$ *y se cumple también que* $\bar{q} = \max_{x \in U}\{\sum_{l \in L} f(x,l)\} < \infty$*. Con respecto a los datos iniciales, si* $S(0) = p_1$ *e* $I(0) = p_2$*, con* $p_1,p_2 \geq 0$ *tal que* $p_1 + p_2 \leq 1$*.Entonces si tomamos* $(S_0^V, I_0^V) = ([Vp_1],[Vp_2])$ *vale que* $\frac{1}{V}(S_0^V, I_0^V) \to (S(0),I(0))$ *cuando* $V \to \infty$*. Por último, es fácil ver que F es localmente Lipchitz.*

Luego, por el teorema anterior tenemos que $E(|\frac{1}{V}(S_t^V, I_t^V) - (S(t),I(t))|) \to 0$ *cuando* $V \to \infty$*.*

De forma similar se puede ver también que el teorema es aplicable al modelo SIS presentado en la Sección 2.2.1. Veamos un ejemplo en el cual el Teorema 3.0.3 no se puede aplicar

Ejemplo 3.7.2 *Tomemos el modelo poblacional determinístico dado por la ecuación 2.1.2 y su correspondiente modelo estocástico presentados en la Sección 2.1. Se puede ver que, en este caso, que F es una función localmente Lipchitz y que existe U acotado tal que si* $x_0 \in U$ *entonces* $x(t) \in U$ *para todo* $t \geq 0$*. Es claro también que se cumple que* $\max_l |l^j| < \infty$ *y se puede ver que como U es acotado* $\bar{q} = \max_{x \in U}\{\sum_{l \in L} f(x,l)\} < \infty$*. Podemos tomar también* $(X_0^V)_{V \in \mathbb{N}}$ *tal que*

$X_0^V \to x_0$ cuando $V \to \infty$. Pero no es cierto que existe algún U_1 acotado tal que $X_t^V \in U_1$ para todo $t \geq 0$ ya que, por ejemplo, $q(k,k+e_1) = \lambda_1 k_1 > 0$ para todo k tal que $k_1 > 0$. Es decir hay una probabilidad positiva de que haya n conejos para cualquier $n \in \mathbb{N}$ y, por lo tanto, dado U_1 acotado hay una probabilidad positiva de que $X_t^V \notin U$ para algun $t \geq 0$. Más allá de eso, como $x(t) \in U$ para todo $t \geq 0$ se puede ver que con probabilidad alta $X_t^V \in U$ por mucho tiempo ya que vimos que F es el drift de X_t^V. Este resultado, sin embargo, excede el trabajo de esta tesis.

Los modelos de la cola M/M/∞ y la cola M/M/1, presentados en las subsecciones 2.3.2 y 2.3.1, son similares al Ejemplo 3.7.2 ya que $x(t)$ es una función acotada, F es una funcioń localmente Lipchitz y se cumplen las demás condiciones pero la cadenas $(X_t^V)_{t\geq 0}$ no son acotadas. Lo mismo ocurre con el modelo poblaciónal dado por la ecuación (2.1.1), presentado en la Sección 2.1. Se puede ver, sin embargo, que el teorema demostrado por Kurtz en [4] se puede aplicar a cualquiera de los ejemplos que vimos.

Capítulo 4

Apéndice

En este apéndice damos una prueba de un resultado similar al obtenido por Kurtz (ver [4]). Esta prueba fue realizada por Darling y Norris en una publicación intitulada "Differential aproximations for Markov chains"[2]. Una diferencia importante entre la prueba de Darling y Norris y la del Teorema 3.0.3 es el uso de la proposición 4.1.1, esta proposición da una desigualdad con Martingalas similar a la desigualdad de Doob. Damos también en este Capítulo algunos detalles de las simulaciones junto con los correspondientes códigos.

4.1. Prueba de Darling y Norris

Sea $(X_t)_{t\geq 0}$ una cadena de Markov a tiempo continuo con matriz de tasas Q en espacio de estados a lo sumo numerable. Asumimos que el conjunto $\{q(x): x \in S\}$ es acotado. Elegimos funciones

$$z^i : S \to \mathbb{R}, \text{ para } i = 1, ..., d.$$

Escribimos

$$z = (z^1, ..., z^d) : S \to \mathbb{R}^d$$

y consideramos el proceso a valores en $\mathbb{R}^d$, $\mathbf{X}_t = z(X_t)$. Llamamos β al drift del proceso $\mathbf{X}_t$.

Sean $U \subset \mathbb{R}^d$ y $x_0 \in U$. Dada $F : U \to \mathbb{R}^d$ una función Lipschitz es sabido que la ecuación diferencial $\dot{x}(t) = F(x(t))$ tiene solución única $(x(t))_{t\leq \zeta}$, empezando en x_0, y tal que $x(t) \in U$ para todo $t < \zeta$.

Para aproximar el proceso y la ecuación diferencial comparamos las siguientes igualdades

$$\mathbf{X}_t = \mathbf{X}_0 + M_t + \int_0^t \beta(X_s)ds, \;\; 0 \leq t \leq T_1, \tag{4.1.1}$$

$$x(t) = x_0 + \int_0^t F(X_s)ds, \;\; 0 \leq t \leq \zeta. \tag{4.1.2}$$

Donde $T_1 = \text{ínf}\{t \geq 0/\beta(X_t) = \infty\}$ y la primera ecuación sirve para definir el proceso $(M_t)_{t \leq T_1}$. Definimos también para $x \in S$

$$\alpha(x) = \Sigma_{x' \neq x} |z(x') - z(x)|^2 q(x,x').$$

Vamos a usar la siguiente proposición

Proposición 4.1.1 *Sean T_1 y $(M_t)_{t \leq T_1}$ definidos como en* (4.1.1)*. Entonces si $T \leq T_1$*

$$E\left(\sup_{t \leq T} |M_t|^2\right) \leq 4 \int_0^T \alpha(X_t)dt.$$

Prueba. Ver [2, pag76]. ■

Llamemos K a la constante de Lipschitz de F en U con respecto a la norma euclidiana $|.|$. Fijemos $t_0 < \zeta$ y $\varepsilon > 0$ y asumamos que

$$\forall x \in S \text{ y } t \leq t_0, |z(x) - x(t)| \leq \varepsilon \Longrightarrow z(x) \in U. \tag{4.1.3}$$

Sea $\delta = \varepsilon \frac{e^{-Kt_0}}{3}$ y fijemos $A > 0$. Para que el resultado sea útil vamos a necesitar que A sea chico con respecto a ε^2. Sea $T = \text{ínf}\{t \geq 0/\mathbf{X}_t \notin U\}$.

Consideremos los siguientes eventos

$$\Omega_0 = \{|\mathbf{X}_0 - x_0| \leq \delta\}, \quad \Omega_1 = \left\{\int_0^{T \wedge t_0} |\beta(X_t) - F(z(X_t))|dt \leq \delta\right\}, \tag{4.1.4}$$

$$\Omega_2 = \left\{\int_0^{T \wedge t_0} \alpha(X_t) \leq At_0\right\}, \quad \Omega_2' = \{T \wedge t_0 \leq T_1 \;\text{ y }\; \sup_{t \leq T \wedge t_0} |M_t| \leq \delta\}. \tag{4.1.5}$$

Teorema 4.1.2 *Sea $(X_t)_{t \geq 0}$ una cadena de Markov a tiempo continuo con matriz de tasas Q en espacio de estados S tal que $\sup_{x \in S}\{q(x)\} < \infty$. Sean $z = (z^1, ..., z^d) : S \to \mathbb{R}^d$ y $\mathbf{X}_t = z(X_t)$. Sean $U \subset \mathbb{R}^d$, $x_0 \in U$, F una función Lipchitz*

en U con constante K y $x(t)$ la solucón de la ecuación diferencial $\dot{x}(t) = F(x(t))$ con valor inicial x_0. Asumamos que se cumple la condición (4.1.3) *entonces*

$$\mathbb{P}(\sup_{t\leq t_0}|\mathbf{X}_t - x(t)| > \varepsilon) \leq 4\frac{At_0}{\delta^2} + \mathbb{P}(\Omega_0^c \cup \Omega_1^c \cup \Omega_2^c)$$

con Ω_0, Ω_1 y Ω_2 definidos como en (4.1.4) *y* (4.1.5).

Prueba. Consideramos la función aleatoria $f(t) = \sup_{s\leq t}|\mathbf{X}_s - x(s)|$ definida en el intervalo $[0, T \wedge t_0]$. Usando las ecuaciones (4.1.1) y (4.1.2), sumando y restando $F(z(X_s))$ en la integral, tenemos que

$$f(t) \leq |\mathbf{X}_0 - x(0)| + \sup_{s\leq t}|M_s| + \int_0^t |\beta(X(s)) - F(z(X_s))|ds + \int_0^t |F(\mathbf{X}) - F(x(s))|ds.$$

Ahora, si estamos en el evento $\Omega_0 \cap \Omega_1 \cap \Omega_2'$, tenemos que

$$f(t) \leq 3\delta + \int_0^t K|\mathbf{X}_s - x(s)|ds \leq 3\delta + K\int_0^t f(s)ds$$

O sea, que f cumple las hipótesis del Teorema de Gronwall con $C = 3\delta$ y $D = K$ y, por lo tanto, $f(T \wedge t_0) \leq 3\delta e^K t_0 = \varepsilon$. En particular, vale que

$$|z(X_{T\wedge t_0}) - x(T \wedge t_0)| \leq \varepsilon$$

y entonces $z(X_{T\wedge t_0}) \in U$. Como esto último implica que $T > t_0$ vale que $f(t_0) \leq \varepsilon$.

Consideramos ahora el tiempo de parada

$$\tilde{T} = T \wedge t_0 \wedge \text{ínf}\{t \geq 0 / \int_0^t \alpha(X_s)ds > At_0\}$$

Usando la desigualdad de Cauchy-Schwarz, tenemos que, para $x \in S$

$$\begin{aligned} |\beta(x)|^2 &= \Sigma_{x\neq x'}(z(x') - z(x))(q(x,x'))^{\frac{1}{2}}(q(x,x'))^{\frac{1}{2}} \\ &\leq \Sigma_{x\neq x'}q(x,x')\Sigma_{x\neq x'}(z(x') - z(x))^2 q(x,x') \\ &= q(x)\alpha(x). \end{aligned}$$

Entonces si $t < \tilde{T}$

$$\begin{aligned} \int_0^t |\beta(X_s)|^2 ds &\leq \sup_{x\in S}\{q(s)\}\int_0^t \alpha(X_s)ds \\ &\leq \sup_{x\in S}\{q(x)\}At_0, \end{aligned}$$

y por lo tanto $\beta(X_s) < \infty$ en $[0,t]$ y $\tilde{T} \leq T_1$. Ahora usando la Proposición 4.1.1 tenemos que

$$E\left(\sup_{t\leq\tilde{T}}|M_t|^2\right) \leq 4At_0.$$

En Ω_2, tenemos que $\tilde{T} = T \wedge t_0$ y, por lo tanto, $\Omega_2\backslash\Omega_2' \subseteq \{\sup_{t\leq\tilde{T}}|M_t| > \delta\}$. Usando la desigualdad de Chebyshev llegamos a que $\mathbb{P}(\Omega_2\backslash\Omega_2') \leq 4\frac{At_0}{\delta^2}$. ■

4.1.1. Aplicación del Teorema a un modelo epidémico

En esta sección vamos a aplicar el Teorema 4.1.2 a un modelo epidémico similar al modelo SIR presentado en la subsección 2.2.2. Dado $V \in \mathbb{N}$ tenemos la cadena de Markov $(X_t^V)_{t\geq 0}$ con espacio de estados $S = \{k = (k_1,k_2) \in \mathbb{Z}^2_{\geq 0}/k_1 + k_2 \leq V\}$ y tasas de salto

$$\begin{array}{ll} q(k,k+e_2-e_1) = \lambda\frac{k_1k_2}{V} & \text{Se infecta un susceptible,} \\ q(k,k-e_2) = \mu k_2 & \text{Se recupera un infectado.} \end{array}$$

Donde (k_1,k_2) representa la cantidad individuos susceptibles e infectados, respectivamente, $\lambda > 0$ es la tasa de infección y μ la tasa de recuperación suponiendo que si un individuo se recupera entonces pasa a ser inmune. Supongamos que empezamos con una proporción p de infectados, o sea, $X_0^V = (V(1-p), Vp)$. Además vamos a suponer que $\mu = 1$. La diferencia con el modelo SIR presentado en la sección 2.2.2 es que en este modelo no se tienen en cuenta las muertes de los individuos.

Definimos $z : S \to \mathbb{R}^2$ como $z(k_1,k_2) = \frac{1}{V}(k_1,k_2)$ y $\mathbf{X}_t^V = z(X_t^V)$. Tenemos entonces que

$$\begin{aligned} \beta(k_1,k_2) &= (-\frac{1}{V}\lambda\frac{k_1k_2}{V}, \frac{1}{V}\lambda\frac{k_1k_2}{V}) - (0, \frac{1}{V}k_2) \\ &= (-\frac{1}{V}\lambda\frac{k_1k_2}{V}, \frac{1}{V}\lambda\frac{k_1k_2}{V} - \frac{1}{V}k_2). \end{aligned}$$

Tomamos entonces para $(x_1,x_2) \in U = [0,1]^2$,

$$F(x_1,x_2) = (-\lambda x_1x_2, \lambda x_1x_2 - \mu x_2).$$

Si tomamos $x_0 = (1-p,p)$ podemos escribir la ecuación diferencial $\dot{x}(t) = F(x(t))$ como

$$\begin{cases} \dot{x}_1(t) = -\lambda x_1(t)x_2(t), \\ \\ \dot{x}_2(t) = \lambda x_1(t)x_2(t) - x_2(t). \end{cases}$$

Esta ecuación tiene solución única para $t \geq 0$. Notemos que como $z(S) \subseteq U$ la condición 4.1.3 se cumple para todo $\varepsilon > 0$ y t_0. Notemos también que como $\beta(X_t) = F(z(x_t))$ se cumple que $\Omega_1 = \Omega$ y que como $\mathbf{X}_0 = x_0$ se cumple que $\Omega_0 = \Omega$. La constante de Lipschitz para F en U es $K = \lambda + \text{máx}\{\lambda, 1\}$. Si tomamos $A = \frac{2\lambda+1}{V}$ tenemos que

$$\int_0^{t_0} \alpha(x_t)dt = \int_0^{t_0} |\frac{1}{V}(-1,1)|^2 \lambda \frac{X_t^{V,1} X_t^{V,2}}{V} + |\frac{1}{V}(-1,0)|^2 X_t^{V,2} dt \leq At_0$$

con lo cual $\Omega_2 = \Omega$ y aplicando el Teorema 4.1.2 vale que

$$\mathbb{P}(\sup_{t \leq t_0} |\mathbf{X}_t^V - x(t)| > \varepsilon) \leq 4\frac{At_0}{\delta^2}$$

Ahora como, para todo $\varepsilon > 0$, el término de la derecha tiende a 0 cuando $V \to \infty$ tenemos que $\sup_{t \leq t_0} |\mathbf{X}_t^V - x(t)|$ tiende en probabilidad a cero y, en particular, que $\mathbf{X}_t^V$ tiende en probabilidad a $x(t)$ para todo $t \geq 0$. De manera muy similar se puede ver que el Teorema 4.1.2 se puede aplicar al modelo SIR.

4.2. Las simulaciones

En esta sección vamos mostrar las simulaciones usadas para las figuras que aparecen el Capťulo 2. Las simulaciones están hechas con Matlab, este programa con el comando rand da un numero al azar distribuido uniformemente en el $[0,1]$. A partir de eso, podemos generar la distribución exponencial para simular el tiempo de espera y también elegir a que estado vamos a saltar. Por ejemplo, si podemos saltar desde el 1 al 2 ó al 3 con probabilidad $\frac{1}{2}$ entonces si el numero aleatorio que nos da el Matlab es menor que $\frac{1}{2}$ saltamos al 2, caso contrario, saltamos al 3. Entonces si tenemos las tasas del proceso y estamos en el lugar x basta con dividir el intervalo $[0,1]$ en intervalos $I(x,y)$ de longitud $\frac{q(x,y)}{q(x)}$. Así, si el numero al azar que nos da el Matlab cae en el intervalo $I(x,y)$ saltamos al lugar y. Más formalmente, estamos usando el Teorema 1.1.5.

Este es el código de la simulación para modelo epidémico.

```
close all
clear all

%par\'ametros de la ecuaci\'on
beta=1;
gamma=1;
b=1;

% datos iniciales
V=10000;
k1=4000;
k2=4000;
x(1)=0.4;
y(1)=0.4;
T(1)=0;

%cantidad de pasos que vamos a hacer de la cadena de Markov
iteraciones=100000;
termina_programa=0;

while(termina_programa<=iteraciones)

    qdelpuntoactual= (gamma+b)*k2+ k1*k2/V+ b*(V-k1-k2) ;

                if(qdelpuntoactual>0)
                    i=termina_programa+1;
%espero un tiempo exponencial de par\'ametro qdelpuntoactual
                    T(i+1)=T(i)-log((1-rand))/qdelpuntoactual;
                    azar=qdelpuntoactual*rand;

%elijo a donde voy a saltar
                    if(azar<=k1*k2/V)
                        k1=k1-1;
                        k2=k2+1;

                    elseif(azar<=k1*k2/V+ gamma*k2 )
                        k2=k2-1;
                    elseif(azar<= k1*k2/V+ gamma*k2+b*k2)
```

```
                        k1=k1+1;
                        k2=k2-1;
                    else
                        k1=k1+1;
                    end
                else
                    termina_programa=iteraciones +1;
                end
                x(i+1)=k1/V;
                y(i+1)=k2/V;

    termina_programa=termina_programa + 1;

end

%gr\'afico la primer coordenada de la cadena
% de Markov y de la ecuaci\'on
hold on
[t,z]=ode45(@F1,[0,T(i+1)],[0.4 0.4]);
plot(T,x,'r');
hold on
plot(t,z(:,1));
print -depsc simulacion_epid\'emica_x

%gr\'afico la segunda coordenada de la cadena
% de Markov y de la ecuaci\'on
figure
hold on
plot(T,y,'r');
hold on
plot(t,z(:,2));
print -depsc simulacion_epid\'emica_y

hold off
```

Usamos esta función para graficar la ecuación diferencial

```
function xdot = F1(t,x);
```

```
X = x(1,:);  %La primera fila tiene los valores de x
Y = x(2,:);  %La segunda fila tiene los valores de y

U = -X.*Y+ones(size(X))-X;
V =  X.*Y-2*Y;

xdot = [U; V];
```

Las simulaciones restantes son similares a esta pero tenemos que cambiar la función F que usamos.

Bibliografía

[1] Fred Brauer, Pauline van den Driessche, and Jianhong Wu (eds.), *Mathematical epidemiology*, Lecture Notes in Mathematics, vol. 1945, Springer-Verlag, Berlin, 2008, Mathematical Biosciences Subseries. MR 2452129 (2009k:92004)

[2] R. W. R. Darling and J. R. Norris, *Differential equation approximations for Markov chains*, Probab. Surv. **5** (2008), 37–79. MR 2395153 (2009m:60196)

[3] M. Jonckheere, *Introduction to stochastic networks*, 2010, Preprint.

[4] Thomas G. Kurtz, *Solutions of ordinary differential equations as limits of pure jump Markov processes*, J. Appl. Probability **7** (1970), 49–58. MR 0254917 (40 #8124)

[5] J. R. Norris, *Markov chains*, Cambridge Series in Statistical and Probabilistic Mathematics, vol. 2, Cambridge University Press, Cambridge, 1998, Reprint of 1997 original. MR 1600720 (99c:60144)

[6] T. Seppalainen, *Translation invariance exclusion processes*, Preprint.

[7] H. Thorisson, *Coupling, stationarity and regenaration*, Springer, New York, 2000.