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

# Procesos de ramificación multitipo, procesos de urna y simulación de distribuciones cuasiestacionarias en espacios numerables

Tesis presentada para optar al título de Doctor de la Universidad de Buenos Aires en el área Ciencias Matemáticas

**Analia Soledad Ferrari**

Director de tesis: Pablo Groisman.
Consejero de estudios: Pablo Groisman.

Fecha de defensa: 16 de junio de 2017.

# Procesos de ramificación multitipo, procesos de urna y simulación de distribuciones cuasiestacionarias en espacios numerables


**Resumen**

En esta tesis abordamos principalmente dos temas: procesos de ramificación multitipo y distribuciones cuasiestacionarias. Primero estudiamos procesos de ramificación multitipo en espacios de tipos numerables. Trabajamos en el caso en que la población tiene probabilidad positiva de no extinguirse y analizamos el comportamiento asintótico de la población. Más específicamente probamos que la proporción de individuos de cada tipo converge en probabilidad a una medida en el espacio de tipos. Este resultado es bien conocido cuando el espacio de tipos es finito y bajo fuertes hipótesis en caso infinito.

Como aplicación del estudio de estos procesos consideramos procesos de urna con infinitos colores y probamos que la proporción de la cantidad de bolas de cada color dentro de la urna converge en probabilidad a una medida determinística.

En segundo lugar estudiamos distribuciones cuasiestacionarias (QSD) y, en particular, métodos que las aproximan. Cuando el espacio de estados es finito, Aldous, Flannery y Palacios (AFP) propusieron un proceso no condicionado, cuya medida empírica converge a la QSD deseada. Generalizamos (bajo ciertas hipótesis sobre la matriz de transición) el método de AFP para espacios de estados numerable. La demostración de la convergencia de este método se basa en los resultados obtenidos para procesos de ramificación multitipo.

Por último consideramos el proceso de nacimiento y muerte perezoso con deriva hacia el origen. Usando la desigualdad de Holley probamos que las QSD de las cadenas truncadas convergen a la QSD minimal monótonamente.



**Palabras Clave:** Distribuciones cuasiestacionarias, simulaciones, procesos de ramificación, procesos de urna.

# Multi-type branching processes, urn process and simulation of quasi-stationary distributions on countable spaces

**Abstract**

This thesis deals mainly with two topics: multi-type branching processes and quasi-stationary distributions. We first study multi-type branching processes with countable types. We work in the case where the population has positive probability of survival and focus on the study of the asymptotic behavior of the population. More specifically, we prove that the proportion of individuals of each type converges in probability to a measure on the type space. This result is well known when the type space is finite or under strong hypotheses in the infinite case.

As an application, we consider urn processes with infinite colors and prove that the proportion of the amount of balls of each color in the urn converges in probability to a deterministic probability measure.

We also study quasi-stationary distributions (QSD) and methods to approximate them. When the state space is finite, Aldous, Flannery and Palacios (AFP) proposed an unconditioned process for which the empirical measure converges to the desired QSD. We generalize (under certain assumptions on the transition matrix) the AFP method to countable spaces. The proof of convergence of the method is based on the results obtained for multi-type branching processes.

Finally we consider the lazy birth and death process with drift towards the origin. By means of Holley inequality we prove that the QSD of the truncated chains converge to the minimal QSD.

**Keywords:** Quasi-stationary distributions, simulations, branching processes, urn processes.

# Agradecimientos

**Gracias:**

Al Conicet por las becas.
A la UBA por toda la formación que me dió.
A Patu, por dirigirme, por enseñarme, por las ideas y la mucha paciencia.
A Ferrari por ayudarme y estar siempre dispuesto a escucharme.
Al jurado por tomarse el tiempo de leer esta tesis.
Al grupo de proba, las probabilidades son más divertidas cuando se aprenden con ustedes.
A Santi y Juli por estar siempre dispuestos a escuchar mis dudas.
A Marie, que hace muchos años me mostró el mundo de las probabilidades.
A la 2103 por el aguante.
A Juli y Pau porque trabajar con ustedes fue un placer que ya extraño.
A la gente con la que trabajo, CBC, UTDT y FIUBA, por interesarse por mi doctorado y por darme tiempo para que pueda terminar.
A mis amigos de la facu Carito, Flopa, Lau, Marian, Nicos, Manu, Juli, Sergio por todo lo que compartimos.
A mis amigos de siempre Guada, Dani, Sol, Sil, Fer, Ini, Inés, Mara, Agus, Marian con ustedes mi vida es mas divertida.
A mi familia por su apoyo constante y por ponerse felices con mis logros.
A César, por acompañarme en todas, por ayudarme tanto, por tu inmensa paciencia, por aguantar mi enojo cada vez que trabaja en la tesis, por poner de premio lo que más quiero en mi vida. Sin tu ayuda terminar no hubiese sido posible, a vos gracias totales.

*A mi papá.*

# Índice general

# Introducción

Desde que Markov introduce su famoso modelo en 1906 [Mar06], este ha sido aplicado en diversas áreas de la ciencia; en biología, genética, dinámica de poblaciones, en ciencias sociales, psicología, física, investigación operativa, etc. Cabe destacar que si bien la formalización de los procesos de Markov fue novedosa, años antes ya se venían estudiando dichos procesos. Por ejemplo, en 1873 Galton y Watson estudiaron la supervivencia o extinción de los apellidos aristocráticos dando lugar a los procesos de Markov llamados procesos Galton-Watson [WG75]. Estos procesos fueron ampliamente estudiados, destacando Steffensen [Ste33] donde se determina la probabilidad de extinción y Kolmogorov [Kol38] que estudia el comportamiento asintótico del proceso cuando las familias no se extinguen.

Bartlett [Bar49] y Kolmogorov y Dmitriev [KD47] extendieron este modelo para considerar la posibilidad de que haya distintos tipos de individuos con diferentes reglas de reproducción. En este modelo se consideran variables aleatorias independientes $(L_{ij}^{n,l})$ a valores en $\mathbb{N}$ que se utilizan para determinar las cantidad de hijos de tipo $j$ que tiene el $l$-ésimo individuo de tipo $i$ en la generación $n$. El proceso queda determinado entonces por la condición inicial y la ecuación

$$Y_j(n+1) = \sum_i \sum_{l=1}^{Y_i(n)} L_{ij}^{n,l}.$$

La variable $Y_j(n)$ representa la cantidad de individuos de tipo $j$ en la generación $n$ y se asume que para cada $i, j$ fijo, las variables $(L_{ij}^{n,l})_{n,l}$ son idénticamente distribuidas. Cuando el espacio de tipos $\mathcal{S}$ es finito el comportamiento del proceso cuando $n \to \infty$ es bien conocido [AN04] y está determinado, a primer orden, por el primer autovalor de la matriz de medias $M = (m_{ij})$ dada por $m_{ij} = \mathbb{E}(L_{ij})$. Si éste es menor o igual que uno la población se extingue en tiempo finito con probabilidad uno, mientas que si es mayor que uno, con probabilidad positiva no se extingue y crece exponencialmente. En ese caso se tiene que cuando $n \to \infty$

$$\frac{Y_j(n)}{\sum_j Y_j(n)} \xrightarrow{cs} \nu_j,$$

donde el vector $\nu = (\nu_j)$ es el autovector a izquierda asociado al primer autovalor de $M$.

Cuando el espacio de tipos es infinito este resultado deja de ser cierto con toda generalidad y es de interés saber bajo qué condiciones se puede obtener este mismo tipo de convergencia. En esta tesis consideramos el caso en que el espacio de tipos es numerable y la matriz $M$ es $R$-positiva y probamos bajo condiciones adecuadas que se puede obtener este mismo tipo de límite. Tanto para esta versión del proceso como para la versión a tiempo continuo que presenta varias dificultades técnicas. Estos resultados son fundamentales para tratar dos problemas distintos. Por un lado los usaremos para probar el comportamiento asintótico de procesos de urnas con bolas de infinitos colores y por otro para probar la convergencia de un método de simulación introducido por Aldous, Flannery y Palacios [AFP88] para aproximar distribuciones cuasiestacionarias.

El estudio de las distribuciones cuasiestacionarias (QSD) comenzó con los trabajos de Kolmogorov, Yaglom y Sevastyanov [Kol38, Yag47, Sev51]. Dada una cadena de Markov en espacio de estados $\mathcal{S} \cup \{0\}$ una distribución cuasiestacionaria $\nu$ es una medida de probabilidad en $\mathcal{S}$ que satisface que cuando la cadena empieza con distribución $\nu$ la distribución del proceso a tiempo $n$ condicionada a no haber tocado 0 sigue siendo $\nu$.

En el artículo fundacional de Yaglom [Yag47] se estudian QSD para procesos de Galton-Watson subcríticos. Desde entonces, se ha estudiado la existencia, unicidad y otras propiedades de las distribuciones cuasiestacionarias para diversos procesos, [DS65, SVJ66, Cav78, ST85, FMP91, FKMP95, FKM96, GJ13, MV12, AFGJ16, CV16].

El cálculo de las distribuciones cuasiestacionarias involucra la resolución de un problema de autovalores y debido a la complejidad de este problema cuando el espacio de estados es muy grande o incluso infinito, se han desarrollado diversos métodos de simulación y aproximación [FV79, AFP88, BH00].

La idea central del método de AFP es construir un proceso no condicionado que sirva para aproximar distribuciones cuasiestacionarias de cadenas de Markov. Esto se hace de la siguiente manera: se corre la cadena de Markov hasta que llega al estado 0. En ese momento se sortea uno de los estados de $\mathcal{S}$ con probabilidades proporcionales a la cantidad de veces que la cadena ha pasado por cada estado (denominamos a esta distribución medida empírica). Se vuelve a correr la cadena de Markov inicializada en el estado sorteado y de forma independiente hasta llegar a 0 nuevamente. Luego se vuelve a sortear un estado pero considerando ambas cadenas.

Se repite el procedimiento teniendo en cuenta que se sorteará el estado inicial de cada cadena de Markov considerando todas las cadenas anteriores. Se crea un nuevo proceso sin el 0 concatenando las cadenas generadas. Este nuevo proceso puede ser identificado con un proceso de ramificación multitipo a tiempo continuo. Partiendo del comportamiento asintótico de este proceso de ramificación, se obtiene que la medida empírica del nuevo proceso converge a la QSD minimal de la cadena de Markov original.

Inspirados en el método de AFP hemos desarrollado otro método que involucra procesos de ramificación multitipo a tiempo discreto. Este nuevo método consiste en correr una cadena de Markov hasta llegar al 0 y luego, en vez de sortear el nuevo estado inicial, se impone que todos los estados por los que pasó la cadena se reproduzcan en simultáneo. Esto da lugar a un proceso de ramificación multitipo a tiempo discreto.

Por un lado, este nuevo método tiene la ventaja de que puede ser fácilmente paralelizado y por otro lado hemos podido probar la convergencia del mismo a la QSD minimal para una familia de procesos mucho más grande que en el caso de AFP. Este familia incluye a los procesos de Galton-Watson para los cuales no hemos podido probar la convergencia de AFP con total generalidad.

En otra dirección, partiendo de las ideas de Ferrari y Rolla [FR15], hemos desarrollado un método para aproximar QSD en espacios numerables por QSD en espacios finitos. La idea central es utilizar dominación estocástica, que es un orden en el espacio de medidas, para construir una sucesión monótona creciente de QSD en espacios finitos y convergente a la QSD del proceso definido en espacio de estados numerable. Este método lo aplicamos a procesos de nacimiento y muerte cuyas QSD son conocidas pero cabe destacar que el método puede ser aplicado a procesos de nacimiento y muerte más generales.

**Organización de la Tesis.** En el Capítulo 1 tratamos procesos de ramificación. Introducimos y damos propiedades del proceso de Galton-Watson y procesos de ramificación multitipo en espacio numerables. Lo construimos y probamos sus principales propiedades para tiempo discreto y continuo. Nos enfocamos en el comportamiento asintótico del proceso cuando este no se extingue. En particular damos condiciones bajo las cuales las proporciones de individuos de cada tipo convergen a una medida de probabilidad.

En el Capítulo 2 describimos los procesos de urna y mostramos la relación con los procesos de ramificación. Probamos el comportamiento asintótico de la urna cuando el conjunto de los posibles colores de las bolas es numerable bajo hipótesis adecuadas.

En el Capítulo 3 introducimos las distribuciones cuasiestacionarias, el tema central de esta tesis. Probamos que si la matriz de probabilidades de transición cumple una condición de Döeblin entonces el proceso es $R$-positivo, lo que implica la existencia de al menos una QSD y posibilita además la aplicación de los métodos desarrollados en el Capítulo 1 para probar la convergencia de los métodos de aproximación y simulación de la misma.

En el Capítulo 4 nos centramos en la simulación de QSD. Definimos el método de AFP, y probamos la convergencia del mismo para procesos definidos en espacios de estados numerables bajo la condición de Döeblin definida en el Capítulo 3. Además damos un nuevo método similar al de AFP, pero el cual involucra procesos de ramificación multitipo a tiempo discreto. Damos condiciones bajo las cuales el método converge y probamos que se puede aplicar este método a procesos de Galton-Watson subcríticos.

En el Capítulo 5 aproximamos la QSD minimal de procesos de nacimiento y muerte en espacios de estados numerable mediante las QSD de procesos de nacimiento y muerte en espacios finitos. Además probamos que esta convergencia es monótona.

# Capítulo 1

# Procesos de ramificación

El origen de las investigaciones sobre procesos de ramificación se atribuye a los estudios realizados por Bienaymé en Francia [Bie74] y por Galton y Watson en Inglaterra [WG75]. Estas surgieron por la preocupación en la extinción de los apellidos aristocráticos. De estos estudios surgió el proceso de ramificación llamado proceso de Galton-Watson.

Los procesos de ramificación ofrecen un modelo matemático para el estudio de poblaciones en las que los individuos viven, se reproducen y mueren independientemente unos de otros. Dentro de la teoría de procesos de ramificación, hay diferentes modelos. El más simple es el llamado Galton-Watson donde se comienza con un individuo, el cual constituye la generación cero, sus hijos forman parte de la primera generación, sus nietos de la segunda, y así sucesivamente. Todos los individuos pertenecientes a una generación, se mueren y se reproducen en simultáneo dando lugar a la siguiente generación.

En lugar de suponer que todos los individuos mueren y se reproducen en simultáneo, es posible considerar un proceso similar al de Galton-Watson, en el cual el tiempo de vida y la cantidad de descendientes de cada individuo son aleatorios e independientes entre sí.

Una posible generalización de los procesos antes mencionados son los procesos de ramificación multitipo en los cuales se consideran distintos tipos de individuos. Estos se pueden separar según si el conjunto de los distintos tipos de individuos es finito, numerable o no numerable. Al igual que en el caso simple, se puede suponer que el tiempo de vida de cada individuo es fijo (tiempo discreto) o que tienen un tiempo aleatorio de vida asociado (tiempo continuo). En general, hay varias condiciones que se pueden considerar de un proceso de ramificación y de la manera en que evoluciona. En nuestro caso, supondremos que cada individuo es independiente de los otros individuos y de la historia del proceso. Además cuando consideramos el caso continuo, supondremos que el tiempo de vida de cada individuo tiene distribución exponencial, lo que resultará en un proceso de Markov.

Una de las preguntas básicas que se plantean en el estudio de los procesos de ramificación es si la población se extingue o sobrevive con probabilidad positiva. Cuando la población sobrevive, interesa conocer la proporción de individuos de cada tipo.

A continuación mencionaremos algunos trabajos relevantes para nuestro estudio. En el caso de procesos de ramificación con finitos tipos de individuos, destacamos los libros [AN04, Har63]. Para estudiar comportamientos asintóticos de la población es utilizada la teoría de Perron-Frobenius sobre matrices positivas [Sen06]. El caso de tiempo discreto puede verse en [KLPP97] y el caso de tiempo continuo en [GB03]. Otra de las herramientas muy utilizadas para conocer propiedades de los procesos es la función generadora de momentos, [AN04, Har63]. Por ejemplo, cuando el proceso multitipo finito no se extingue la función generadora tiene solo dos puntos fijos, uno de los cuales representa la probabilidad de extinción.

La bibliografía sobre el caso multitipo con numerable tipos de individuos es menor a la del caso finito. Solo algunos de los resultados han sido extendidos y otros permanecen abiertos. Típicamente se asumen hipótesis de $R$-positividad sobre la matriz de medias. En el trabajo [VJ67] se estudian las propiedades de las matrices $R$-positivas; la $R$-positividad de una matriz es una condición que garantiza que el comportamiento asintótico de las potencias de la misma es similar al del caso finito dado por la teoría de Perron-Frobenius. En [Moy67] se estudia el comportamiento asintótico del proceso cuando este no se extingue bajo hipótesis de $R$-positividad sobre la matriz de medias.

Un ejemplo en donde la generalización del caso finito al caso numerable falla es en el estudio de los puntos fijos de la función generadora de momentos. Como hemos mencionado en el caso finito, si la población no se extingue, la función tiene solo dos puntos fijos, pero en el caso numerable, esto es falso en general [BZ14]. Pueden verse contraejemplos en [BZ17].

El resultado principal de este capítulo es el Teorema 1.35 sobre procesos de ramificación multitipo a tiempo continuo para numerables tipos de individuos. El teorema dice que, cuando el proceso no se extingue, la proporción de individuos de cada tipo converge en probabilidad a una medida de probabilidad en $\mathbb{N}$. Más aún, se caracteriza dicha medida.

El esquema del capítulo es el siguiente. Comienza con el proceso de ramificación llamado Galton-Watson, se dan varias propiedades, se define la función generadora y se analiza el comportamiento del proceso a tiempos grandes. Luego se estudian procesos de ramificación multitipo con numerables tipos de individuos, separándolos en tiempo discreto y continuo.

A tiempo discreto se analiza la función generadora, las propiedades de la matriz de medias, el teorema de convergencia de [Moy67] y la probabilidad de extinción entre otros. En la Proposición 1.29 se muestra, en el caso en el que el proceso no se extingue, la convergencia en probabilidad de la proporción de individuos de cada tipo al autovector a izquierda asociado a la matriz de medias (Subsección 1.2.2). Luego se da la construcción del proceso de ramificación multitipo a tiempo continuo, se calcula la media de la cantidad de individuos de cada tipo, se estudia el proceso discretizado en el tiempo para luego generalizar al caso continuo la Proposición 1.29 dando lugar al Teorema 1.35.

## 1.1. Procesos de Galton-Watson

Los resultados principales sobre procesos de Galton-Watson pueden verse en [AN04, Capítulo 1]. Daremos las demostraciones de los resultados que consideramos pertinentes para el resto del capítulo.

El proceso de Galton-Watson es un proceso de Markov a tiempo discreto que representa la cantidad de individuos de las generaciones sucesivas. Denotamos $Y(n)$ a la variable aleatoria que representa la cantidad de individuos en la generación $n$-ésima. Suponemos que el número de hijos que tiene un individuo es una variable aleatoria con distribución $(p_i)_{i\in\mathbb{N}_0}$, donde $p_i$ es la probabilidad de que un individuo deje $i$ descendientes, con $i \in \mathbb{N}_0$. Para cada $j \in \mathbb{N},\ k \in \mathbb{N}_0,\ L_j^k$ representa el número de hijos del $j$-ésimo miembro de la $k$-ésima generación. Podemos expresar a $Y(n+1)$ a partir de $\{L_j^n,\ j \geq 1\}$ e $Y(n)$,

$$Y(n+1) = L_1^n + .. + L_{Y(n)}^n.$$

Asumimos que $\{L_j^k,\ k \geq 0,\ j \geq 1\}$ son variables aleatorias i.i.d., donde $\mathbb{P}(L_1^1 = i) = p_i \quad \forall i \in \mathbb{N}_0$. El proceso construido cumple $\mathbb{E}(s^{Y(n+1)}|Y(n) = i) = \left[\mathbb{E}(s^{L_1^1})\right]^i$.

Si denotamos

$$p_{ij} := \mathbb{P}(Y(n+1) = j|Y(n) = i) = \mathbb{P}\big(\sum_{l=1}^{i} L_l^n = j\big),$$

el proceso verifica

$$\sum_{j\in\mathbb{N}_0} p_{ij}s^j = \big[\sum_{j\in\mathbb{N}_0} p_j s^j\big]^i.$$

**Definición 1.1.** *Una cadena de Markov $\{Y(n), n \in \mathbb{N}_0\}$ es un Proceso de Galton-Watson si las probabilidades de transición cumplen*

$$\sum_{j\in\mathbb{N}_0} p_{ij}s^j = \big[\sum_{j\in\mathbb{N}_0} p_j s^j\big]^i,$$

*para todo $s \in [0,1]$, donde $p_{ij} = \mathbb{P}(Y(n+1) = j|Y(n) = i)$ para $i, j \in \mathbb{N}_0$.*

Notemos que la definición da una propiedad aditiva sobre el proceso, significa que si $Y(n) = i$ entonces la distribución $Y(n+1)$ tiene la distribución de la suma de $i$ variables aleatorias independientes, todas con distribución $(p_i)_{i\in\mathbb{N}_0}$. Si en la generación $n$ no hay individuos, tampoco habrá en las generaciones siguientes, es decir $\mathbb{P}(Y(n+1) = 0|Y(n) = 0) = 1$.

Si el proceso comienza con $i$ individuos, denotamos $Y^{(i)}(n),\ n \in \mathbb{N}$. Este proceso es la suma de $i$ procesos de ramificación independientes que se inician con un solo individuo. Para evitar situaciones triviales, vamos a suponer que $p_0+p_1 < 1$. Es decir, la probabilidad de tener a lo sumo un individuo es menor que 1.

Estamos interesados en estudiar si la población perdura en el tiempo. Para esto tendríamos que conocer la distribución de $Y(n)$, para así poder calcular $\mathbb{P}(Y(n) = 0)$ para $n$ suficientemente grande. Observemos que la distribución de la cantidad de individuos en la generación $n$-ésima depende de las generaciones anteriores.

Una función que será muy útil en el estudio de procesos de ramificación es la función generadora de momentos de la variable $Y(1)$.

Sea $f : [0,1] \to \mathbb{R}_+$ definida por $f(s) := \mathbb{E}(s^{Y(1)}) = \mathbb{E}(s^{L_1^1})$. Consideramos también sus composiciones,

$$f_0(s) = s \qquad f_1(s) = f(s) \tag{1.1}$$

$$f_{n+1}(s) = f(f_n(s)), \quad n \in \mathbb{N} \tag{1.2}$$

**Proposición 1.2.** *La función generadora de $Y(n)$ es $f_n(s)$, es decir, componer $n$ veces la función generadora de $Y(1)$.*

*Demostración.* Sea $f_{(n)}$ la función generadora asociada a $Y(n)$:

$$f_{(n)}(s) = \sum_{k \in \mathbb{N}_0} \mathbb{P}(Y(n) = k)s^k.$$

Como $f(s) = f_{(1)}(s)$, basta probar que la sucesión de funciones $f_{(n)}$ cumple una relación de recurrencia como (1.2). Por como la definimos,

$$f_{(n+1)}(s) = \mathbb{E}(s^{Y(n+1)}) = \mathbb{E}(s^{L_1^n + .. + L_{Y(n)}^n}) = \sum_{k \in \mathbb{N}_0} \mathbb{E}(\mathbf{1}\{Y(n) = k\} s^{L_1^n + .. + L_k^n}).$$

La variable $Y(n)$ es independiente de las variables $L_j^n, j \geq 1$, por lo tanto,

$$\mathbb{E}(\mathbf{1}\{Y(n) = k\} s^{L_1^n + .. + L_k^n}) = \mathbb{P}(Y(n) = k)\mathbb{E}(s^{L_1^n + .. + L_k^n}).$$

Como las variables, $L_1^n, .., L_k^n$ son i.i.d.

$$\mathbb{E}(s^{L_1^n + .. + L_k^n}) = \prod_{j=1}^{k} \mathbb{E}(s^{L_j^n}) = (\mathbb{E}(s^{L_1^1}))^k = (f(s))^k.$$

Luego,

$$f_{(n+1)}(s) = \sum_{k \in \mathbb{N}_0} \mathbb{P}(Y(n) = k)(f(s))^k = f_{(n)}(f(s)).$$

□

Estas funciones nos van a ayudar para calcular los momentos del proceso. Cuando estos existan, los podemos expresar en términos de las derivadas de $f_n(s)$ para $s = 1$.

**Teorema 1.3.** *Para el proceso de ramificación ya definido, si $\mathbb{E}(Y(1)) = m$, entonces $\mathbb{E}(Y(n)) = m^n$.*

*Demostración.* Para la media de la primera generación tenemos,

$$\mathbb{E}(Y(1)) = \sum_{j \in \mathbb{N}_0} p_j j = f'(1) = m.$$

Para la generación $n$-ésima,

$$\begin{aligned}\mathbb{E}(Y(n)) &= \sum_{j\in\mathbb{N}_0} \mathbb{P}(Y(n)=j)j = f_n'(1) = (f_{n-1}(f(1)))' = f_{n-1}'(f(1))f'(1)\\ &= f_{n-1}'(1)f'(1) = ... = (f'(1))^n = m^n.\end{aligned}$$

□

Queremos analizar el comportamiento del proceso, para ello necesitamos algunas propiedades de la funciones generadoras. Luego, veremos que de lo único que depende que el proceso se extinga es de la media.

**Lema 1.4.** *La función $f$ cumple las siguientes propiedades:*

1. *$f$ es estrictamente creciente y convexa.*
2. *$f(0) = p_0 \qquad f(1) = 1 \qquad f'(1) = \mathbb{E}(L) \qquad f''(1) = \mathbb{E}(L(L-1))$.*
3. *Si $m \leq 1$ entonces $f(s) > s$, para todo $s \in [0,1)$.*
4. *Si $m > 1$ entonces $f(s) = s$ tiene una única solución en $[0,1)$.*

*Demostración.* 1. Veamos las derivadas de $f$,
$f(s) = \sum_{k=0}^{\infty} p_k s^k$
$f'(s) = \sum_{k=1}^{\infty} p_k k s^{k-1} > 0 \qquad \forall s \in (0,1)$
$f''(s) = \sum_{k=2}^{\infty} p_k k(k-1)s^{k-2} > 0 \qquad \forall s \in (0,1)$ (pues supusimos $p_0 + p_1 < 1$).
Por lo tanto $f$ es estrictamente creciente y convexa.

2. Es inmediato.

3. Consideremos la función

$$g : [0,1] \to \mathbb{R} \qquad g(s) = f(s) - s.$$

Sus derivadas, $g'(s) = f'(s) - 1$, $g''(s) = f''(s)$. Como $f$ es estrictamente convexa en $[0,1]$ y $f(1) = 1$, la función $g(s)$ también es estrictamente convexa y $g(1) = 0$. Luego la derivada $g'(s)$ es una función estrictamente creciente, y si suponemos $m \leq 1$, $g'(1) = f'(1) - 1 = m - 1 \leq 0$, por lo tanto $g$ es estrictamente decreciente en $[0,1]$. Como $g(1) = f(1) - 1 = 0$, $g(s) > 0$ para $s \in [0,1)$ y $f(s) > s$ en $[0,1)$.

4. Sea $m > 1$, y $g$ la función definida anteriormente. Como $g''(s) > 0$ $\forall s \in (0,1)$, $g$ es convexa en $(0,1)$ y $g'$ es estrictamente creciente, por lo tanto $g$ tiene a lo sumo un punto crítico en $(0,1)$. Como $g'(1) = m - 1 > 0$ y $g(1) = f(1) - 1 = 0$, existe $s_0 < 1$ con $g(s_0) < 0$. Por otro lado, $g(0) = p_0 > 0$. Luego existe $q \in (0,1)$ tal que $g(q) = 0$, es decir, $f(q) = q$.

□

**Teorema 1.5.** *Sea $(Y(n))_{n\geq 0}$ un proceso de Galton-Watson. Llamamos $q$ a la probabilidad de que el proceso se extinga, es decir*

$$\mathbb{P}\left(\bigcup_{n\in\mathbb{N}}\{Y(n)=0\}\right)=q.$$

*Entonces $q$ es el punto fijo más chico de la función generadora.*

*Demostración.* Primero veamos que, $f_n(0)\nearrow q$, cuando $n\to\infty$. Observemos que $f(0)>0$, y $f$ es creciente, por lo tanto, $0<f(0)<f(q)=q$. Aplicando nuevamente $f$, $0<f(0)<f_2(0)<q$. Repitiendo este procedimiento, obtenemos que $f_n(0)$ es creciente y acotada. Sea $L=\lim_{n\to\infty} f_n(0)$, $L\leq q$. Como $f$ es continua,

$$f(L)=f\big(\lim_{n\to\infty} f_n(0)\big)=\lim_{n\to\infty} f_{n+1}(0)=L$$

Luego, $L=q$. Como $\{Y(n)=0\}\subseteq\{Y(n+1)=0\}$ y $f_n(0)=\mathbb{P}(Y(n)=0)$ tenemos

$$\mathbb{P}\left(\bigcup_{n\in\mathbb{N}}\{Y(n)=0\}\right)=\lim_{n\to\infty}\mathbb{P}(Y(n)=0)=\lim_{n\to\infty} f_n(0)=q.$$

□

Del teorema se obtiene que si $m\leq 1$ el proceso se extingue en tiempo finito con probabilidad 1. Cuando $m<1$ se lo llama caso subcrítico, y cuando $m=1$ caso crítico. Por otro lado, si $m>1$ (caso supercrítico) hay una probabilidad positiva de que los descendientes permanezcan en todas las generaciones futuras. Veremos que en este caso, no solo hay probabilidad positiva de no extinguirse, sino que además el tamaño de la población crece exponencialmente. Para ello debemos introducir el concepto de martingala.

**Definición 1.6.** *El proceso estocástico $(X(n))_{n\in\mathbb{N}}$ a valores en $\mathbb{R}$ es una martingala respecto de la filtración $(\mathcal{F}_n)_{n\in\mathbb{N}}$ si, para todo $n\geq 0$*

1. $\mathbb{E}(|X(n)|)<\infty$
2. $X(n)\in\mathcal{F}_n$
3. $\mathbb{E}(X(n+1)|\mathcal{F}_n)=X(n)$.

*Para abreviar, decimos que $X(n)$ es $\mathcal{F}_n$-martingala.*

**Observación 1.7.** *Si $X(n)$ es una martingala, $\mathbb{E}(X(n))=\mathbb{E}(X(0))$ para todo $n\in\mathbb{N}$.*

En la siguiente proposición se dan condiciones bajo las cuales las martingalas convergen.

**Proposición 1.8.** *[Kal02, Proposición 6.19] Sea $(X(n))_{n\in\mathbb{N}}$ una $\mathcal{F}_n$-martingala, tal que $\sup_{n\in\mathbb{N}}\{\mathbb{E}(|X(n)|)\}<\infty$. Entonces existe $\lim\limits_{n\to\infty} X(n)$ casi seguramente.*

A continuación damos una martingala asociada al proceso de Galton-Watson.

**Teorema 1.9.** *[AN04, Teorema 1, p. 9] Sea* $\mathcal{F}_n = \sigma(L_i^k,\ i \geq 1,\ 1 \leq k \leq n)$, *y* $m = \mathbb{E}(L_i^k)$, *entonces* $W(n) = \frac{Y(n)}{m^n}$ *es una* $\mathcal{F}_n$*-martingala.*

*Demostración.* Claramente $W(n) \in \mathcal{F}_n$. Como $m^n$ es una constante,

$$\mathbb{E}(W_{n+1}|\mathcal{F}_n) = \frac{1}{m^n}\mathbb{E}(Y(n+1)|\mathcal{F}_n).$$

Además

$$\mathbb{E}(Y(n+1)|\mathcal{F}_n) = \mathbb{E}\left(\sum_{i\in\mathbb{N}} \mathbf{1}\{i \leq Y(n)\}L_1^n|\mathcal{F}_n\right).$$

La suma es de variables positivas, por lo tanto

$$\mathbb{E}(Y(n+1)|\mathcal{F}_n) = \sum_{i\in\mathbb{N}} \mathbb{E}\left(\mathbf{1}\{i \leq Y(n)\}L_1^n|\mathcal{F}_n\right).$$

Como $\mathbf{1}\{i \leq Y(n)\} \in \mathcal{F}_n$ y $L_1^n$ es independiente de $\mathcal{F}_n$,

$$\mathbb{E}(Y(n+1)|\mathcal{F}_n) = \sum_{i\in\mathbb{N}} \mathbf{1}\{i \leq Y(n)\}\mathbb{E}(L_1^n) = mY(n).$$

Luego,

$$\mathbb{E}(W(n+1)|\mathcal{F}_n) = \frac{1}{m^{n+1}}\mathbb{E}(Y(n+1)|\mathcal{F}_n) = \frac{1}{m^{n+1}}mY(n) = W(n).$$

Como $\mathbb{E}(W(0)) = 1$, $\mathbb{E}(W(n)) = 1$ para todo $n$. □

**Observación 1.10.** *Existe una variable aleatoria* $W$*, tal que cuando* $n \to \infty$

$$\frac{Y(n)}{m^n} \xrightarrow{cs} W.$$

De esta convergencia se ve que $Y(n)$ crece como $m^n W$, si $W \neq 0$. Si $\mathbb{P}(W = 0) = 1$ nos dice simplemente que $m^n$ crece mas rápido que $Y(n)$. El siguiente teorema muestra que $\mathbb{P}(W = 0)$ es exactamente la probabilidad de extinción del proceso.

**Teorema 1.11.** *[AN04, Teorema 2, p. 9] Si* $m > 1$, $V(Y(1)) = \sigma^2 < \infty$ *e* $Y(0) = 1$, *entonces*

1. $\lim_{n\to\infty} \mathbb{E}(W(n) - W)^2 = 0$,
2. $\mathbb{E}(W) = 1$,
3. $\mathbb{P}(W = 0) = \mathbb{P}(Y(n) = 0$ *para algún* $n) = q$.

## 1.2. Procesos de ramificación multitipo a tiempo discreto

Los procesos de ramificación multitipo son una generalización del modelo de Galton-Watson, en el cual, se mantiene la independencia en la probabilidad de generar nuevos individuos y se introduce la posibilidad de que haya diferentes tipos de individuos que coexistan en la población. Las diferencias entre los distintos tipos quedarán puestas de manifiesto en las pautas de reproducción de los mismos. Para analizar procesos de ramificación multitipo es necesario introducir notación, definiciones y algunos resultados.

### 1.2.1. Construcción del proceso

Daremos una construcción del proceso similar a la que se puede encontrar, a tiempo continuo, en [GB03].

Sea $\mathcal{S}$ el conjunto de los distintos tipos de individuos. Suponemos que $\mathcal{S}$ es un conjunto numerable. Para $i, j \in \mathcal{S}$ la variable aleatoria $L_{ij}$ representa la cantidad de hijos de tipo $j$ que tiene un individuo de tipo $i$. Para cada $\mathbf{k} = (k_1, k_2, ..)$, $\mathbf{k} \in \mathbb{N}_0^{\mathcal{S}}$ sea $p_{i\mathbf{k}} = \mathbb{P}(L_{ij} = k_j, \forall j \in \mathcal{S})$, es decir la probabilidad de que un individuo de tipo $i$ tenga $k_1$ individuos de tipo 1, $k_2$ individuos de tipo 2, etc. Notamos $\mathbf{p}_i = \{p_{i\mathbf{k}}, \mathbf{k} \in \mathbb{N}_0^{\mathcal{S}}\}$ a la distribución de la cantidad de hijos de cada tipo que tiene un individuo de tipo $i$.

Construimos el árbol genealógico. Consideramos el espacio de estados

$$\mathbb{X} = \bigcup_{n\in\mathbb{N}} \mathbb{X}^n,$$

donde $\mathbb{X}^n$ es el espacio donde vive la generación $n$-ésima. Esto es $\mathbb{X}^0 = \mathcal{S}$ y $i_0 \in \mathbb{X}$ es la raíz del árbol. El siguiente es $\mathbb{X}^1 = \mathcal{S} \times \mathbb{N}$ y el elemento $x = (i_1, l_1) \in \mathbb{X}^1$ es el $l_1$-ésimo hijo de tipo $i_1$ de la raíz. Finalmente, para $n > 1$, $\mathbb{X}^n = \mathcal{S}^n \times \mathbb{N}^n$ y $x = (i_1, .., i_n, l_1, .., l_n) \in \mathbb{X}^n$ es $l_n$-ésimo hijo de tipo $i_n$ del padre $\hat{x} = (i_1, .., i_{n-1}, l_1, .., l_{n-1})$. Escribimos $\sigma(x) = i_n$ para el tipo de $x \in \mathbb{X}^n$. Observemos que $\sigma(x) \in \mathcal{S}$.

A cada $x \in \mathbb{X}$ le asociamos descendencia aleatoria, $\mathbf{L}_{\sigma(x)} = (L_{\sigma(x)j})_{j\in\mathcal{S}} \in \mathbb{N}_0^{\mathcal{S}}$ con distribución $\mathbf{p}_{\sigma(x)}$ tal que las variables $\{\mathbf{L}_{\sigma(x)} : x \in \mathbb{X}\}$ son independientes. Las variables aleatorias $\mathbf{L}_{\sigma(x)}$ indican como son actualizados los individuos $x \in \mathbb{X}$. Por lo tanto el conjunto $X = \bigcup_{n\in\mathbb{N}} X^n$ esta definido recursivamente por

$$X^0 = \{i_0\} \qquad X^n = \{x = (\hat{x}, i_n, l_n) \in \mathbb{X}^n : \hat{x} \in \mathbb{X}^{n-1}, l_n \leq L_{\sigma(\hat{x}),i_n}\}.$$

El árbol completo está determinado por el proceso $X$. Cuando comenzamos con una distribución inicial $\rho$, la distribución de $X^n$ es $\mathbb{P}_\rho$ y la esperanza es $\mathbb{E}_\rho$, cuando $i_0 = i$ los denotamos $\mathbb{P}_i$ y $\mathbb{E}_i$.

Consideremos

$$\mathbf{Y}(n) := \sum_{x\in X^n} \delta_{\sigma(x)}. \tag{1.3}$$

Es decir, el vector $\mathbf{Y}(n) = (Y_1(n), \ldots, Y_k(n), \ldots)$ describe cuántos individuos de cada tipo hay en la $n$-ésima generación. En particular, $Y_i(n)$ es el cardinal de $X_i^n = \{x \in X^n :$

$\sigma(x) = i\}$. Notemos que

$$\mathbf{Y}(n+1) = \sum_{i \in \mathcal{S}} \sum_{l=1}^{Y_i(n)} \mathbf{L}_i^{n,l}, \tag{1.4}$$

donde $\mathbf{L}_i^{n,l}$ denota el vector que indica la cantidad de descendientes del $l$-ésimo individuo de tipo $i$ en la generación $n$, su distribución no depende ni de $l$ ni de $n$. De la ecuación se desprende que el proceso es de Markov.

Introducimos la siguiente notación para el producto $\mathbf{s}^{\mathbf{i}} = \prod_{j \in \mathcal{S}} s_j^{i_j}$ con $\mathbf{s} = (s_1, s_2, \ldots) \in [0,1]^{\mathcal{S}}$ e $\mathbf{i} \in \mathbb{N}^{\mathcal{S}}$.

Observemos que por la independencia de las variables aleatorias $\mathbf{L}_k$ se verifica

$$\mathbb{E}(\mathbf{s}^{\mathbf{Y}(n+1)} | \mathbf{Y}(n) = \mathbf{i}) = \prod_{k \in \mathcal{S}} \left(\mathbb{E}(\mathbf{s}^{\mathbf{L}_k})\right)^{i_k}.$$

Que es equivalente a

$$\sum_{\mathbf{j} \in \mathbb{N}^{\mathcal{S}}} p_{\mathbf{i},\mathbf{j}} \mathbf{s}^{\mathbf{j}} = \prod_{k \in \mathcal{S}} [\sum_{\mathbf{j} \in \mathbb{N}^{\mathcal{S}}} p_{k,\mathbf{j}} \mathbf{s}^{\mathbf{j}}]^{i_k}, \tag{1.5}$$

donde $p_{\mathbf{ij}} := \mathbb{P}(\mathbf{Y}(n+1) = \mathbf{j} | \mathbf{Y}(n) = \mathbf{i})$.

**Definición 1.12.** *Una cadena de Markov $\{\mathbf{Y}(n), n \in \mathbb{N}_0\}$ es un proceso de ramificación multitipo ($\mathcal{S}-$tipos) si las probabilidades de transición cumplen*

$$\sum_{\mathbf{j} \in \mathbb{N}^{\mathcal{S}}} p_{\mathbf{i},\mathbf{j}} \mathbf{s}^{\mathbf{j}} = \prod_{k \in \mathcal{S}} [\sum_{\mathbf{j} \in \mathbb{N}^{\mathcal{S}}} p_{k,\mathbf{j}} \mathbf{s}^{\mathbf{j}}]^{i_k},$$

*para todo $\mathbf{s} \in [0,1]^{\mathcal{S}}$, donde $p_{\mathbf{ij}} := \mathbb{P}(\mathbf{Y}(n+1) = \mathbf{j} | \mathbf{Y}(n) = \mathbf{i})$ para $\mathbf{i}, \mathbf{j} \in \mathbb{N}^{\mathcal{S}}$.*

Esta es una propiedad aditiva, significa que si $\mathbf{Y}(n) = \mathbf{i}$, entonces $\mathbf{Y}(n+1)$ tiene la distribución de la suma de $i_1 + i_2 + ...$ variables aleatorias independientes, que toman valores en $\mathcal{S}$, de las cuales $i_1$ tienen distribución $p_1$, $i_2$ tienen distribución $p_2$, etc. Observemos que si en la generación $n$ no hay individuos, tampoco habrá en las generaciones siguientes, es decir $\mathbb{P}(\mathbf{Y}(n+1) = \mathbf{0} | \mathbf{Y}(n) = \mathbf{0}) = 1$ para todo $n \in \mathbb{N}$, donde denotamos con $\mathbf{0} := (0, 0, \ldots)$.

**Observación 1.13.** *Consideremos $\mathbf{Y}$ un proceso de Markov que verifica la Definición 1.12.*

- *Si hay un solo posible tipo de individuo, es decir el cardinal de $\mathcal{S}$ es 1, entonces $\mathbf{Y}$ verifica la Definición 1.1.*
- *Todo proceso de ramificación $\mathbf{Y}$ puede ser construido como en la ecuación 1.4.*

Cuando el proceso comienza con $\mathbf{k}$ individuos, notamos $Y_j^{(\mathbf{k})}(n)$ a la cantidad de individuos de tipo $j$ que hay en la $n$-ésima generación. Por simplicidad en la notación, cuando no sea esencial mostrar la dependencia en $\mathbf{k}$ omitiremos el superíndice.

Al igual que en proceso de Galton-Watson vamos a trabajar con la función generadora de momentos.

**Definición 1.14.** *Si el proceso empieza con un individuo de tipo $i$, la función generadora de momentos de $\mathbf{Y}(n)$ viene dada por:*

$$f_{ni}(\mathbf{s}) := \mathbb{E}_i(\mathbf{s}^{\mathbf{Y}(n)}) = \mathbb{E}_i(\prod_{j\in\mathcal{S}} s_j^{Y_j(n)}),$$

*con $\mathbf{s} = (s_1, s_2, \ldots)$, y $s_j \in [0,1]$ para todo $j \in \mathcal{S}$. Denotamos*

$$\mathbf{f}_n(\mathbf{s}) = (f_{n1}(\mathbf{s}), f_{n2}(\mathbf{s}), \ldots)$$

*y definimos $f_{ni}(\mathbf{0}) = p_{i\mathbf{0}}$.*

Para simplificar la notación, cuando nos estamos refiriendo a la función generadora de $\mathbf{Y}(1)$, usaremos $f_i(\mathbf{s}) = f_{1i}(\mathbf{s})$.

Observemos que las funciones generadoras de $\mathbf{Y}(n)$ verifican que $\mathbb{E}_{\mathbf{k}}(\mathbf{s}^{\mathbf{Y}(n)}) = (\mathbf{f}_n(\mathbf{s}))^{\mathbf{k}}$ y que $\mathbf{f}_n(\mathbf{s})$ es componer $n$ veces la función generadora de $\mathbf{Y}(1)$. También para cada $i \in \mathcal{S}$, $f_{ni}$ es continua pues $\mathbb{P}_i(|\mathbf{Y}(n)| < \infty) = 1$ y $\mathbb{E}_i(\mathbf{s}^{\mathbf{Y}(n)})$ resulta un polinomio en $\mathbf{s}$. Más aún $\mathbf{f}_n$ es continua respecto a la topología producto [Mun75, Teorema 19.6].

### 1.2.2. Matriz de medias

Sea $m_{ij} := \mathbb{E}(L_{ij})$ el número esperado de descendientes de tipo $j$ que tiene un individuo de tipo $i$. Definimos la matriz de medias

$$M := \{m_{ij} : i, j \in \mathcal{S}\}.$$

Claramente

$$m_{ij} = \frac{\partial f_i(\mathbf{s})}{\partial s_j}|_{\mathbf{s}=\mathbf{1}},$$

donde $\mathbf{1} = (1, 1, \ldots)$. Definimos inductivamente $m_{ij}^{(n)}$, para todo $i, j \in \mathcal{S}$ y $n \in \mathbb{N}$:

$$m_{ij}^{(1)} = m_{ij}, \quad m_{ij}^{(n+1)} = \sum_{k\in\mathcal{S}} m_{ik}^{(n)} m_{kj}.$$

Es decir

$$m_{ij}^{(n)} = (M^n)_{ij}.$$

De donde se deducen las siguientes propiedades:

$$\mathbb{E}(Y_j(k+n)|\mathbf{Y}(k) = \delta_i) = m_{ij}^{(n)}$$

$$\mathbb{E}(Y_j(k+n)|\mathbf{Y}(k) = \mathbf{z}) = (\mathbf{z}M^n)_j = \sum_{i\in\mathcal{S}} z_i m_{ij}^{(n)}.$$

Si $\gamma = (\gamma_1, \gamma_2, \ldots)^t \in \mathbb{R}^{\mathcal{S}\times 1}$ es un vector columna no negativo, entonces

$$\mathbb{E}(\mathbf{Y}(k+n)\gamma|\mathbf{Y}(k) = \mathbf{z}) = \sum_{j\in\mathcal{S}} (\mathbf{z}M^n)_j \gamma_j = \mathbf{z}M^n\gamma. \tag{1.6}$$

**Definición 1.15.** *Dada $M = \{m_{ij} : i, j \in \mathcal{S}\}$ una matriz con entradas no negativas tal que $m_{ij}^{(n)} < \infty$ para todo $i, j \in \mathcal{S}$ y $n \in \mathbb{N}$. Decimos que:*

- *La matriz $M$ es irreducible si para cada par de índices $i, j \in \mathcal{S}$ existe $n \in \mathbb{N}$ tal que $m_{ij}^{(n)} > 0$.*
- *El estado $i \in \mathcal{S}$ es aperiódico si mcd$\{n \in \mathbb{N} / m_{ii}^{(n)} > 0\} = 1$.*
- *La matriz $M$ es aperiódica si para todo $i \in \mathcal{S}$, $i$ es aperiódico.*

**Definición 1.16.** *Dada una matriz $M = \{m_{ij} : i, j \in \mathcal{S}\}$ diremos que $\lambda$ es autovalor con autovector a izquierda $\nu$ si*

$$\nu M = \lambda \nu$$

*es decir para todo $i \in \mathcal{S}$*

$$\sum_{j \in \mathcal{S}} \nu_j m_{ji} = \lambda \nu_i.$$

*Análogamente definimos autovector a derecha $\mu$, como el vector columna que verifica*

$$M\mu = \lambda \mu$$

*es decir para todo $i \in \mathcal{S}$*

$$\sum_{j \in \mathcal{S}} m_{ij} \mu_j = \lambda \mu_i.$$

### 1.2.3. Propiedades ergódicas de la matriz de medias

En [VJ67, Teorema A] se muestra que si la matriz $M$ es no negativa, irreducible, aperiódica y verifica que $m_{ij}^{(n)} < \infty$ para todo $i, j \in \mathcal{S}$ y para todo $n \in \mathbb{N}$, entonces existe

$$\lim_{n \to \infty} (m_{ij}^{(n)})^{\frac{1}{n}}$$

y es el mismo para todo $i, j \in \mathcal{S}$. De donde existe un número $R$ que es el radio de convergencia común de las series $\sum_{n \in \mathbb{N}} m_{ij}^{(n)} s^n$.

**Definición 1.17.** *Sea $M$ una matriz no negativa, irreducible, aperiódica tal que $m_{ij}^{(n)} < \infty$ para todo $i, j \in \mathcal{S}$ y para todo $n \in \mathbb{N}$.*

*Sea $R > 0$, decimos que $M$ es $R$-transiente (respectivamente $R$-recurrente) si para todo $i, j \in \mathcal{S}$ las series $\sum_{n \in \mathbb{N}} m_{ij}^{(n)} R^n$ son todas finitas (respectivamente infinitas).*

*Si $M$ es $R$-recurrente, decimos que es $R$-positiva (respectivamente $R$-nula) si las sucesiones $(m_{ij}^{(n)} R^n)_n$ no tienden a 0 (respectivamente tienden a 0).*

La $R$-positividad de una matriz puede ser caracterizada en términos de sus autovectores como muestran los siguientes teoremas.

**Teorema 1.18.** *[VJ67, Teorema D] Si $M$ es $R$-positiva existen únicos (salvo constantes) autovectores positivos a izquierda y derecha $((\mu_k),(\nu_k))_{k\in\mathcal{S}}$, asociados al autovalor $1/R$ tal que $\sum_{k\in\mathcal{S}}\mu_k\nu_k<\infty$ y se verifica*

$$\lim_{n\to\infty} m_{ij}^{(n)}R^n=\frac{\mu_i\nu_j}{\sum_{k\in\mathcal{S}}\mu_k\nu_k}.$$

Sin pérdida de generalidad podemos suponer $\sum_{k\in\mathcal{S}}\mu_k\nu_k=1$, con lo cual para todo $i,j\in\mathcal{S}$, tenemos que $\lim_{n\to\infty} m_{ij}^{(n)}R^n=\mu_i\nu_j>0$.

**Teorema 1.19.** *[Sen06, Teorema 6.4] Sea $M$ una matriz no nula, irreducible que verifica $m_{ij}^{(n)}<\infty$ para todo $i,j\in\mathcal{S}$ y para todo $n\in\mathbb{N}$. Asumimos que el radio de convergencia $R>0$. Sean $\nu,\mu$ autovectores a izquierda y derecha respectivamente asociados al mismo autovalor $\alpha$ tal que $\sum_{k\in\mathcal{S}}\mu_k\nu_k<\infty$. Entonces $M$ es $R$-positiva y $R=1/\alpha$.*

Los procesos de ramificación multitipo con matriz de transición $R$-positiva los podemos separar en 3 casos.

- Si $R>1$ decimos que el proceso es subcrítico. En este caso, $\lim_{n\to\infty} m_{ij}^{(n)}=0$ lo que significa que en media la población tiende a extinguirse.
- Si $R=1$ es el caso crítico. En este caso obtenemos que $\lim_{n\to\infty} m_{ij}^{(n)}=\mu_i\nu_j>0$.
- Si $R<1$ el proceso es supercrítico, $\lim_{n\to\infty} m_{ij}^{(n)}=\infty$ lo que significa que en media la cantidad de individuos de cada tipo de la población tiende a infinito.

Estamos interesados en el último caso, que es cuando el tamaño de la población en media crece exponencialmente. En el evento en que la población no se extingue nos interesa conocer la proporción de cada tipo en la población para tiempos grandes, es decir estudiar el comportamiento asintótico de $\frac{\mathbf{Y}(n)}{|\mathbf{Y}(n)|}$. Al igual que en el caso de Galton-Watson trabajaremos con una martingala asociada al proceso.

**Proposición 1.20.** *Sean $\{\mathbf{Y}(n),n\in\mathbb{N}\}$ un proceso de ramificación multitipo y $M$ la matriz de medias. Supongamos que $\mu$ es autovector a derecha de $M$ asociado al autovalor $\lambda$. Consideremos $W(n)=\mathbf{Y}(n)\mu\lambda^{-n}$ y $\mathcal{F}_n=$ la $\sigma$-álgebra generada por $\mathbf{Y}(n)$. Entonces la familia $\{(W(n),\mathcal{F}_n),n\in\mathbb{N}\}$ es una martingala no negativa.*

*Demostración.* Claramente $W(n)$ es no negativa y $W(n)\in\mathcal{F}_n$. Observemos que

$$\begin{aligned}\mathbb{E}(W(n+1)|\mathbf{Y}(n))&=\lambda^{-(n+1)}\mathbb{E}(\mathbf{Y}(n+1)\mu|\mathbf{Y}(n))=\lambda^{-(n+1)}\mathbf{Y}(n)M\mu\\&=\lambda^{-(n+1)}\mathbf{Y}(n)\lambda\mu=W(n).\end{aligned}$$

Luego $W(n)$ es una martingala. □

**Corolario 1.21.** *Sea $W(n)=\mathbf{Y}(n)\mu\lambda^{-n}$. Existe una variable aleatoria $W$, tal que cuando $n\to\infty$*

$$W(n)\xrightarrow{cs}W.$$

*Demostración.* Se deduce de aplicar el Teorema 1.8 a la martingala $W(n)$. □

**Definición 1.22.** *Sean $\{\mathbf{Y}(n),n\in\mathbb{N}\}$ un proceso de ramificación multitipo y $M$ la matriz de medias. Si $M$ es $R$-positiva, denotaremos por $W$ al límite de la martingala $\mathbf{Y}(n)\mu R^n$ con $\mu$ autovector a derecha asociado a $1/R$.*

### 1.2.4. Probabilidad de extinción

En esta sección veremos, al igual que en los procesos de Galton-Watson, que la probabilidad de extinción es un punto fijo de la función generadora de momentos. Además relacionaremos la probabilidad de extinción con la variable aleatoria $W$ generada a partir de la martingala asociada al proceso.

**Lema 1.23.** *Sea* $\{\mathbf{Y}(n), n \in \mathbb{N}\}$ *un proceso de ramificación multitipo con* $M = \{m_{ij}, i, j \in \mathcal{S}\}$ *la matriz de medias del proceso. Suponemos que* $M$ *es una matriz irreducible, aperiódica, R-positiva con* $R \in (0,1)$ *y que* $m_{ij}^{(n)}$ *es finito para todo* $n \in \mathbb{N}$, $i, j \in \mathcal{S}$.

*Si* $\mathbf{f}(\mathbf{s})$ *es la función generadora de* $\mathbf{Y}(1)$, *obtenemos los siguientes resultados:*

*a) La función* $\mathbf{f}$ *tiene al menos dos puntos fijos:*

- $\mathbf{1} = (1, 1, \dots)$
- $\mathbf{q} = (q_1, q_2, \dots)$ *donde* $q_i$ *es la probabilidad de que el proceso se extinga, dado que comenzó con un individuo de tipo* $i$.

*b) Si* $\mathbf{s}$ *es un punto fijo de* $\mathbf{f}$ *distinto de* $\mathbf{1}$ *entonces todas sus coordenadas son menores estrictas que 1.*

*Demostración.* $a$) Claramente $\mathbf{1}$ es solución de $\mathbf{f}(\mathbf{s}) = \mathbf{s}$. Consideremos el evento de extinción

$$\bigcup_{n\in\mathbb{N}} \{\mathbf{Y}(n) = \mathbf{0}\}.$$

Llamamos $\mathbf{q} = (q_i, i \in \mathcal{S})$ a la probabilidad de que el proceso se extinga, donde para cada $i \in \mathcal{S}$

$$q_i = \mathbb{P}_i\left(\bigcup_{n\in\mathbb{N}} \{\mathbf{Y}(n) = \mathbf{0}\}\right).$$

Dado $\mathbf{Y}_0 = \delta_i$

$$q_i = \mathbb{P}_i\left(\bigcup_{n\in\mathbb{N}} \{\mathbf{Y}(n) = \mathbf{0}\}\right) = \lim_{n\to\infty} \mathbb{P}_i(\mathbf{Y}(n) = \mathbf{0}) = \lim_{n\to\infty} f_{ni}(\mathbf{0}),$$

de donde obtenemos el siguiente límite coordenada a coordenada $\mathbf{q} = \lim_{n\in\mathbb{N}} \mathbf{f}_n(\mathbf{0})$. Por lo tanto, como $\mathbf{f}_{n+1}(\mathbf{0}) = \mathbf{f}(\mathbf{f}_n(\mathbf{0}))$ y $\mathbf{f}$ es continua tenemos

$$\mathbf{q} = \mathbf{f}(\mathbf{q}).$$

$b$) Sea $\mathbf{s} \in [0,1]^{\mathcal{S}}$ un punto fijo de $\mathbf{f}$ con $s_j < 1$. Supongamos que existe $k \in \mathcal{S}$ tal que $f_k(\mathbf{s}) = 1 = s_k$. Si $\mathbf{s}$ es punto fijo de $\mathbf{f}$ lo es también de $\mathbf{f}_n$ para cualquier $n$. Sea $\mathbf{t} \in [0,1]^{\mathcal{S}}$ tal que

$$t_l = 1 \quad \text{si } l \neq j$$

$$t_l = s_j \quad \text{si } l = j.$$

Como $\mathbf{s} \leq \mathbf{t}$, $f_{nk}(\mathbf{s}) = 1$ y la función $f_{nk}$ es creciente, obtenemos que $1 = f_{nk}(\mathbf{t})$ y por lo tanto

$$1 = \sum_{\mathbf{i} \in \mathbb{N}^{\mathcal{S}}} \mathbb{P}_k(\mathbf{Y}(n) = \mathbf{i})\mathbf{t}^{\mathbf{i}} = \sum_{\mathbf{i} \in \mathbb{N}^{\mathcal{S}}} \mathbb{P}_k(\mathbf{Y}(n) = \mathbf{i}) s_j^{i_j}.$$

Luego $\mathbb{P}_k(\mathbf{Y}(n) = \mathbf{i}) = 0$ si $i_j > 0$, pues $\sum_{\mathbf{i} \in \mathbb{N}^{\mathcal{S}}} \mathbb{P}_k(\mathbf{Y}(n) = \mathbf{i}) = \mathbf{1}$. Con lo cual,

$$m_{kj}^{(n)} = \mathbb{E}_k(Y_j(n)) = \sum_{\mathbf{i} \in \mathbb{N}^{\mathcal{S}}} \mathbb{P}_k(\mathbf{Y}(n) = \mathbf{i}) i_j = 0$$

para cualquier $n \in \mathbb{N}$. Lo cual contradice la irreducibilidad de $M$, luego $\mathbf{s} < \mathbf{1}$. □

**Teorema 1.24.** *Sea $\{\mathbf{Y}(n), n \in \mathbb{N}\}$ un proceso de ramificación multitipo, bajo las condiciones del Lema 1.23. Si existe $i \in \mathcal{S}$ para el cual cuando $n \to \infty$ se tiene $Y_i(n) \xrightarrow{cs} \infty$, la función generadora tiene solo dos puntos fijos: $\mathbf{1}$ y $\mathbf{0}$. En particular, la probabilidad de extinción $\mathbf{q}$ es igual a $\mathbf{0}$.*

*Demostración.* Sea $\mathbf{s}$ un punto fijo de la función generadora distinto de $\mathbf{1}$. Por el Lema 1.23 todas las coordenadas son menores estrictas que 1. Sea $i$ tal que $Y_i(n) \xrightarrow{cs} \infty$, entonces

$$s_k = f_{nk}(\mathbf{s}) = \mathbb{E}_k \left( \prod_{j \in \mathcal{S}} s_j^{Y_j(n)} \right) = \mathbb{E}_k(s_i^{Y_i(n)} \prod_{j \neq i} s_j^{Y_j(n)}) \leq \mathbb{E}_k(s_i^{Y_i(n)}).$$

Por el Teorema de convergencia mayorada, siendo que $s_i < 1$ y que $Y_i(n) \to \infty$ obtenemos que cuando $n \to \infty$, $\mathbb{E}_k(s_i^{Y_i(n)}) \to 0$ y por lo tanto $s_k = 0$ para cualquier $k \in \mathcal{S}$. Luego el único punto fijo distinto de $\mathbf{1}$ es $\mathbf{0}$. □

**Proposición 1.25.** *Sea $W$ la variable aleatoria definida en el Corolario 1.21 y $\mathbf{r} = (r_1, r_2, \ldots)$ el vector con coordenadas $r_i = \mathbb{P}_i(W = 0)$. La función generadora $\mathbf{f}$ tiene a $\mathbf{r}$ como punto fijo.*

*Demostración.* Veamos que $\mathbf{r}$ es un punto fijo de la función generadora. Sea $i \in \mathcal{S}$,

$$\begin{aligned} r_i = \mathbb{P}_i(W = 0) &= \sum_{\mathbf{j} \in \mathbb{N}^{\mathcal{S}}} \mathbb{P}_i(W = 0 | \mathbf{Y}(1) = \mathbf{j}) \mathbb{P}_i(\mathbf{Y}(1) = \mathbf{j}) \\ &= \sum_{\mathbf{j} \in \mathbb{N}^{\mathcal{S}}} \mathbb{P}_i(\mathbf{Y}(1) = \mathbf{j}) \prod_{k \in \mathcal{S}} \mathbb{P}_k(W = 0)^{j_k} \\ &= f_i(\mathbb{P}(W = 0)) = f_i(\mathbf{r}) \end{aligned}$$

□

**Corolario 1.26.** *Sea $\{\mathbf{Y}(n), n \in \mathbb{N}\}$ un proceso de ramificación que verifica las hipótesis del Teorema 1.24. Entonces $\mathbf{r} = \mathbf{0}$ o $\mathbf{r} = \mathbf{1}$.*

### 1.2.5. Teoremas de convergencia

En [Moy67] se muestra el comportamiento asintótico para los procesos de ramificación multitipo en espacios de tipos numerable en el caso supercrítico. Vamos a enunciar este teorema y luego dar el resultado obtenido sobre la convergencia de la proporción de cada tipo individuos.

**Teorema 1.27.** *[Moy67, Teorema 1] Sea $\{\mathbf{Y}(n), n \in \mathbb{N}_0\}$ un proceso de ramificación multitipo. Para $i, j \in \mathcal{S}$ la variable aleatoria $L_{ij}$ representa la cantidad de hijos de tipo $j$ que tiene un individuo de tipo $i$. Sea $m_{ij} = \mathbb{E}(L_{ij})$ y $M = \{m_{ij}, i, j \in \mathcal{S}\}$ la matriz de medias del proceso. Suponemos que $M$ es una matriz irreducible, aperiódica, R-positiva con $R \in (0, 1)$ y que $m_{ij}^{(n)}$ es finito para todo $n \in \mathbb{N}$, $i, j \in \mathcal{S}$. Notamos $\nu$ y $\mu$ los autovectores a izquierda y derecha respectivamente, asociados al autovalor $1/R$ que satisfacen que $\nu\mu = 1$. Si $\mathbb{P}(\mathbf{Y}(0) = \mathbf{y}) = 1$ para $\mathbf{y} \in \mathbb{N}^{\mathcal{S}}$ con $|\mathbf{y}| < \infty$ y se verifica*

$$\sum_{j \in \mathcal{S}} \mathbb{E}_j((\mathbf{Y}(1)\mu)^2)\nu_j < \infty, \tag{1.7}$$

*entonces la variable aleatoria $W$ (Definición 1.22) cumple $\mathbb{E}_{\mathbf{y}}(W^2) < \infty$ y*

$$\lim_{n\to\infty} \mathbb{E}_{\mathbf{y}}((R^n\mathbf{Y}(n)g - \nu g W)^2) = 0,$$

*para todo vector columna $g = (g_1, g_2, \ldots)^t \in \mathbb{R}^{\mathcal{S}\times 1}$ tal que $(g_i/\mu_i)_{i\in\mathcal{S}}$ está acotado. Además, $\mathbb{E}_{\mathbf{y}}(W) = \mathbf{y}\mu$.*

**Definición 1.28.** *Dada una sucesión de variables aleatorias $(\mathbf{X}(n))_{n\in\mathbb{N}}$, con $\mathbf{X}(n) \in \mathbb{R}^{\mathcal{S}}$ y $\gamma \in \mathbb{R}^{\mathcal{S}}$. Notamos $\mathbf{X}(n) \xrightarrow{P} \gamma$ cuando para todo $i \in \mathcal{S}$ obtenemos que cuando $n \to \infty$, $X_i(n) \xrightarrow{P} \gamma_i$. Es decir, consideramos la convergencia en probabilidad coordenada a coordenada.*

A continuación damos condiciones bajo las cuales, a partir del Teorema 1.27, podemos hallar el comportamiento asintótico de la proporción de individuos.

**Proposición 1.29.** *Sea $\{\mathbf{Y}(n), n \in \mathbb{N}_0\}$ un proceso de ramificación multitipo bajo las condiciones del Teorema 1.27. Además suponemos que el autovector a derecha $\mu$ verifica que $\liminf_{i\to\infty} \mu_i \neq 0$ y que el autovector a izquierda $\nu$ cumple $\sum_{j\in\mathcal{S}} \nu_j = 1$.*

*Si para todo $i \in \mathcal{S}$ $\mathbb{P}_i(W = 0) = 0$, cuando $n \to \infty$ se tiene*

$$\frac{\mathbf{Y}(n)}{|\mathbf{Y}(n)|} \xrightarrow{P} \nu.$$

*Demostración.* Por el Teorema 1.27 se tienen en particular los siguientes límites.

- Tomando $g = \delta_i$ entonces

$$\lim_{n\to\infty} \mathbb{E}((R^n Y_i(n) - \nu_i W)^2) = 0.$$

- Tomando $g \equiv 1$ y usando que $\liminf_{i\to\infty} \mu_i \neq 0$

$$\lim_{n\to\infty} \mathbb{E}((R^n \sum_{i\in\mathcal{S}} Y_i(n) - W)^2) = 0.$$

Luego valen los límites en probabilidad:

$$\lim_{n\to\infty} R^n Y_i(n) = \nu_i W \quad \text{y} \quad \lim_{n\to\infty} R^n |\mathbf{Y}(n)| = W.$$

De este último límite se deduce

$$\{|\mathbf{Y}(n)| \to 0\} \subseteq \{W = 0\}. \tag{1.8}$$

Como $\mathbb{P}_i(W = 0) = 0$ para todo $i$ por (1.8) obtenemos que $\mathbb{P}_i(|\mathbf{Y}(n)| \to 0) = 0$.

Siendo que $\mathbb{P}_i(|\mathbf{Y}(n)| \nrightarrow 0) = 1$ obtenemos para cada $j \in \mathcal{S}$ el siguiente límite en probabilidad:

$$\lim_{n\to\infty} \frac{\mathbf{Y}_j(n)}{|\mathbf{Y}(n)|} = \lim_{n\to\infty} \frac{R^n \mathbf{Y}_j(n)}{R^n \sum_{i\in\mathcal{S}} Y_i(n)} = \frac{\nu_j W}{W} = \nu_j.$$

□

**Proposición 1.30.** *Bajo las condiciones del Teorema 1.27, una condición suficiente para que se verifique que $r_i = \mathbb{P}_i(W = 0) = 0$ para todo $i \in \mathcal{S}$ es que exista $j \in \mathcal{S}$ para el cual cuando $n \to \infty$, $Y_j(n) \xrightarrow{cs} \infty$. La variable aleatoria $W$ está definida en 1.22.*

*Demostración.* Notemos que si existe $j \in \mathcal{S}$ para el cual cuando $n \to \infty$ $Y_j(n) \xrightarrow{cs} \infty$, se tiene que la función generadora tiene solo dos puntos fijos, $\mathbf{0}$ y $\mathbf{1}$ (Teorema 1.24). Por el Teorema 1.27 obtenemos que $\mathbb{E}_i(W) = \mu_i > 0$ de donde $r_i = \mathbb{P}_i(W = 0) \neq 1$ y como $\mathbf{r}$ es punto fijo de la función generadora resulta $r_i = 0$ para todo $i$. □

## 1.3. Procesos de ramificación multitipo a tiempo continuo

Una extensión natural es considerar que el tiempo de vida de los individuos es aleatorio. Es decir, considerar el tiempo de vida de cada individuo como una variable aleatoria continua, al término del cual se muere y se produce un número aleatorio de individuos descendientes. En lugar de considerar la cadena de Markov $\{\mathbf{Y}(n), n \in \mathbb{N}\}$ consideráramos el proceso $\{\mathbf{Z}(t), t \geq 0\}$. Para que este proceso sea de Markov debemos imponer que los tiempos de vida tengan distribución exponencial. Para comenzar daremos la construcción del proceso y luego su definición.

### 1.3.1. Construcción del proceso

La siguiente construcción fue extraída de [GB03].

Sea $\mathcal{S}$ el espacio de los distintos tipos de individuos. Cada individuo $i \in \mathcal{S}$ vive un tiempo exponencial de parámetro $a_i > 0$. Para $i, j \in \mathcal{S}$ sea $L_{ij}$ el número de hijos de tipo $j$ que tiene el individuo de tipo $i$. Para cada $i \in \mathcal{S}$ la variable $\mathbf{L}_i = (L_{ij})_{j\in\mathcal{S}}$ tiene distribución $\mathbf{p}_i = (p_{i1}, .., p_{ik}, ..)$, y media finita $m_{ij} = \mathbb{E}(L_{ij})$, $\forall i, j \in \mathcal{S}$.

Construimos el árbol genealógico. Para empezar consideramos el espacio de estados

$$\mathbb{X} = \bigcup_{n\in\mathbb{N}} \mathbb{X}^n,$$

donde $\mathbb{X}^n$ es el espacio donde vive la generación $n$-ésima. Esto es $\mathbb{X}^0 = \mathcal{S}$ y $i_0 \in \mathbb{X}$ es la raíz del árbol. El siguiente es $\mathbb{X}^1 = \mathcal{S} \times \mathbb{N}$ y el elemento $x = (i_1, l_1) \in \mathbb{X}^1$ es el $l_1$-ésimo hijo de tipo $i_1$ de la raíz. Finalmente, para $n > 1$, $\mathbb{X}^n = \mathcal{S}^n \times \mathbb{N}^n$ y $x = (i_1, .., i_n, l_1, .., l_n) \in \mathbb{X}^n$ es $l_n$-ésimo hijo de tipo $i_n$ del padre $\hat{x} = (i_1, .., i_{n-1}, l_1, .., l_{n-1})$. Escribimos $\sigma(x) = i_n$ para el tipo de $x \in \mathbb{X}^n$. Observemos que $\sigma(x) \in \mathcal{S}$.

A cada $x \in \mathbb{X}$ le asociamos

- un tiempo de vida aleatorio $\tau_x$, con distribución exponencial de parámetro $a_{\sigma(x)}$
- descendencia aleatoria, $\mathbf{L}_{\sigma(x)} = (L_{\sigma(x)j})_{j\in\mathcal{S}} \in \mathbb{N}_0^{\mathcal{S}}$ con distribución $\mathbf{p}_{\sigma(x)}$ tal que las variables $\{\tau_x, \mathbf{L}_{\sigma(x)} : x \in \mathbb{X}\}$ son independientes.

Las variables aleatorias $\mathbf{L}_{\sigma(x)}$ indican como son actualizados los individuos $x \in \mathbb{X}$. Por lo tanto el conjunto $X = \bigcup_{n\in\mathbb{N}} X^n$ esta definido recursivamente por

$$X^0 = \{i_0\} \qquad X^n = \{x = (\hat{x}, i_n, l_n) \in \mathbb{X}^n : \hat{x} \in \mathbb{X}^{n-1}, l_n \leq L_{\sigma(\hat{x}),i_n}\}.$$

La variable aleatoria $\tau_x$ es el tiempo de vida de $x \in \mathbb{X}$. Sea $T_x$ el instante en que muere y se reproduce $x$. Lo definimos recursivamente por $T_x = T_{\hat{x}} + \tau_x$, con $T_{\hat{i_0}} = \tau_{\hat{i_0}}$ si $\hat{i_0} \in \mathbb{X}^0$, es decir el tiempo en el que esta vivo $x \in \mathbb{X}$ es: $[T_{\hat{x}}, T_x)$. Por lo tanto, $X(t) = \{x \in X / T_{\hat{x}} \leq t < T_x\}$ es la población a tiempo $t$.

El árbol completo esta determinado por el proceso $X((0,\infty)) = (X(t))_{t\geq 0}$ definido en el espacio Skorohod $\Omega =: D([0,+\infty), \beta(\mathbb{X}))$, es decir el espacio formado por todas las funciones cadlag (continua a derecha con límite a izquierda) de $[0,+\infty)$ a valores en el conjunto $\beta(\mathbb{X})$ que son los subconjuntos finitos de $\mathbb{X}$.

Cuando comenzamos con una distribución inicial $\rho$, la distribución de $X(t)$ en $\Omega$ es $\mathbb{P}_\rho$ y la esperanza es $\mathbb{E}_\rho$, cuando $i_0 = i$ notamos, $\mathbb{P}_i$ y $\mathbb{E}_i$.

Para cada $x \in X(s)$ consideramos, $X(x,t) = \{y \in X : xy \in X(t)\}$, en palabras el conjunto $X(x,t)$ son los descendientes de $x$ que están vivos a tiempo $t$. Por la propiedad de falta de memoria de la distribución exponencial, los descendientes de $X(x,[s,\infty)) = (X(x,t))_{t\geq s}$ con $x \in X(s)$, son independientes de $X([0,s])$.

Consideremos la medida,

$$\mathbf{Z}(t) := \sum_{x\in X(t)} \delta_{\sigma(x)}. \tag{1.9}$$

Es decir, $\mathbf{Z}(t) = (Z_1(t), .., Z_k(t), ..)$ donde $Z_i(t)$ es la cantidad de individuos de tipo $i$ que viven a tiempo $t$. En particular, $Z_i(t)$ es el cardinal de $X_i(t) = \{x \in X(t) : \sigma(x) = i\}$. El total de la población a tiempo $t$ lo denotamos, $|\mathbf{Z}(t)| = \sum_{j\in\mathcal{S}} Z_j(t) = |X(t)|$.

Notemos que el proceso construido verifica

$$\mathbb{E}(\mathbf{s}^{\mathbf{Z}(t)}|\mathbf{Z}(0) = \mathbf{i}) = \prod_{k\in\mathcal{S}} \left(\mathbb{E}(\mathbf{s}^{\mathbf{Z}(t)}|\mathbf{Z}(0) = \delta_k)\right)^{i_k}.$$

Si denotamos $P(\mathbf{i},\mathbf{j},t) := \mathbb{P}(\mathbf{Z}(t)=\mathbf{j}|\mathbf{Z}(0)=\mathbf{i}) = \mathbb{P}_{\mathbf{i}}(\mathbf{Z}(t)=\mathbf{j})$, se verifica

$$\sum_{\mathbf{j}\in\mathbb{N}^{\mathcal{S}}} P(\mathbf{i},\mathbf{j},t)\mathbf{s}^{\mathbf{j}} = \prod_{k\in\mathcal{S}}[\sum_{\mathbf{j}\in\mathbb{N}^{\mathcal{S}}} P(\delta_k,\mathbf{j},t)\mathbf{s}^{\mathbf{j}}]^{i_k}.$$

**Definición 1.31.** *Un proceso de Markov* $\{\mathbf{Z}(t), t\geq 0\}$ *es un proceso de ramificación multitipo (*$\mathcal{S}$*-tipos) a tiempo continuo si las probabilidades de transición cumplen:*

$$\sum_{\mathbf{j}\in\mathbb{N}^{\mathcal{S}}} P(\mathbf{i},\mathbf{j},t)\mathbf{s}^{\mathbf{j}} = \prod_{k\in\mathcal{S}}[\sum_{\mathbf{j}\in\mathbb{N}^{\mathcal{S}}} P(\delta_k,\mathbf{j},t)\mathbf{s}^{\mathbf{j}}]^{i_k},$$

*para todo* $\mathbf{s}\in[0,1]^{\mathcal{S}}$.

Observemos que un proceso de Markov que verifica esta definición puede ser construido como (1.9) y que $\mathbb{P}(\mathbf{Z}(t+h)=\mathbf{0}|\mathbf{Z}(t)=\mathbf{0})=1$ para todo $h,t\geq 0$.

**Proposición 1.32.** *Sea* $\{\mathbf{Z}(t), t\geq 0\}$ *un proceso de ramificación multitipo a tiempo continuo. Cada individuo de tipo* $i\in\mathcal{S}$ *vive un tiempo exponencial de parámetro 1 y luego se reproduce aleatoriamente con distribución* $\mathbf{p}_i$*. Notamos* $L_{ij}$ *a la cantidad de individuos de tipo* $j$ *que tiene un individuo de tipo* $i$*. Sea* $m_{ij}=\mathbb{E}(L_{ij})$ *y* $M=\{m_{ij}, i,j\in\mathcal{S}\}$ *la matriz de medias del proceso, suponemos que* $M$ *es un operador acotado. Para cada* $t\geq 0$ *consideramos la matriz* $C(t)=\{C_{ij}(t), i,j\in\mathcal{S}\}$ *donde* $C_{ij}(t)=\mathbb{E}_i(Z_j(t))$.

*Si* $M$ *es irreducible, aperiódica y* $R$*-positiva* $(R>0)$*, obtenemos para* $t\geq 0$,

$$C(t)=e^{t(M-Id)}$$

*donde* $Id := \mathbf{1}\{i=j, i,j\in\mathcal{S}\}$*. Además para todo* $t\geq 0$*,* $C(t)$ *es irreducible,* $\alpha_t$*-positiva* $(\alpha_t = e^{-\frac{t(1-R)}{R}})$ *y verifica que* $C_{ij}^{(n)}(t)$ *es finito para todo* $n\in\mathbb{N}$*,* $i,j\in\mathcal{S}$*. Los autovectores asociados a* $1/\alpha_t$ *son los autovectores de la matriz* $M$ *asociados a* $1/R$.

*Demostración.* Inicializamos el proceso $(\mathbf{Z}(t))_{t\geq 0}$ con un individuo de tipo $i$. La cantidad de individuos de tipo $j$ a tiempo $t$ la podemos expresar separando en si el individuo inicial se reproduce y tiene $L_{il}$ hijos de tipo $l$ o no se reproduce y solo él sigue vivo:

$$Z_j^{(i)}(t) = \mathbf{1}\{A<t\}\sum_{l\in\mathcal{S}}\sum_{n=1}^{L_{il}} Z_j^{(l,n)}(t-A) + \mathbf{1}\{A>t\}\delta_{ij}$$

donde $A$ es una variable aleatoria con distribución exponencial de parámetro 1 y $Z_j^{(l,n)}(t)$ es la cantidad de hijos de tipo $j$ que tiene el $n$-ésimo individuo de tipo $l$. Tomando esperanza obtenemos

$$\begin{aligned} C_{ij}(t) &= \int_0^t e^{-s}\sum_{l\in\mathcal{S}}\mathbb{E}(L_{il})C_{lj}(t-s)ds + e^{-t}\delta_{ij} \\ &= e^{-t}\int_0^t \sum_{l\in\mathcal{S}}\mathbb{E}(L_{il})e^{u}C_{lj}(u)du + e^{-t}\delta_{ij}. \end{aligned}$$

Por [BE84, Teorema VII.3], este sistema de ecuaciones es equivalente a

$$\begin{aligned}\frac{d}{dt}C_{ij}(t) &= -e^{-t}\int_0^t \sum_{l\in\mathcal{S}} \mathbb{E}(L_{il})e^u C_{lj}(u)du + e^{-t}\sum_{l\in\mathcal{S}}\mathbb{E}(L_{il})e^t C_{lj}(t) - e^{-t}\delta_{ij}\\ &= \sum_{l\in\mathcal{S}}\mathbb{E}(L_{il})C_{lj}(t) - C_{ij}(t).\end{aligned}$$

Es decir, para cada $t \geq 0$, $\frac{d}{dt}C(t) = (M - Id)C(t)$ y con condición inicial $C(0) = Id$. Siendo que $M$ es un operador acotado esta ecuación tiene solución única, [BE84, Teorema VII.3] dada por

$$C_{ij}(t) = (e^{t(M-Id)})_{ij} = \left(\sum_{k\in\mathbb{N}} \frac{t^k(M-Id)^k}{k!}\right)_{ij}$$

donde $e^{t(M-Id)}$ es la función exponencial aplicada a $t(M - Id)$, [Lax02, Sección 17.2].

Dado que $C(t)^n = C(nt)$, los coeficientes $C_{ij}^{(n)}(t)$ son finitos para todo $n \in \mathbb{N}$. Es fácil ver que los autovectores de $C(t)$ asociados a $1/\alpha_t$ son los autovectores de la matriz $M$ asociados a $1/R$ y luego, por el Teorema 1.19, se obtiene que $C(t)$ es $\alpha_t$-positiva, con $\alpha_t = e^{-t(\frac{1}{R}-1)}$. Veamos que $C(t)$ es irreducible,

$$C_{ij}(t) = (e^{t(M-Id)})_{ij} = e^{-t}(e^{tM})_{ij} = e^{-t}\sum_{n=0}^{\infty} m_{ij}^{(n)}\frac{t^n}{n!}.$$

Luego, dado que $M$ es irreducible, $C(t)$ resulta irreducible. □

A continuación definimos el proceso discreto inmerso en el proceso de ramificación $(\mathbf{Z}(t))_{t\geq 0}$, es decir para $\delta > 0$ miramos el proceso continuo a tiempo $n\delta$ con $n \in \mathbb{N}$ lo que da lugar a un proceso de ramificación multitipo a tiempo discreto.

**Lema 1.33.** *Sea $\mathbf{Z} = \{\mathbf{Z}(t) : t \geq 0\}$ un proceso de ramificación multitipo a tiempo continuo con probabilidades de transición $p_{\mathbf{i},\mathbf{j}} = P(\mathbf{i}, \mathbf{j}, t)$. Para cualquier $\delta > 0$, el proceso discretizado $\mathbf{Y}^\delta = \{\mathbf{Y}^\delta(n), n \in \mathbb{N}\}$ con $\mathbf{Y}^\delta(n) = \mathbf{Z}(n\delta)$ para todo $n \in \mathbb{N}$, es un proceso de ramificación multitipo a tiempo discreto con probabilidades de transición $P(\mathbf{i}, \mathbf{j}, \delta)$.*

*Demostración.* Dado $\delta > 0$, consideramos el proceso $\mathbf{Y}^\delta(n) = \mathbf{Z}(n\delta), n \in \mathbb{N}_0$. Este proceso tiene probabilidades de transición $p_{\mathbf{ij}} = P(\mathbf{i}, \mathbf{j}, \delta)$ que verifican la Definición 1.12. □

Por simplicidad al proceso discretizado $\{\mathbf{Y}^\delta(n)), n \in \mathbb{N}\}$ lo notaremos simplemente $\{\mathbf{Y}(n)), n \in \mathbb{N}\}$.

### 1.3.2. Teoremas de convergencia

Enunciamos el siguiente resultado que lo usaremos en el estudio del comportamiento asintótico de $\mathbf{Z}(t)$.

**Teorema 1.34.** *[Kin63, Teorema 2] Sea $g(t)$ una función continua definida en $(0,\infty)$ tal que para cada $\delta > 0$,*

$$\lim_{n\to\infty} g(n\delta) = 0.$$

*Entonces,*

$$\lim_{t\to\infty} g(t) = 0.$$

A continuación damos el teorema principal de este capítulo. Si bien tiene muchas hipótesis técnicas, en los capítulos siguientes mostraremos que éstas se verifican en todos los casos de nuestro interés.

**Teorema 1.35.** *Sea $\{\mathbf{Z}(t), t \geq 0\}$ un proceso de ramificación multitipo a tiempo continuo que verifica las condiciones de la Proposición 1.32. Sean $\nu$ y $\mu$ los autovectores a izquierda y derecha respectivamente asociados al autovalor $1/R$ de la matriz de medias. Supongamos que $R \in (0,1)$, $\liminf_{i\to\infty} \mu_i \neq 0$ y $\sum_{j\in\mathcal{S}} \nu_j = 1$. Si para todo $\delta > 0$*

$$\sum_{j\in\mathcal{S}} \mathbb{E}_j\left((\mathbf{Z}(\delta)\mu)^2\right)\nu_j < \infty \tag{1.10}$$

*y si la variable aleatoria $W$ (Definición 1.22) del proceso discretizado $\mathbf{Y}(n)$ verifica que $\mathbb{P}_i(W=0)=0$ para todo $i \in \mathcal{S}$, entonces cuando $t \to \infty$*

$$\frac{\mathbf{Z}(t)}{|\mathbf{Z}(t)|} \xrightarrow{P} \nu.$$

*Demostración.* Dado $\epsilon > 0$, para cada $j \in \mathcal{S}$ consideremos

$$g_j(t) = \mathbb{P}\left(\left|\frac{Z_j(t)}{|\mathbf{Z}(t)|} - \nu_j\right| > \epsilon\right).$$

Veamos que $g_j$ es una función continua. Sean $0 \leq s < t$

$$\begin{aligned}\left|\mathbb{P}\left(\left|\frac{Z_j(t)}{|\mathbf{Z}(t)|} - \nu_j\right| > \epsilon\right) - \mathbb{P}\left(\left|\frac{Z_j(s)}{|\mathbf{Z}(s)|} - \nu_j\right| > \epsilon\right)\right| \leq \\ \mathbb{E}\left(\left|\mathbf{1}\left\{\left|\frac{Z_j(t)}{|\mathbf{Z}(t)|} - \nu_j\right| > \epsilon\right\} - \mathbf{1}\left\{\left|\frac{Z_j(s)}{|\mathbf{Z}(s)|} - \nu_j\right| > \epsilon\right\}\right|\right) \leq \\ \mathbb{E}\left(\mathbf{1}\left\{\frac{Z_j(t)}{|\mathbf{Z}(t)|} \neq \frac{Z_j(s)}{|\mathbf{Z}(s)|}\right\}\right).\end{aligned}$$

Esta última indicadora vale 0 si no suena ninguna exponencial entre tiempo $s$ y $t$. Condicionando a que, a tiempo $s$, haya en total $|\mathbf{Z}(s)|$ individuos, la última expresión es menor o igual que

$$\mathbb{E}(1 - e^{-|\mathbf{Z}(s)|(t-s)}).$$

Tomando límite y usando convergencia mayorada obtenemos que $g_j(t)$ es continua.

Por otro lado, observemos que dado $\delta > 0$ el proceso de ramificación multitipo $\mathbf{Y}(n) = \mathbf{Z}(\delta n)$ tiene como matriz de medias $C(\delta)$. Esta matriz verifica las condiciones de la Proposición 1.32 con lo cual es irreducible, $\alpha_\delta$-positiva (con $\alpha_\delta = e^{\delta \frac{1-R}{R}} > 1$) y para todo $n \in \mathbb{N}$, $C_{ij}^{(n)}(\delta)$ es finito. Además, $C(\delta)$ tiene todas sus entradas positivas, luego es aperiódica.

Como $\mathbb{P}_i(W = 0) = 0$ para todo $i \in \mathcal{S}$, se verifican las hipótesis de la Proposición 1.29, y cuando $n \to \infty$

$$\frac{\mathbf{Y}_j(n)}{|\mathbf{Y}(n)|} \xrightarrow{P} \nu_j, \qquad \forall j \in \mathcal{S}.$$

Es decir, $g_j(n\delta) \overset{n\to\infty}{\longrightarrow} 0$. Con lo cual podemos aplicar el Teorema 1.34 y extender la convergencia de la proporción al proceso $\mathbf{Z}(t)$,

$$g_j(t) = \mathbb{P}\left(\left|\frac{Z_j(t)}{|\mathbf{Z}(t)|} - \nu_j\right| > \epsilon\right) \overset{t\to\infty}{\longrightarrow} 0.$$

$\square$

**Definición 1.36.** *Dadas dos variables aleatorias $X$ e $Y$ decimos que $X$ está dominada estocásticamente por $Y$, y denotamos $X \preceq Y$ si existen dos variables aleatorias $\hat{X}$ e $\hat{Y}$ ambas definidas en un mismo espacio de probabilidad tal que $F_X = F_{\hat{X}}$, $F_Y = F_{\hat{Y}}$ y $\mathbb{P}(\hat{X} \leq \hat{Y}) = 1$.*

La siguiente proposición da condiciones suficientes sobre la matriz de medias bajo las cuales la condición (1.10) se verifica.

**Proposición 1.37.** *Sea $\{\mathbf{Z}(t), t \geq 0\}$ un proceso de ramificación multitipo a tiempo continuo que verifica las condiciones de la Proposición 1.32. Sean $\nu$ y $\mu$ los autovectores a izquierda y derecha respectivamente asociados al autovalor $1/R$ de la matriz de medias. Supongamos que*

- $R \in (0,1), \quad |\nu| = 1,$
- *existe una constante $C \in (0,\infty)$ que verifica que para todo $i \in \mathcal{S}$ $\mu_i \leq C$,*
- *existe una variable aleatoria $L$ con segundo momento finito tal que para todo $i \in \mathcal{S}$, $\sum_{j\in\mathcal{S}} L_{ij} \preceq L$.*

*Entonces*

$$\sum_{j\in\mathcal{S}} \mathbb{E}_j\left((\mathbf{Z}(\delta)\mu)^2\right)\nu_j < \infty.$$

*Demostración.* Dado $\delta > 0$ la condición $\sum_{j\in\mathcal{S}} \mathbb{E}_j\left((\mathbf{Z}(\delta)\mu)^2\right)\nu_j < \infty$ se reduce a :

$$\sum_{j\in\mathcal{S}} \mathbb{E}_j\left(\left(\sum_{i\in\mathcal{S}} Z_i(\delta)\mu_i\right)^2\right)\nu_j \leq C^2 \mathbb{E}_\nu(|\mathbf{Z}(\delta)|^2).$$

Este último término es el segundo momento del total de individuos a tiempo $\delta$, comenzando con un individuo de tipo $j$ con probabilidad $\nu_j$. Siendo que para todo $i \in \mathcal{S}$, $\sum_{j\in\mathcal{S}} L_{ij} \preceq L$ con $L$ una variable aleatoria tal que $\mathbb{E}(L) < \infty$, podemos acotar la cantidad de individuos del proceso $\mathbf{Z}(t)$ por un proceso de ramificación con un solo tipo de individuo, el cual cada vez que se reproduce tiene $L$ individuos. Llamemos $\{\hat{Z}(t), t \geq 0\}$ a un proceso de ramificación continuo con un solo tipo de individuo, que se inicializa con un individuo. Cada individuo espera un tiempo exponencial de parámetro 1 y luego se reproduce teniendo $L$ individuos. Claramente $|\mathbf{Z}(t)| \preceq \hat{Z}(t)$. En [AN04, Corolario 1, p. 111] se prueba que

$$\mathbb{E}(\hat{Z}(t)^2) < \infty \text{ si y sólo si } \mathbb{E}(L^2) < \infty.$$

Concluimos que

$$\mathbb{E}_\nu(|\mathbf{Z}(\delta)|^2) = \sum_{i\in\mathcal{S}} \nu_i \mathbb{E}_i(|\mathbf{Z}(\delta)|^2) \leq \sum_{i\in\mathcal{S}} \nu_i \mathbb{E}(\hat{Z}(t)^2) = \mathbb{E}(\hat{Z}(t)^2) < \infty.$$

□

# Capítulo 2

# Procesos de urna

De una urna que contiene bolas de distintos colores, se extrae una bola al azar, se observa el color y se la devuelve a la urna junto con un número de bolas de diferentes colores. La cantidad de bolas de cada color que se agregan a la urna sólo depende del color de la bola extraída. Se repite este procedimiento indefinidamente. ¿Cuál es la proporción de cada color dentro de la urna después de muchas extracciones?

Los modelos de urna han sido ampliamente estudiados, comenzando con Pólya (1921), que introdujo un modelo para estudiar la propagación de enfermedades infecciosas. En 1954 Pólya hace la siguiente observación al respecto:

> *Cualquier problema en probabilidad se puede hacer comparable a un problema acerca de urnas que contengan bolas, y cualquier fenómeno aleatorio puede hacerse similar, en sus aspectos esenciales, a extracciones sucesivas de bolas de un sistema combinado de urnas.*

El proceso de urna más simple es llamado *Urna de Pólya*. De una urna con $a$ bolas azules y $b$ blancas se extrae una al azar y se la repone a la urna agregando además $d$ bolas del mismo color que la extraída. Similarmente se puede considerar que cuando se extrae una bola de color azul se la repone y se agregan $d_{11}$ bolas azules y $d_{12}$ bolas blancas y si se extrae una bola de color blanca, se la repone junto con $d_{21}$ bolas azules y $d_{22}$ blancas. La cantidad de bolas de cada color dentro de la urna depende de la urna inicial, de las extracciones y de la matriz de reposición $D = \{d_{ij} : i, j \in \{1, 2\}\}$.

Una generalización inmediata a este proceso es considerar una cantidad finita o numerable de colores de bolas. En el caso finito, si la matriz de reposición $D$ es irreducible y aperiódica, la proporción de bolas de cada color converge casi seguramente a una medida de probabilidad asociada a la matriz $D$. Este resultado es conocido y demostrado usando diferentes técnicas. En [AN04, AK68] se lo demuestra mediante la relación entre procesos de urna y procesos de ramificación multitipo a tiempo continuo introducida por [Fri49]. En [BT14] se demuestra el mismo resultado asociando una cadena de Markov a la matriz de reposición y dando un acoplamiento entre esta cadena y el proceso de urna.

Cuando el conjunto de posibles colores de las bolas es numerable, en [BT14], se prueba la convergencia puntual de la esperanza de la proporción de bolas de cada color. Además se menciona que, al igual que en el caso finito, la proporción de bolas de cada color debería converger casi seguramente a una medida pero sólo lo verifica en un ejemplo. En este capítulo probaremos (bajo ciertas hipótesis) la convergencia en probabilidad.

Usando la relación entre procesos de ramificación y procesos de urna, probamos en el Teorema 2.7, que si la matriz de reposición es múltiplo de una matriz estocástica, recurrente positiva, irreducible y aperiódica $Q$, la proporción de la cantidad de bolas de cada color converge en probabilidad a la única distribución invariante de la matriz $Q$.

Primero damos conceptos básicos sobre cadenas de Markov, luego describimos el modelo de urna, introducimos notación y damos la relación entre procesos de urna y procesos de ramificación. Analizamos el comportamiento asintótico de la proporción de bolas de cada color dentro de la urna. Enunciamos el resultado existente en el caso de una cantidad finita de colores, damos el teorema principal del capítulo, Teorema 2.7 y terminamos el capítulo con algunos ejemplos.

## 2.1. Conceptos básicos de cadenas de Markov

**Definición 2.1.** *Sea $(X(n))_{n\in\mathbb{N}}$ una cadena de Markov definida en un espacio de estados $\mathcal{S}$ finito o numerable con matriz de transición $P$. El vector $\nu$ se dice que es una distribución estacionaria de la cadena si:*

1. $\nu_i \geq 0, \quad \forall i \in \mathcal{S} \quad y \quad \sum_{i\in\mathcal{S}} \nu_i = 1.$
2. $\nu P = \nu.$

La primera condición muestra que $\nu$ es una probabilidad en $\mathcal{S}$ y la segunda implica que es una medida invariante para el proceso.

**Definición 2.2.** *Sea $(X(n))_{n\in\mathbb{N}}$ una cadena de Markov definida en un espacio de estados $\mathcal{S}$. Sean $i, j \in \mathcal{S}$.*

- *Si $X(0) = i$ definimos*
$$T_{ij} = \text{ínf}\{n \in \mathbb{N}/X(n) = j\}.$$
*En palabras, $T_{ij}$ es la primera vez que la cadena visita el estado $j$, siendo que $X(0) = i$. En particular decimos que $T_{ii}$ es el tiempo de retorno del estado $i$.*
- *Un estado $i$ se dice recurrente si $\mathbb{P}(T_{ii} < \infty) = 1$.*
- *Un estado $i$ se dice recurrente positivo si $\mathbb{E}(T_{ii}) < \infty$.*
- *Una cadena de Markov (o la matriz asociada $P$) se dicen recurrente (positiva) si todos sus estados son recurrentes (positivos).*

El siguiente resultado garantiza la existencia y unicidad de la medida invariante.

**Teorema 2.3.** *[Bré99, Teorema 3.1] Sea $(X(n))_{n\in\mathbb{N}}$ una cadena de Markov definida en un espacio de estados $\mathcal{S}$ finito o numerable con matriz de transición $P$ irreducible. La cadena $(X(n))_{n\in\mathbb{N}}$ es recurrente positiva si y sólo si $P$ tiene una única medida invariante $\nu$.*

**Observación 2.4.** *Sea $P$ una matriz positiva, irreducible y recurrente positiva. Entonces $P$ es 1-positiva con $\nu$ y $\mathbf{1}$ autovectores a izquierda y derecha respectivamente. Además $\sum_i \nu_i < \infty$.*

## 2.2. Descripción del modelo

De una urna que contiene bolas de distintos colores, se extrae una bola al azar, se observa el color y se devuelve a la urna junto con un número de bolas de diferentes colores. Notamos $K \in \mathbb{N}$ a la cantidad de colores. El proceso comienza con composición inicial

$$\mathbf{X}(0) = (x_1, \ldots, x_K) \qquad \text{con } x_i \in \mathbb{N}_0.$$

Luego de $n \in \mathbb{N}$ extracciones, denotamos la configuración de la urna por

$$\mathbf{X}(n) = (X_1(n), \ldots, X_K(n)),$$

donde $X_j(n)$ representa el número de bolas de color $j$ a tiempo $n$. La dinámica del proceso depende únicamente de la forma en que se reemplazan las bolas, que se da a través de la matriz de reposición $D = \{d_{ij}, i, j \in \{1, \ldots, K\}\}$, donde $d_{ij}$ es el número de bolas de color $j$ que se agregan en la urna cuando se extrajo una bola de color $i$. Notamos $J(n)$ al color de la bola seleccionada en el paso $n$-ésimo y $D_{i,\cdot}$ a la fila $i$-ésima de la matriz $D$, es decir $D_{i,\cdot} := (d_{i1}, \ldots, d_{iK})$. Describimos la dinámica del proceso a través de

$$\mathbf{X}(n+1) = \mathbf{X}(n) + D_{J(n),\cdot}.$$

Si $J(n)$ es el color de la bola extraída en el $n$-ésimo paso, entonces

$$\mathbb{P}(J(n+1) = i | \mathbf{X}(0), \mathbf{X}(1), \ldots, \mathbf{X}(n)) = \frac{X_i(n)}{\sum_{i=1}^K X_i(n)}, \quad \forall i \in \{1, \ldots, K\}$$

lo que implica que

$$\mathbb{P}(J(n+1) = i) = \mathbb{E}\left(\frac{X_i(n)}{\sum_{i=1}^K X_i(n)}\right).$$

Cuando los posibles colores son un conjunto $\mathcal{S}$ numerable, la matriz de reposición $D$ es $D = \{d_{ij}/i, j \in \mathcal{S}\}$. El proceso comienza con composición inicial

$$\mathbf{X}(0) = (x_1, x_2, \ldots) \qquad \text{con } x_i \in \mathbb{N}_0.$$

Luego de $n \in \mathbb{N}$ extracciones, denotamos la configuración de la urna por

$$\mathbf{X}(n) := (X_1(n), X_2(n), \ldots),$$

donde $X_j(n)$ representa el número de bolas de color $j$ a tiempo $n$. Sin pérdida de generalidad podemos suponer que se comienza con solo una bola en la urna $X(0) = \delta_i$ para algún $i \in \mathcal{S}$. Luego de la $n$-ésima reposición habrá $\sum_{i\in\mathcal{S}} X_i(n)$ bolas en la urna y se obtiene que $\frac{\mathbf{X}(n)}{\sum_{i\in\mathcal{S}} X_i(n)}$ es la proporción de los diferentes colores a tiempo $n$.

Para cada $n \in \mathbb{N}$, notamos $J(n)$ al color de la bola seleccionada en el paso $n$, y $D_{i,\cdot}$ a la fila $i$-ésima de la matriz $D$. Describimos la dinámica del proceso a través de

$$\mathbf{X}(n+1) = \mathbf{X}(n) + D_{J(n),\cdot}.$$

Si $J(n)$ es el color de la bola extraída en el $n$-ésimo paso, entonces

$$\mathbb{P}(J(n+1) = i | \mathbf{X}(0), \mathbf{X}(1), \ldots, \mathbf{X}(n)) = \frac{X_i(n)}{\sum_{i\in\mathcal{S}} X_i(n)}, \quad \forall i \in \mathcal{S}$$

lo que implica que

$$\mathbb{P}(J(n+1) = i) = \mathbb{E}\left(\frac{X_i(n)}{\sum_{i\in\mathcal{S}} X_i(n)}\right).$$

## 2.3. Relación con los procesos de ramificación

El proceso de urna se puede pensar como el esqueleto de un proceso de ramificación multitipo a tiempo continuo, en donde cada individuo de tipo $i$ vive un tiempo exponencial y luego se reproduce teniendo un hijo de su mismo tipo y exactamente $d_{ij}$ hijos de tipo $j$. El siguiente teorema puede verse en [AN04, Teorema 2, p. 221] que da la relación entre ambos procesos cuando hay una cantidad finita de colores. A continuación damos la extensión inmediata para numerables colores.

**Teorema 2.5.** *Sea $\mathbf{X} = (\mathbf{X}(n))_{n\in\mathbb{N}}$ el proceso que representa la cantidad de bolas en una urna con $\mathcal{S}$ el conjunto de los posibles colores de las bolas y $D = \{d_{ij}/i, j \in \mathcal{S}\}$ la matriz de reposición de la urna. Sea $\mathbf{Z} = (\mathbf{Z}(t))_{t\geq 0}$ un proceso de ramificación multitipo a tiempo continuo donde cada individuo de tipo $\mathcal{S}$ tiene asociado un tiempo de vida exponencial de parámetro 1 independiente del resto de los individuos. Suponemos que cada individuo de tipo i muere y tiene $d_{ij} + \delta_{ij}$ descendientes de tipo $j$. Para cada $n \in \mathbb{N}$ sea $\tau_n$ la $n$-ésima vez que un individuo se reproduce.*

*El proceso de urna generalizado $\mathbf{X} = (\mathbf{X}(n))_{n\in\mathbb{N}}$ y el proceso $(\mathbf{Z}(\tau_n))_{n\in\mathbb{N}}$ tienen la misma distribución.*

*Demostración.* Notemos primero que ambos procesos son a tiempo discreto, y son cadenas de Markov en espacio de estados discreto. Debemos ver que que los procesos tienen las mismas probabilidades de transición. El proceso de ramificación comienza con $Z_i(0) = X_i(0)$ individuos de tipo $i \in \mathcal{S}$. Los tiempos de vida de los individuos son independientes entre sí y tienen distribución exponencial de parámetro 1, con lo cual la probabilidad de que el primer individuo en reproducirse sea de tipo $i$ es simplemente

$$\frac{Z_i(0)}{\sum_{i\in\mathcal{S}} Z_i(0)} = \frac{X_i(0)}{\sum_{i\in\mathcal{S}} X_i(0)}.$$

Luego se crean exactamente $d_{ij} + \delta_{ij}$ individuos de tipo $j$. El proceso comienza de nuevo pues el tiempo de vida de los individuos tiene pérdida de memoria. Esto muestra que si $\mathbf{X}(0) = \mathbf{Z}(0)$, luego $\mathbf{X}(1)$ tiene igual distribución que $\mathbf{Z}(\tau_1)$. Iterando este procedimiento se obtiene que ambos procesos tienen igual distribución. □

## 2.4. Comportamiento asintótico

El principal resultado de este capítulo es la caracterización del comportamiento asintótico de la proporción de bolas de colores dentro de la urna. Primero veremos el caso conocido de finitos colores y luego probaremos el caso numerable.

**Teorema 2.6.** *[AN04, Sección 9.3, Capítulo 5] Sean $K \in \mathbb{N}$ y $D = \{d_{ij}/i, j \in \{1, \ldots, K\}\}$ una matriz irreducible y aperiódica con $d_{ij} \in \mathbb{N}_0$. Consideremos el proceso de urna $\mathbf{X}$, con matriz de reposición $D$. Entonces para cada $i \in \mathcal{S}$ se verifica que si $n \to \infty$*

$$\frac{X_i(n)}{\sum_{i \in \mathcal{S}} X_i(n)} \xrightarrow{cs} \nu_i.$$

*Demostración.* La demostración es una aplicación del Teorema 2.5 y del resultado análogo al Teorema 1.35 en el caso finito. □

El teorema anterior se prueba de forma alternativa en [BT14, Corolario 5.2.1]. Cabe destacar que en [BT14], se extienden las condiciones del proceso permitiendo reponer a la urna proporciones de bolas y se restringe suponiendo que siempre se agrega la misma cantidad de bolas.

A continuación damos la convergencia en probabilidad de la proporción de bolas a una medida en $\mathbb{N}$. Supondremos que siempre se repone a la urna la misma cantidad de bolas.

**Teorema 2.7.** *Consideremos el modelo de urna con $\mathcal{S}$ un conjunto numerable que indica los posibles colores de las bolas y $D = \{d_{ij}/i, j \in \mathcal{S}\}$ la matriz de reposición de la urna que verifica que $d_{ij} \in \mathbb{N}_0$ y que $\sum_{j \in \mathcal{S}} d_{ij} = \beta$ para todo $i \in \mathcal{S}$ y para algún $\beta \in \mathbb{N}$. Suponemos que la matriz estocástica $D/\beta$ es irreducible, aperiódica y recurrente positiva. Sea $\mathbf{X} = (\mathbf{X}(n))_{n \in \mathbb{N}}$ el proceso que representa la cantidad de bolas en la urna, entonces para todo $i \in \mathcal{S}$ se verifica que cuando $n \to \infty$*

$$\frac{X_i(n)}{\beta n + 1} \xrightarrow{P} \nu_i,$$

*donde $\nu$ es la única solución de $\nu D = \beta \nu$.*

*Demostración.* Observemos que la matriz $Q = D/\beta$ verifica el Teorema 2.3 con lo cual existe una única medida invariante $\nu$. Además $Q$ es 1-positiva con $\nu$ y $\mathbf{1}$ autovectores a izquierda y derecha respectivamente.

Consideremos $\mathbf{Z} = (\mathbf{Z}(t))_{t \geq 0}$ el proceso de ramificación definido en el Teorema 2.5. Para cada $n \in \mathbb{N}$ sea $\tau_n$ el tiempo de reproducción del $n$-ésimo individuo. Siendo que el proceso $\mathbf{Z}$ no explota, es decir para todo tiempo hay una cantidad finita de individuos,

$\tau_n \to \infty$ cuando $n \to \infty$. Cada individuo espera un tiempo exponencial de parámetro 1 y luego se reproduce teniendo exactamente $\beta + 1$ descendientes. En particular un individuo de tipo $i$ tiene exactamente $L_{ij} = d_{ij} + \delta_{ij}$ hijos de tipo $j$ y siendo que tiene descendientes de forma determinística la matriz de medias del proceso de ramificación es $M = D + Id$. El proceso $\mathbf{Z}$ verifica que $\sum_{j \in \mathcal{S}} L_{ij} = \beta + 1 < \infty$ para todo $i \in \mathcal{S}$ con lo cual $M$ resulta un operador acotado como muestra el lema a continuación. Además como $D$ es irreducible, aperiódica y $\beta$-positiva, resulta que $M$ es irreducible, aperiódica y $(\beta+1)$-positiva. Luego $\mathbf{Z}$ verifica las condiciones de la Proposición 1.32 y para cada $t \geq 0$

$$\mathbb{E}(\mathbf{Z}(t)) = e^{tD}.$$

Más aún $\mathbb{E}(\mathbf{Z}(t))$ es $\alpha_t$-positiva con $\alpha_t = e^{-t\beta}$ y autovectores $\nu$, $\mathbf{1}$ a izquierda y derecha respectivamente asociados al autovalor $1/\alpha_t$.

Claramente $\mathbf{Z}$ está bajo las condiciones de la Proposición 1.37, con lo cual se verifica la condición (1.10):

$$\sum_{i \in \mathcal{S}} \mathbb{E}_i(|\mathbf{Z}(\delta)|^2)\nu_i < \infty.$$

Por otro lado, dado $\delta > 0$ consideremos el proceso discretizado $\mathbf{Y}(n) = \mathbf{Z}(\delta n)$ que es un proceso de ramificación multitipo a tiempo discreto (ver Lema 1.33) con matriz de medias $\mathbb{E}(\mathbf{Z}(\delta))$. Consideremos $W$ la variable aleatoria que es el límite de la martingala asociada a $\mathbf{Y}$ (Definición 1.22). Veamos que $W$ verifica que $\mathbb{P}_i(W = 0) = 0$ para todo $i \in \mathcal{S}$. Para ello probaremos que la cantidad de individuos de algún tipo tiende a infinito. Si el proceso comienza con $\mathbf{X}(0) = \delta_i$, definimos para cada $n \in \mathbb{N}$ los eventos

$$A_n := \{ \text{ se extrae una bola de color } i \text{ en la } n\text{-ésima extracción } \}$$

y la $\sigma$-álgebra generada por las bolas extraídas

$$\mathcal{F}_n := \sigma(\{J(k), \quad 1 \leq k \leq n\}).$$

Claramente $A_n \in \mathcal{F}_n$. Notemos que

$$\mathbb{P}(A_n|\mathcal{F}_{n-1}) \geq \frac{1}{\beta n + 1}.$$

Luego $\sum_{n \in \mathbb{N}} \mathbb{P}(A_n|\mathcal{F}_{n-1}) = \infty$ y por el Teorema de Borel Cantelli condicionado [Dur96, Teorema 5.3.2], se obtiene que con probabilidad 1 el evento $A_n$ ocurre infinitas veces, es decir, infinitas veces se extrae una bola del color inicial. Notemos que existe $l \in \mathbb{N}$ tal que $d_{il} > 0$, con lo cual cada vez que se extrae una bola de color $i$ se repone al menos una bola de color $l$. Siendo que el evento $A_n$ ocurre infinitas veces, entonces infinitas veces se repone una bola de color $l$, con lo cual $X_l(n) \underset{n\to\infty}{\longrightarrow} \infty$, y se verifica

$$\mathbb{P}_i(Z_l(\tau_n) \to \infty) = 1.$$

La función $\mathbf{Z}(t)$ es monótona creciente en $t$, y $\tau_n \to \infty$ con lo cual obtenemos que $\mathbb{P}_i(Z_l(t) \to \infty) = 1$ lo que implica $\mathbb{P}_i(Y_l(n) \to \infty) = 1$. Además el proceso $\mathbf{Y}$ verifica las hipótesis del Teorema 1.27, con lo cual por la Proposición 1.30 se obtiene que $\mathbb{P}_i(W = 0) = 0$ para todo $i$.

Finalmente por el Teorema 1.35, cuando $t \to \infty$

$$\frac{Z_i(t)}{|\mathbf{Z}(t)|} \xrightarrow{P} \nu_i, \quad \forall\, i \in \mathcal{S}.$$

Obtenemos que para cada $i \in \mathcal{S}$ cuando $n \to \infty$

$$\frac{Z_i(\tau_n)}{|\mathbf{Z}(\tau_n)|} \xrightarrow{P} \nu_i.$$

Luego por el Teorema 2.5 se obtiene que cuando $n \to \infty$

$$\frac{X_i(n)}{\beta n + 1} = \frac{X_i(n)}{|\mathbf{X}(n)|} = \frac{Z_i(\tau_n)}{|\mathbf{Z}(\tau_n)|} \xrightarrow{P} \nu_i.$$

□

La prueba de arriba queda completa gracias al siguiente lema.

**Definición 2.8.** *Se define el espacio $\ell^1$ como*

$$\ell^1 := \{x \in \mathbb{R}^{\mathcal{S}} : \sum_{i \in \mathcal{S}} |x_i| < \infty\}.$$

*La norma en este espacio es $\|x\|_1 = \sum_{i \in \mathcal{S}} |x_i|$. Dado un operador lineal acotado $T : \ell^1 \to \ell^1$ su norma viene dada por*

$$\|T\|_1 := \sup_x \frac{\|T(x)\|_1}{\|x\|_1}.$$

**Lema 2.9.** *Sean $Q = \{q_{ij}, i, j \in \mathcal{S}\}$ y $K \in \mathbb{R}$ que verifican $\sum_{j \in \mathcal{S}} |q_{ij}| \leq K$ para todo $i \in \mathcal{S}$. Consideremos $\widetilde{Q} : \ell^1 \to \ell^1$ definido como $\widetilde{Q}(x) = xQ$. Entonces $\|\widetilde{Q}\|_1 \leq K$.*

*Más aún, debido a*

$$\|\widetilde{Q}^n\|_1 \leq \|\widetilde{Q}\|_1^n < \infty,$$

*las coordenadas de $Q^n$, denotadas $q_{ij}^{(n)}$, son finitas para todo $n$.*

*Demostración.*

$$\widetilde{Q}(x) = \left( \sum_{i \in \mathcal{S}} x_i q_{i1}, \sum_{i \in \mathcal{S}} x_i q_{i2}, \, \ldots \right) \Longrightarrow$$

$$\|\widetilde{Q}(x)\|_1 = \sum_{j \in \mathcal{S}} |\sum_{i \in \mathcal{S}} x_i q_{ij}| \leq \sum_{i \in \mathcal{S}} |x_i| \sum_{j \in \mathcal{S}} |q_{ij}| \leq K \|x\|_1.$$

□

Mediante un abuso de notación denotaremos $Q$ en vez de $\widetilde{Q}$.

## 2.5. Ejemplos

En esta sección damos algunas matrices a las cuales se le puede aplicar el Teorema 2.7 y se puede dar de forma explícita la medida invariante.

En ambos ejemplos consideramos un proceso de urna con bolas de $\mathbb{N}$ colores distintos y matriz de reposición $D$.

**Ejemplo 1:**

$$D = \begin{pmatrix} a & \beta - a & 0 & \dots \\ a & 0 & \beta - a & \dots \\ \vdots & & & \ddots \end{pmatrix},$$

con $\beta \in \mathbb{N}$ y $a \in \{1, \dots, \beta - 1\}$. Es decir

$$d_{ij} = \begin{cases} a & \text{si } i = 1 \\ \beta - a & \text{si } j = i + 1 \\ 0 & \text{si no.} \end{cases}$$

Puede verse que $D/\beta$ es irreducible, aperiódica y recurrente positiva. Con lo cual existe una única medida invariante $\nu = (\nu_i)_{i \in \mathbb{N}}$, donde

$$\nu_j = \frac{a}{\beta}\left(1 - \frac{a}{\beta}\right)^{j-1} \quad \text{para todo } j \in \mathbb{N}.$$

Por el Teorema 2.7 se tiene que para cada $i \in \mathbb{N}$ cuando $n \to \infty$

$$\frac{X_i(n)}{\beta n + 1} \xrightarrow{P} \frac{a}{\beta}\left(1 - \frac{a}{\beta}\right)^{j-1}.$$

**Ejemplo 2:**

$$D = \begin{pmatrix} 0 & a + b + c & 0 & 0 & \dots \\ a & c & b & 0 & \dots \\ 0 & a & c & b & \dots \\ \vdots & 0 & & & \ddots \end{pmatrix},$$

con $a, b, c \in \mathbb{N}$. Es decir para $i \geq 2$

$$d_{ij} = \begin{cases} a & \text{si } j = i - 1 \\ c & \text{si } j = i \\ b & \text{si } j = i + 1 \\ 0 & \text{si no.} \end{cases}$$

Puede verse que $\frac{D}{a+b+c}$ es irreducible, aperiódica y recurrente positiva. Con lo cual existe una única medida invariante $\nu = (\nu_i)_{i \in \mathbb{N}}$, con

$$\nu_i = \alpha \left(\frac{b}{a}\right)^{i-1}$$

donde $\alpha$ es una constante de normalización, [Bré99, Ejemplo 5.7]. Por el Teorema 2.7 se tiene que para cada $i \in \mathbb{N}$ cuando $n \to \infty$

$$\frac{X_i(n)}{\beta n + 1} \xrightarrow{P} \nu_i.$$

# Capítulo 3

# Distribuciones Cuasiestacionarias

Las distribuciones cuasiestacionarias han sido ampliamente estudiadas desde los trabajos de Kolmogorov [Kol38], Yaglom [Yag47] y Sevastyanov [Sev51]. Aparecen como un objeto natural al considerar los procesos de Markov que son absorbidos dado que éstas se definen como el conjunto de distribuciones invariantes para la evolución del proceso condicionada a la no ocurrencia de este hecho. Un estudio amplio de distribuciones cuasiestacionarias puede verse en el libro [CMSM13].

Podemos separar el estudio de las QSD según si su espacio de estados es finito, numerable o no numerable. En esta tesis nos centramos en los casos en que el espacio de estados es finito o numerable.

Uno de los primeros problemas que surge en el estudio de las QSD es su existencia. Cuando el espacio de estados es finito, Darroch y Seneta [DS65] probaron que existe una única QSD independientemente de la distribución inicial. En cambio, cuando el espacio de estados es numerable hay resultados parciales pudiendo existir 0, 1 o infinitas distribuciones cuasiestacionarias. Hay varias condiciones para garantizar que exista al menos una QSD, entre ellas la $R$-positividad de la matriz de transición (Sección 3.2).

En la primer sección de este capítulo tratamos cadenas de Markov con estados absorbentes, para las cuales la cadena es absorbida con probabilidad uno. Para estos procesos introducimos la evolución condicionada y vemos cómo calcular su comportamiento asintótico. En la segunda sección consideramos en particular los procesos $R$-positivos y damos condiciones para la existencia del límite de Yaglom y de las distribuciones cuasiestacionarias. En la última sección damos el resultado principal de este capítulo, Teorema 3.19, donde se demuestra que bajo una condición de Döeblin (Definición 3.18) el proceso es $R$-positivo, tiene una única QSD y además el autovector a derecha $\mu$ asociado a $1/R$ de la matriz de transición cumple $0 < \tilde{c} \leq \mu_j \leq \tilde{C} < \infty$ para todo $j$ y para ciertas constantes $\tilde{c}, \tilde{C}$.

## 3.1. Distribuciones cuasiestacionarias

Comenzamos definiendo y enunciando las principales propiedades de las QSD. Si la cadena está definida en un espacio de estados $\mathcal{S}$, cualquier elemento o subconjunto de

$\mathcal{S}$ podrían ser estados absorbentes. Sin pérdida de generalidad, en este trabajo el estado absorbente será el 0. Los resultados de esta sección pueden verse con más detalle en [CMSM13].

**Definición 3.1.** *Sea $Y = (Y(n))_{n\in\mathbb{N}}$ una cadena de Markov definida en un espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$, con estados $\mathcal{S}_0 := \mathcal{S} \cup \{0\}$ numerable y matriz de transición $P_0$ con entradas $p_{ij}, i, j \in \mathcal{S}_0$. Decimos que 0 es un estado absorbente si $p_{0i} = 0$ para todo $i \in \mathcal{S}$.*

Notamos $P_0$ a la matriz de transición con entradas $p_{ij}$ con $i, j \in \mathcal{S}_0$, y simplemente $P$ a la matriz restringida a $\mathcal{S}$, es decir, la matriz $P_0$ sin la fila y la columna correspondientes al 0.

**Definición 3.2.** *Sea $Y = (Y(n))_{n\in\mathbb{N}}$ una cadena de Markov a valores en un espacio $\mathcal{S}_0$, y $(\mathcal{F}_n)_{n\in\mathbb{N}}$ la filtración natural asociada. Dado $A \subset \mathcal{S}_0$, definimos el tiempo de parada $\tau_A$ :*

$$\tau_A = \min\{n \in \mathbb{N} / Y(n) \in A\}.$$

*En particular, definimos el tiempo de parada $\tau_0$ como el tiempo de absorción, es decir:*

$$\tau_0 = \min\{n \in \mathbb{N} / Y(n) = 0\}.$$

**Definición 3.3.** *Para un proceso que empieza con distribución inicial $\gamma$, definimos la evolución del proceso condicionada a no ser absorbido como:*

$$\varphi_j^\gamma(n) = \mathbb{P}_\gamma(Y(n) = j | \tau_0 > n).$$

Equivalentemente, podemos expresar a la evolución del proceso condicionada como:

$$\varphi_j^\gamma(n) = \frac{\mathbb{P}_\gamma(Y(n) = j)}{\mathbb{P}_\gamma(Y(n) \neq 0)} = \frac{(\gamma P^n)_j}{\sum_{i\in\mathcal{S}} (\gamma P^n)_i}.$$

De aquí se ve que para cada $n \in \mathbb{N}$, $(\varphi_j^\gamma(n))_{j\in\mathcal{S}}$ es una medida de probabilidad en $\mathcal{S}$.

**Definición 3.4.** *Cuando existe $\lim_{n\to\infty} \varphi_j^{\delta_i}(n)$ para todo $i \in \mathcal{S}$ y es una medida de probabilidad en $\mathcal{S}$ que no depende del estado $i$ se lo denomina límite de Yaglom. Denotamos*

$$\lim_{n\to\infty} \varphi_j^{\delta_i}(n) = \nu_j.$$

Al igual que con las medidas invariantes, cuando este límite existe, es invariante para la evolución de la probabilidad condicionada. A estas medidas se las llama distribuciones cuasiestacionarias.

**Definición 3.5.** *Una medida $\nu$ en $\mathcal{S}$ se dice distribución cuasiestacionaria (QSD) si verifica:*

- *$\nu_i \geq 0$ para $i \in \mathcal{S}$ y $\sum_{i\in\mathcal{S}} \nu_i = 1$.*
- *$\varphi_i^\nu(n) = \nu_i$, para $n \in \mathbb{N}$ e $i \in \mathcal{S}$.*

La primer condición nos dice que es una medida de probabilidad y la segunda que es invariante para el proceso condicionado.

**Proposición 3.6.** *Una medida $\nu$ es una QSD si y sólo si*

$$\nu_j = \sum_{i \in \mathcal{S}} p_{ij} \nu_i + \sum_{i \in \mathcal{S}} p_{i0} \nu_j.$$

*Demostración.* Dada $\gamma$ una distribución inicial

$$\begin{aligned}
\varphi_j^\gamma(n+1) &= \frac{\mathbb{P}_\gamma(Y(n+1)=j)}{\mathbb{P}_\gamma(Y(n+1) \neq 0)} = \frac{\sum_{i\in\mathcal{S}} \mathbb{P}(Y(n+1)=j|Y(n)=i)\mathbb{P}_\gamma(Y(n)=i)}{\sum_{z\in\mathcal{S}} \mathbb{P}_\gamma(Y(n+1)=z)} \\
&= \frac{\sum_{i\in\mathcal{S}} p_{ij}\varphi_i^\gamma(n)}{\sum_{z\in\mathcal{S}}\sum_{i\in\mathcal{S}} p_{iz}\varphi_i^\gamma(n)} = \frac{\sum_{i\in\mathcal{S}} p_{ij}\varphi_i^\gamma(n)}{\sum_{z\in\mathcal{S}}\varphi_i^\gamma(n)\sum_{i\in\mathcal{S}} p_{iz}} \\
&= \frac{\sum_{i\in\mathcal{S}} p_{ij}\varphi_i^\gamma(n)}{\sum_{i\in\mathcal{S}}\left(\varphi_i^\gamma(n)(1-p_{i0})\right)} = \frac{\sum_{i\in\mathcal{S}} p_{ij}\varphi_i^\gamma(n)}{1-\sum_{i\in\mathcal{S}}\varphi_i^\gamma(n)p_{i0}}.
\end{aligned}$$

Luego se tiene

$$\varphi_j^\gamma(n+1) = \sum_{i\in\mathcal{S}} p_{ij}\varphi_i^\gamma(n) + \sum_{i\in\mathcal{S}} \varphi_i^\gamma(n) p_{i0} \varphi_j^\gamma(n+1). \tag{3.1}$$

Por lo tanto $\nu$ es una distribución cuasiestacionaria si y sólo si

$$\nu_j = \sum_{i\in\mathcal{S}} p_{ij}\nu_i + \sum_{i\in\mathcal{S}} p_{i0}\nu_i\nu_j \quad \Leftrightarrow \quad \nu_j\left(1-\sum_{i\in\mathcal{S}} p_{i0}\nu_i\right) = \sum_{i\in\mathcal{S}} p_{ij}\nu_i. \tag{3.2}$$

□

**Observación 3.7.** *De la Proposición anterior se desprende que una QSD $\nu$, es autovector a izquierda de la matriz $P$, con autovalor asociado $1-\sum_{i\in\mathcal{S}}\nu_i p_{i0}$.*

El límite de Yaglom, cuando existe, es único. En cambio, existen procesos con infinitas QSD (por ejemplo el proceso de nacimiento y muerte en la Sección 5.2). De aquí se deduce que hay QSD que no son límites de Yaglom, pero sí ocurre que si el límite de Yaglom existe entonces es una QSD. Este hecho puede verse fácilmente cuando la matriz de transición $P$ verifica $\sum_{i\in\mathcal{S}} p_{ij} < \infty$ para todo $j \in \mathcal{S}$. De hecho, si $\nu$ es el límite de Yaglom entonces tomando límite en la ecuación (3.1) y usando convergencia mayorada se obtiene que

$$\nu_j = \sum_{i\in\mathcal{S}} p_{ij}\nu_i + \sum_{i\in\mathcal{S}} p_{i0}\nu_i\nu_j, \tag{3.3}$$

y por lo tanto es una distribución cuasiestacionaria. Pero cabe destacar que este resultado vale con total generalidad [MV12, Proposición 1].

En el caso en que $\mathcal{S}$ sea finito, la teoría de Perron-Frobenius asegura la existencia de un autovalor de módulo máximo con multiplicidad simple y autovectores a izquierda y derecha asociados a este autovalor.

**Teorema 3.8** (Perron-Frobenius). *[Sen06] Sea $A$ una matriz no negativa e irreducible, de $r \times r$. Entonces, existe un autovalor $\lambda_1$ real, positivo, con multiplicidad 1 y tal que $\lambda_1 > |\lambda_j|$ para todo $\lambda_j$ autovalor de $A$. Además sean $\nu, \mu$ autovectores a izquierda y derecha respectivamente asociados a $\lambda_1$, podemos elegirlos positivos y tal que $|\nu| = 1$ y $\nu\mu = 1$. Sean $\lambda_2, .., \lambda_r$ los demás autovalores de $A$, tales que $\lambda_1 > |\lambda_2| \geq ... \geq |\lambda_r|$. Sea $m_j$ la multiplicidad de $\lambda_j$. Luego,*

$$A^n = \lambda_1^n \mu\nu + O(n^{m_2-1}|\lambda_2|^n), \quad \forall n \in \mathbb{N},$$

*donde $\mu\nu$ es una matriz de $r \times r$.*

Del Teorema de Perron-Frobenius se deduce

**Teorema 3.9.** *[DS65] Dada una cadena de Markov definida en $\mathcal{S}_0$, con $\mathcal{S}$ un conjunto finito y 0 el estado absorbente. Sea $P_0$ la matriz de transición, con $P$ irreducible. Existe una única distribución cuasiestacionaria, que es el límite de Yaglom.*

Cuando $\mathcal{S}$ es numerable el panorama es completamente distinto, puede haber 0, 1 o infinitas QSD. Por lo que nos interesan resultados de existencia de QSD, y en el caso de que existan infinitas nos interesaremos en particular por la QSD con menor tiempo medio de absorción.

**Definición 3.10.** *Decimos que $\nu^*$ es la QSD minimal si cumple:*

$$\mathbb{E}_{\nu^*}(\tau_0) = \inf\{\mathbb{E}_\nu(\tau_0)/\nu \text{ es una } \ QSD \ \}.$$

Damos algunos resultados sobre el tiempo de absorción. Para mas detalle puede verse en [CMSM13, Capítulo 4].

**Proposición 3.11.** *Si $\nu$ es una QSD, entonces cuando el proceso empieza con $\nu$, el tiempo de absorción $\tau_0$ tiene distribución geométrica. Es decir, existe $\alpha = \alpha(\nu) \in (0, 1)$ tal que*

$$\mathbb{P}_\nu(\tau_0 > n) = \alpha^n.$$

*Demostración.* Primero observemos que una QSD es una medida $\nu$ en $\mathcal{S}$ tal que

$$\mathbb{P}_\nu(Y(n) = i | \tau_0 > n) = \nu_i, \text{ para todo } i \in \mathcal{S}.$$

Esta condición toma la forma:

$$\mathbb{P}_\nu(Y(n) = i) = \nu_i \mathbb{P}_\nu(\tau_0 > n). \tag{3.4}$$

Por otro lado, como $Y$ es una cadena de Markov,

$$\begin{aligned}
\mathbb{P}_\nu(\tau_0 > n + m) &= \sum_{i \in \mathcal{S}} \mathbb{P}_\nu(\tau_0 > n + m | Y_m = i) \mathbb{P}_\nu(Y_m = i) \\
&= \sum_{i \in \mathcal{S}} \mathbb{P}_i(\tau_0 > n) \mathbb{P}_\nu(\tau_0 > m) \nu_i \\
&= \mathbb{P}_\nu(\tau_0 > m) \mathbb{P}_\nu(\tau_0 > n).
\end{aligned}$$

Luego

$$\mathbb{P}_\nu(\tau_0 > n) = \mathbb{P}_\nu(\tau_0 > 1)^n = \alpha^n,$$

es decir $\tau_0$ tiene distribución geométrica con parámetro $1 - \alpha$. □

Consideramos el coeficiente

$$\theta(\nu) := -\ln(\alpha(\nu)) \in (0, \infty).$$

El parámetro $\theta(\nu)$ es comúnmente llamado tasa de decaimiento de la QSD. Este verifica:

$$\mathbb{P}_\nu(\tau_0 > n) = e^{-\theta(\nu)n} \quad \text{para todo} \quad n \in \mathbb{N}. \tag{3.5}$$

De las ecuaciones (3.4) y (3.5) se obtiene que $\nu$ es una QSD si y sólo si existe $\theta \in (0, \infty)$ tal que se verifica

$$\mathbb{P}_\nu(Y(n) = i) = \nu_i e^{-\theta n}.$$

En este caso $\theta = \theta(\nu)$.

**Lema 3.12.** *Para que exista una QSD $\nu$ es necesario que exista $\theta > 0$ tal que*

$$\mathbb{E}_i(e^{\theta\tau_0}) < \infty, \ \textit{para algún } i \in \mathcal{S}.$$

*Si la cadena es irreducible, entonces es necesario que exista $\theta^* > 0$ tal que*

$$\mathbb{E}_i(e^{\theta^*\tau_0}) < \infty, \ \textit{para todo } i \in \mathcal{S}.$$

*Demostración.* Sea $\nu$ una QSD y $\theta(\nu) = -\ln(\mathbb{P}_\nu(\tau_0 > 1))$. Para todo $\theta \leq \theta(\nu)$ obtenemos

$$\mathbb{E}_\nu(e^{\theta\tau_0}) \leq \sum_{n\in\mathbb{N}} e^{\theta n}\mathbb{P}_\nu(\tau_0 \geq n) \leq \sum_{n\in\mathbb{N}} e^{\theta n} e^{-\theta(\nu)n} < \infty.$$

Además siendo que $\mathbb{E}_\nu(e^{\theta\tau_0}) = \sum_{i\in\mathcal{S}} \nu_i \mathbb{E}_i(e^{\theta\tau_0})$ se obtiene que existe $i \in \mathcal{S}$ tal que

$$\mathbb{E}_i(e^{\theta\tau_0}) < \infty.$$

Supongamos ahora que la cadena es irreducible. Siendo que existe $\theta > 0$ tal que

$$\mathbb{E}_\nu(e^{\theta\tau_0}) = \sum_{i\in\mathcal{S}} \nu_i \mathbb{E}_i(e^{\theta\tau_0}) < \infty$$

se obtiene que $\mathbb{E}_i(e^{\theta\tau_0}) < \infty$ para todo $i \in \mathcal{S}$ (pues $\nu_i > 0$ para todo $i$). $\square$

En [FKMP95] se prueba que esta condición no solo es necesaria si no que además es (bajo ciertas condiciones) suficiente.

**Teorema 3.13.** *[FKMP95, Teorema 1.1] Asumimos que el tiempo de absorción es finito; $\mathbb{P}_i(\tau_0 < \infty) = 1$ y que para $n \in \mathbb{N}$,*

$$\lim_{i\to\infty} \mathbb{P}_i(\tau_0 < n) = 0.$$

*Bajo estas condiciones, que exista $\theta > 0$ tal que*

$$\mathbb{E}_i(e^{\theta\tau_0}) < \infty$$

*para todo $i \in \mathcal{S}$, es una condición necesaria y suficiente para la existencia de una QSD.*

## 3.2. $\boldsymbol{R}$-positividad

La $R$-positividad es una condición que asegura la existencia de los autovectores a izquierda y derecha de la matriz de transición asociados al autovalor mayor. Bajo esta condición el comportamiento de las probabilidades asintóticas es similar al dado por el Teorema de Perrón Frobenius para el caso finito y será de utilidad para el análisis de las QSD. Si bien en la Sección 1.2.3 dimos algunos resultados sobre matrices $R$-positivas, en esta sección trabajamos con matrices $R$-positivas que además son subestocásticas.

Recordemos los siguientes conceptos de matrices no negativas.

**Teorema 3.14.** *[VJ67, Teorema A] Sea $P$ una matriz no negativa, irreducible y aperiódica. Entonces existe $R$ tal que*

$$\lim_{n\to\infty} (p_{ij}^{(n)})^{1/n} = 1/R \quad \forall \quad i,j \in \mathcal{S}.$$

Decimos que la matriz $P$ subestocástica es $R$-positiva, si se cumple alguna de las siguientes condiciones equivalentes

- Para todo $i,j$ las sucesiones $p_{i,j}^{(n)} R^n$ tienden a un valor positivo cuando $n \to \infty$.
- existen únicos autovectores positivos a izquierda y derecha $((\nu_k),(\mu_k)\ )$, asociados al autovalor $\frac{1}{R}$, tal que $\sum_{k\in\mathbb{N}} \mu_k \nu_k < \infty$.

Cuando el proceso de Markov esta definido a tiempo continuo, la matriz de tasas del proceso es $R$-positiva y el autovector a izquierda $\nu$ asociado a $1/R$ es sumable, entonces $\nu$ es el límite de Yaglom del proceso, [AEGR15, Teorema 3.1]. A continuación damos una prueba similar para procesos definidos a tiempo discreto.

**Teorema 3.15.** *Sea $Y = \{Y(n), n \in \mathbb{N}\}$ un proceso de Markov con matriz de transición $P_0$. Suponemos que $P$ es irreducible, aperiódica y $R$-positiva. Si el autovector a izquierda $\nu$ asociado a $\frac{1}{R}$ verifica que $\sum_{i\in\mathbb{N}} \nu_i < \infty$ entonces $\nu$ es el límite de Yaglom.*

*Demostración.* Sea $P$ una matriz irreducible, aperiódica y $R$-positiva. Con lo cual, existen $\mu$ y $\nu$ autovectores positivos a derecha e izquierda asociados al autovalor $\frac{1}{R}$ tal que, $\nu P = \frac{1}{R}\nu$, $P\mu = \frac{1}{R}\mu$, $\nu\mu = 1$. Por hipótesis, suponemos que $\sum_{i\in\mathbb{N}} \nu_i = 1$.

Verifiquemos que $\nu$ es el límite de Yaglom. Para ello, consideremos $Z = \{Z(n), n \in \mathbb{N}\}$ una cadena de Markov con matriz de transición $H = RU^{-1}PU$, con $U$ una matriz con $\mu$ en la diagonal y ceros fuera de la diagonal. Observemos que la matriz $H$ en el lugar $(i,j)$ es:

$$h_{ij} = R\mu_i^{-1} p_{ij} \mu_j.$$

Luego $\sum_{j\in\mathbb{N}} h_{ij} = R\mu_i^{-1} \sum_{j\in\mathbb{N}} p_{ij}\mu_j = 1$ para todo $i$ por lo que la matriz $H$ es estocástica. La matriz $P$ es irreducible y $R$-positiva, con lo cual $H$ resulta irreducible, recurrente positiva y el vector $\nu U = (\nu_1\mu_1, \nu_2\mu_2, \dots)$ es la medida invariante $(\nu U H = \nu U)$.

Luego para cada $i \in \mathbb{N}$

$$\lim_{n\to\infty} \mathbb{P}_i(Z(n) = j) = (\nu U)_j = \nu_j \mu_j, \text{ para todo } j.$$

Si denotamos $\mathbf{1}$ al vector columna de unos, en términos de las matrices el límite anterior queda:

$$\lim_{n\to\infty} H^n = \mathbf{1}\nu U.$$

Multiplicando por la matriz $U^{-1}$ a derecha obtenemos $\lim_{n\to\infty} H^n U^{-1} = \mathbf{1}\nu$, y multiplicando por $U$ a izquierda $\lim_{n\to\infty} UH^nU^{-1} = U\mathbf{1}\nu$. Por lo tanto

$$\lim_{n\to\infty} R^n P^n = \mu\nu,$$

donde $\mu\nu$ es la matriz que tiene en el lugar $(i,j)$ a $\mu_i\nu_j$. Se deduce el resultado

$$\mathbb{P}_i(Y(n) = j|\tau_0 > n) = \frac{\mathbb{P}_i(Y(n) = j)}{\sum_{j\in\mathcal{S}} \mathbb{P}_i(Y(n) = j)} = \frac{R^n P^n_{ij}}{\sum_j R^n P^n_{ij}} \overset{n\to\infty}{\longrightarrow} \frac{\mu_i\nu_j}{\mu_i} = \nu_j,$$

donde $\sum_j R^n P^n_{ij} \to \mu_i$ por convergencia mayorada ya que

$$R^n P^n_{ij} = \frac{R^n \nu_i P^n_{ij}}{\nu_i} \leq \frac{\nu_j}{\nu_i} \quad \text{y} \quad \frac{\sum_j \nu_j}{\nu_i} = \frac{1}{\nu_i} < \infty.$$

Luego $\nu$ es el límite de Yaglom y por lo tanto una $QSD$. □

El siguiente teorema da condiciones bajo las cuales un proceso es $R$-positivo, y además asegura que el autovector asociado al autovalor $1/R$ es la QSD minimal.

**Teorema 3.16.** *[FKM96, Teorema 1] Sea $X = \{X(n), n \in \mathbb{N}\}$ una cadena de Markov irreducible definida en el espacio $\mathcal{S}_0$ con 0 el estado absorbente. Si la cadena verifica que para algún $i \in \mathcal{S}$, $\mathbb{P}_i(\tau_0 < \infty) = 1$ y además se cumplen las siguientes condiciones:*

1. *existen un conjunto no vacío $\mathcal{S}' \subseteq \mathcal{S}$, un $\epsilon_0 > 0$ y una constante $C_1$ tal que para todo $i \in \mathcal{S}'$ y para todo $n \geq 0$,*

$$\mathbb{P}_i(\tau_0 > n, \ \textit{pero } X(l) \notin \mathcal{S}' \textit{ para todo } 1 \leq l \leq n) \leq C_1(R + \epsilon_0)^{-n},$$

2. *existen un estado $i_0 \in \mathcal{S}'$ y una constante $C_2$ tal que para todo $j \in \mathcal{S}'$ y $n \geq 0$,*

$$\mathbb{P}_j(\tau_0 > n) \leq C_2 \mathbb{P}_{i_0}(\tau_0 > n)$$

3. *existen un conjunto finito $\tilde{\mathcal{S}} \subseteq \mathcal{S}$ y constantes $0 \leq n_0 < \infty$, $C_3 > 0$ tal que para todo $i \in \mathcal{S}'$*

$$\mathbb{P}_i(X(n) \in \tilde{\mathcal{S}} \textit{ para algún } n \leq n_0) \geq C_3.$$

*Luego, el proceso $X$ es R-positivo con autovectores a izquierda y derecha $\nu$ y $\mu$, tal que $\nu$ verifica $\sum_{i\in\mathcal{S}} \nu_i = 1$ y define la QSD minimal. Además,*

$$\lim_{n\to\infty} R^n \mathbb{P}_i(\tau_0 > n) = \mu_i,$$

$$\lim_{n\to\infty} \mathbb{P}_i(X(n) = j|\tau_0 > n) = \nu_j.$$

**Observación 3.17.** *Si el proceso $X$ es R-positivo con autovector a izquierda $\nu$ asociado a $1/R$ tal que $\nu$ es una probabilidad entonces por Fatou $\nu$ tiene que ser la QSD minimal.*

## 3.3. Condición de Döeblin

En esta sección damos una condición de Döeblin que garantiza que la matriz asociada al proceso es $R$-positiva y además que el proceso tiene una única QSD. Bajo esta condición en [JR95] se prueba, para procesos a tiempo continuo, que existe a lo sumo una QSD. A continuación damos la definición equivalente para tiempo discreto.

**Definición 3.18.** *Sea $P_0 = \{p_{ij} : i, j \in \mathcal{S}_0\}$ una matriz de transición.*

- *Denotamos $\alpha(i) := \text{ínf}_{j \neq i} p_{ji}$. Si $\alpha(i) > 0$ decimos que $i$ es un estado de Döeblin.*
- *Definimos el coeficiente de ergodicidad de la matriz $P_0$ como $\alpha := \sum_{i \in \mathcal{S}} \alpha(i)$.*
- *Definimos la tasa de máxima absorción como $C := \sup_{i \in \mathcal{S}} p_{i0}$.*

En palabras, si un proceso tiene un estado de Döeblin $i$ significa que de cualquier estado se llega al estado $i$ en un paso con probabilidad al menos $\alpha(i)$. Observemos que si se considera un proceso que con probabilidad positiva se absorbe, se tiene que $C > 0$.

El siguiente teorema asegura que un proceso que verifique la condición $\alpha > C$ es $R$-positivo y además acota inferior y superiormente a los coeficientes del autovector a derecha asociado al autovalor $1/R$. Nos será de utilidad en el siguiente capítulo para aplicar el método de AFP a procesos que verifiquen esta condición.

**Teorema 3.19.** *Sea $X = \{X(n), n \in \mathbb{N}\}$ una cadena de Markov definida en el espacio $\mathcal{S}_0$ con 0 el estado absorbente y $P_0$ la matriz de transición. Supongamos que $P = \{p_{ij}; i.j \in \mathcal{S}\}$ es una matriz subestocástica, irreducible y aperiódica que verifica la condición $\alpha > C$. Entonces $P$ es $R$-positiva con $\nu$ y $\mu$ autovectores a izquierda y derecha respectivamente asociados al autovalor $1/R$. El autovector $\nu$ es la única QSD y el autovector $\mu$ verifica*

$$0 < \tilde{c} \leq \mu_j \leq \tilde{C} < \infty,$$

*para todo $j \geq 1$ y para constantes $\tilde{c}, \tilde{C}$.*

*Demostración.* Notemos la siguiente relación,

$$\begin{aligned}
\sum_{j \in \mathcal{S}} \mathbb{P}_l(X_{n-1} = j) &= \sum_{j \in \mathcal{S}} \sum_{i \in \mathcal{S}} \mathbb{P}_l(X_{n-2} = i) \mathbb{P}(X_{n-1} = j | X_{n-2} = i) \\
&= \sum_{i \in \mathcal{S}} \mathbb{P}_l(X_{n-2} = i) \sum_{j \in \mathcal{S}} \mathbb{P}(X_{n-1} = j | X_{n-2} = i) \\
&= \sum_{i \in \mathcal{S}} \mathbb{P}_l(X_{n-2} = i)(1 - p_{i0}) \\
&> \sum_{i \in \mathcal{S}} \mathbb{P}_l(X_{n-2} = i)(1 - C).
\end{aligned}$$

Iterando obtenemos la inecuación:

$$\sum_{j \in \mathcal{S}} \mathbb{P}_l(X_{n-1} = j) > (1 - C)^{n-1}.$$

La matriz $P$ verifica la condición $\alpha > C$, con lo cual existe un estado $l \in \mathcal{S}$ que cumple $\alpha(l) > 0$. Luego se tiene

$$\begin{aligned}(p^{(n)})_{ll} = \mathbb{P}_l(X_n = l) &= \sum_{j\in\mathcal{S}} \mathbb{P}_l(X_{n-1} = j)p_{jl} \\ &> \alpha(l)\sum_{j\in\mathcal{S}} \mathbb{P}_l(X_{n-1} = j) \\ &> \alpha(l)(1-C)^{n-1}.\end{aligned}$$

Por el Teorema 3.14 existe $R$ tal que $\lim_{n\to\infty}(p_{ij}^{(n)})^{1/n} = 1/R$ para todo $i, j \in \mathcal{S}$. En particular

$$\lim_{n\to\infty}(p_{ll}^{(n)})^{1/n} = 1/R.$$

Con lo cual, tomando raíz $n$-ésima de ambos miembros de $(p^{(n)})_{ll} > \alpha(l)(1-C)^{n-1}$ y pasando al límite obtenemos

$$\frac{1}{R} \geq 1 - C.$$

Sin pérdida de generalidad podemos asumir que existe un conjunto finito $\mathcal{S}' \subset \mathcal{S}$ tal que $\sum_{j\in\mathcal{S}'} \inf_{i\in\mathcal{S}} p_{ij} = \alpha > C$. Luego existe $\epsilon > 0$ tal que

$$1 - \alpha = \frac{1}{R+\epsilon}.$$

Verifiquemos que la matriz $P$ está bajo las hipótesis del Teorema 3.16.

- Siendo que $\mathbb{P}_i(\tau_{\mathcal{S}'} > n) = \mathbb{P}(X(k) \notin \mathcal{S}', \text{para } k \in \{1, 2, \ldots, n\}) \leq (1-\alpha)^n$,

  $$\mathbb{P}_i(\tau_{\mathcal{S}'} > n, \tau_0 > n) \leq \mathbb{P}_i(\tau_{\mathcal{S}'} > n) \leq (1-\alpha)^n = \frac{1}{(R+\epsilon)^n}.$$

- El conjunto $\mathcal{S}'$ es finito y $P$ irreducible con lo cual para cualquier $i, j \in \mathcal{S}'$ existe una constante $C_2 > 0$ tal que

  $$\mathbb{P}_i(\tau_0 > n) \leq C_2\mathbb{P}_j(\tau_0 > n).$$

- Siendo que $\mathcal{S}'$ es finito, simplemente observemos que para cualquier $i \in \mathcal{S}'$,

  $$\mathbb{P}_i(X(0) \in \mathcal{S}') = 1.$$

Luego por el Teorema 3.16 el proceso $X$ es $R$-positivo y existen autovectores $\nu$ y $\mu$ a izquierda y derecha asociados al autovalor $1/R$ tal que $\nu$ es la QSD minimal de $X$. Además se verifican los siguientes límites

$$\lim_{n\to\infty} \mathbb{P}_i(X(n) = j|\tau_0 > n) = \nu_j \quad \text{y} \quad \lim_{n\to\infty} R^n\mathbb{P}_i(\tau_0 > n) = \mu_i.$$

Para ver que $\nu$ es la única QSD del proceso $X$ consideremos un proceso de Markov a tiempo continuo con matriz de tasas $Q_0$, donde $q_{ij} := p_{ij}$ para todo $i \neq j$, y resulta $q_{jj} = -\sum_{i \in \mathcal{S}_0 \setminus \{j\}} p_{ji}$. Claramente la matriz $Q$ verifica la condición $\alpha > C$ y en [JR95] se prueba que $Q$ tiene una única QSD. Por lo tanto la ecuación $\gamma Q = \lambda \gamma$ tiene única solución con $\gamma = (\gamma_1, \gamma_2, \ldots)$, $\gamma_i > 0$ para todo $i \in \mathcal{S}$, $|\gamma| = 1$ y $\lambda = -\sum_{k \in \mathcal{S}} \gamma_k q_{k0}$ (la forma del autovalor se desprende de una ecuación similar al caso discreto).

$$\gamma Q = \lambda \gamma \iff \sum_{i \in \mathcal{S}} \gamma_i q_{ij} = \lambda \gamma_j \quad \forall j \in \mathcal{S} \iff$$

$$\sum_{i \neq j} \gamma_i p_{ij} + \gamma_j q_{jj} = -\sum_{k \in \mathcal{S}} \gamma_k q_{k0} \; \gamma_j \quad \forall j \in \mathcal{S} \iff$$

$$\sum_{i \neq j} \gamma_i p_{ij} = \left( -\sum_{k \in \mathcal{S}} \gamma_k p_{k0} + \sum_{i \in \mathcal{S} \setminus \{j\}} p_{ji} + p_{j0} \right) \gamma_j \quad \forall j \in \mathcal{S} \iff$$

$$\sum_{i \neq j} \gamma_i p_{ij} = \left( -\sum_{k \in \mathcal{S}} \gamma_k p_{k0} + 1 - p_{jj} \right) \gamma_j \quad \forall j \in \mathcal{S} \iff$$

$$\sum_{i} \gamma_i p_{ij} = \left( 1 - \sum_{k \in \mathcal{S}} \gamma_k p_{k0} \right) \gamma_j \quad \forall j \in \mathcal{S} \iff \gamma P = (1 + \lambda) \gamma.$$

Siendo que $Q$ tiene una única QSD, resulta que $P$ tiene una única QSD y es $\nu$.

En lo que sigue calculamos las cotas del autovector $\mu$. Observemos que

$$\begin{aligned} \mathbb{P}_i(\tau_0 > n) &= \mathbb{P}_i(\tau_0 > n | \tau_{\mathcal{S}'} > n) \mathbb{P}_i(\tau_{\mathcal{S}'} > n) + \mathbb{P}_i(\tau_0 > n | \tau_{\mathcal{S}'} \leq n) \mathbb{P}_i(\tau_{\mathcal{S}'} \leq n) \\ &\leq \mathbb{P}_i(\tau_{\mathcal{S}'} > n) + \mathbb{E}_i(\mathbb{P}_{X(\tau_{\mathcal{S}'})}(\tau_0 > n - \tau_{\mathcal{S}'}) \mathbf{1}\{\tau_{\mathcal{S}'} \leq n\}) \\ &\leq (1 - \alpha)^n + \mathbb{E}_i(g(n - \tau_{\mathcal{S}'})), \end{aligned}$$

donde $g(k) = \max_{j \in \mathcal{S}'} \mathbb{P}_j(\tau_0 > k)$. Luego

$$\begin{aligned} \mu_i = \lim_{n \to \infty} R^n \mathbb{P}_i(\tau_0 > n) &\leq \limsup_n \left[ R^n (1 - \alpha)^n + \mathbb{E}_i(R^n g(n - \tau_{\mathcal{S}'})) \right] \\ &\leq \limsup_n \mathbb{E}_i \left( R^{n - \tau_{\mathcal{S}'}} g(n - \tau_{\mathcal{S}'}) R^{\tau_{\mathcal{S}'}} \right) \\ &\leq \left( \sup_n R^n g(n) \right) \mathbb{E}_i R^{\tau_{\mathcal{S}'}} =: \tilde{C}. \end{aligned}$$

El primer factor es finito pues $R^n g(n)$ tiene límite y no depende de $i$. Para acotar el segundo factor, como $\mathbb{P}_i(\tau_{\mathcal{S}'} > n) \leq (1 - \alpha)^n = (R + \epsilon)^{-n}$ para todo $n \geq 1$ se cumple que

$$\mathbb{E}_i(R^{\tau_{\mathcal{S}'}}) = \sum_{n \in \mathbb{N}} R^n \mathbb{P}(\tau_{\mathcal{S}'} = n) \leq \sum_{n \in \mathbb{N}} R^n \mathbb{P}(\tau_{\mathcal{S}'} > n - 1) = R \sum_{n \in \mathbb{N}_0} R^n (R + \epsilon)^{-n} < \infty.$$

Queda probado que $\mu_i \leq \tilde{C}$ para todo $i \in \mathcal{S}$ con $\tilde{C}$ una constante que no depende de $i$.

Por otro lado tenemos la cota inferior

$$\mu_i = R \sum_{j \in \mathcal{S}} p_{ij} \mu_j \geq R \sum_{j \in \mathcal{S}} \inf_{i \neq j} p_{ij} \mu_j \geq R \sum_{j \in \mathcal{S}'} \inf_{i \neq j} p_{ij} \mu_j \geq R \alpha \min_{j \in \mathcal{S}'} \mu_j =: \tilde{c} > 0.$$

□

# Capítulo 4

# Simulación de QSD

Un problema importante en el estudio de las QSD es su aproximación ya que cuando el espacio de estados es muy grande o infinito no siempre es posible calcularlas fácilmente. En el caso en que haya más de una QSD nos interesan métodos que simulen la QSD minimal, que es la que tiene menor tiempo medio de absorción (Definición 3.10).

Hay varios métodos para simular distribuciones cuasiestacionarias. Por ejemplo, los procesos de tipo Fleming-Viot, introducidos en [BHM00] y generalizados a espacios numerables en [FM07]. Aldous, Flannery y Palacios (AFP) en [AFP88] proponen otro método (en espacios finitos) donde, dado un proceso de Markov, construyen un nuevo proceso sin estados absorbentes, cuya medida empírica converge a la QSD del proceso original.

En este capítulo damos dos resultados principales. Primero el Teorema 4.7 que prueba que el método de AFP se puede extender a procesos en espacios numerables que verifican la condición de Döeblin dada en el capítulo anterior. La prueba de su convergencia se basa en los resultados obtenidos para procesos de ramificación multitipo a tiempo continuo. Además damos un nuevo método, similar al de AFP, el cual involucra procesos de ramificación multitipo a tiempo discreto. Probamos que para procesos que verifiquen las condiciones de la Proposición 4.9 el nuevo método converge a la QSD minimal. En el Teorema 4.18 aplicamos este método a procesos de Galton-Watson subcríticos y otros procesos $R$-positivos. Cabe destacar que los procesos de Galton-Watson no verifican la condición de Döeblin antes mencionada y los procesos de ramificación multitipo asociados al método de AFP no caen en las hipótesis bajo las cuales obtuvimos convergencia en este caso.

## 4.1. Método de Aldous-Flannery-Palacios (AFP)

Aldous, Flannery y Palacios [AFP88] propusieron un método para obtener una aproximación de la distribución cuasiestacionaria y dieron una prueba de convergencia para espacios de estados finitos. La idea es construir una cadena de Markov sin estados absorbentes cuya medida empírica converja a la QSD del proceso original.

El método es el siguiente: dado un proceso de Markov $X = \{X(n), n \in \mathbb{N}\}$ definido en $\mathcal{S} \cup \{0\}$ con 0 un estado absorbente, se genera un nuevo proceso $V = (V(n))_{n\in\mathbb{N}}$ sobre $\mathcal{S}$. Este proceso $V$ se construye de la siguiente manera: mientras que $X(n) \neq 0$, se concatena $X(n)$ a la cadena $V$. Si $X(n) = 0$ no se cuenta la transición, y se elige un estado $j \in \mathcal{S}$ según la medida empírica de la historia (hasta ese momento) de $V$. Se repite el procedimiento con $X(0) = j$ y de esta forma se obtiene $V$.

Se relaciona la medida empírica de $V$ (ver Definición 4.1) con un proceso de ramificación multitipo supercrítico a tiempo continuo. A partir de ello se obtiene la convergencia de la medida empírica de $V$, con la cual se aproxima la QSD de $X$.

**Definición 4.1.** *Dado un proceso $V = (V(n))_{n\in\mathbb{N}}$ se define $(\frac{\xi(n)}{n})_{n\in\mathbb{N}}$ la medida empírica a tiempo $n-1$, donde para cada $i \in \mathcal{S}$*

$$\xi_i(n) = \sum_{l=0}^{n-1} \mathbf{1}\{V(l) = i\}.$$

En la siguiente definición se da la construcción del proceso generado por el método de AFP.

**Definición 4.2.** *[AFP88] Dada una cadena de Markov $X$ con estado absorbente, se construye el proceso $V$ de la siguiente forma.*
*Inicializamos $T_0 = 0$, $m = 1$, $n = 0$, $X(0) = i$.*
*Se corre una cadena $X$, con matriz de transición $P_0$ e inicializada en $X(0)$.*

- *Si $X(n) \neq 0$,*
  - $V(T_{m-1} + n) = X(n)$,
  - $n = n + 1$.
- *si no*
  - $T_m = n + T_{m-1}$
  - *Con distribución $\frac{\xi(T_m)}{T_m}$ se sortea un estado de $\mathcal{S}$, al que notamos $j$.*
  - *Se borra la cadena $X$ y se corre una nueva cadena independiente $X$ pero con estado inicial $X(0) = j$.*
  - $m = m + 1$, $n = 1$.

*Iterando este procedimiento, obtenemos $V = \{V(n), n \in \mathbb{N}\}$.*

Observemos que $T_m$ es el tiempo de la $m$-ésima absorción y que

$$T_m - T_{m-1} | X(T_{m-1}) = j \sim T_1 | X(0) = j.$$

**Definición 4.3.** *Sean $\frac{\xi(n)}{n}$ la distribución empírica de $V$ a tiempo $n-1$ y $\mathcal{M}$ un conjunto de medidas de probabilidad en $\mathcal{S}$. Definimos la cadena de Markov $(V(n-1), \xi(n))_{n\in\mathbb{N}}$ en $\mathcal{S} \times \mathcal{M}$, con $V_0 = i_1, \xi(1) = e_{i_1}$ y transiciones:*

$$\mathbb{P}(V(n) = j, \xi(n+1) = \xi + e_j | V(n-1) = i, \xi(n) = \xi) = p_{ij} + p_{i0}\frac{\xi_j}{|\xi|}.$$

Demos un ejemplo simple de la construcción de los procesos $V$ y $\mathbf{Z}$ a partir de corridas del proceso $X$.

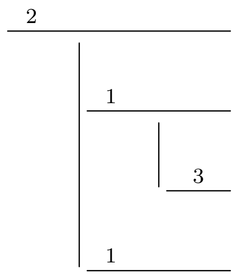
**Ejemplo 4.4.** *Supongamos que $X$ arroja los siguientes estados:*

$$X(0) = 2, X(1) = 1, X(2) = 1, X(3) = 0,$$

*luego se elige con distribución $(2/3, 1/3)$ entre el 1 y el 2. Supongamos que sale el 1. Se vuelve a correr una cadena $X$ inicializada en 1 y se obtiene*

$$X(0) = 1, X(1) = 3, X(2) = 0.$$

*De estos dos experimentos se obtiene que $V$ es $(2, 1, 1, 3)$ y $\mathbf{Z}$ se puede representar de la siguiente manera*




**Lema 4.5.** *Sea $X$ un proceso de Markov con estado absorbente y $P_0$ la matriz de transición. Sea $V$ el proceso construido en la Definición 4.2, $\frac{\xi(n)}{n}$ la distribución empírica (Definición 4.1) y $T_n$ el n-ésimo tiempo de absorción. Si $P$ es R-positiva entonces existe un proceso de ramificación multitipo a tiempo continuo $\mathbf{Z} = (\mathbf{Z}(t))_{t \geq 0}$ que verifica las hipótesis de la Proposición 1.32 y tiene las siguientes propiedades:*

- *la variable aleatoria $L_{ij}$, que representa la cantidad de individuos de tipo $j$ que tiene un individuo de tipo $i$, tiene la misma distribución que $\sum_{k=0}^{T_1-1} \mathbf{1}\{X^{(i)}(k) = j\}$, donde el superíndice $(i)$ denota que $X(0) = i$.*

- $$m_{ij} := \mathbb{E}(L_{ij}) = \sum_{k=0}^{\infty} (P^k)_{ij}, \quad \mathbb{E}(Z(t)) = e^{t(M-Id)}.$$

- *La matriz $M$ es $\frac{R-1}{R}$-positiva y la matriz $\mathbb{E}(Z(t))$ resulta $\left(e^{-t(\frac{1}{R-1})}\right)$-positiva, ambas matrices con los mismos autovectores de la matriz $P$ asociados a $1/R$.*

- *Si $\tau_n$ es la n-ésima reproducción, entonces*

$$Z_j(\tau_n) = \xi_j(T_n), \quad \forall j, \quad y \quad \tau_n \overset{n\to\infty}{\longrightarrow} \infty.$$

*Más aún, si $P$ es irreducible y/o aperiódica también lo es $\mathbb{E}(Z(t))$.*

*Demostración.* Damos la construcción del proceso $\mathbf{Z}$. Si la cadena $V$ comienza en el estado $i$, el proceso de ramificación $\mathbf{Z}$ empieza con un individuo de tipo $i$. La cadena $X$ pasea por los estados usando la matriz de transición $P_0$, hasta llegar al estado absorbente (tiempo $T_1$). El proceso $V$ es igual a $X$ hasta $T_1$ y el proceso $\mathbf{Z}$ lo construimos de forma tal que el individuo de tipo $i$ tiene tantos hijos de tipo $j$ como veces que la cadena $V$ pasó por el estado $j$ antes del tiempo $T_1$. Es decir, el individuo de tipo $i$ tiene $L_{ij}$ descendientes de tipo $j$, donde $L_{ij}$ tiene la misma distribución que $\sum_{k=0}^{T_1-1} \mathbf{1}\{X^{(i)}(k) = j\}$.

Notemos que cuando se reproduce un individuo, este muere y nace al menos un individuo del mismo tipo. Nos referimos a este hecho como que el individuo que se reproduce no se muere.

Se elige un individuo de tipo $j$ con probabilidad $\frac{\xi_j(T_1)}{T_1}$, y se corre un nuevo proceso $X$ el cual se concatena a $V$ salvo el estado inicial y el estado 0. En el proceso de ramificación, se elige entre todos los individuos que están vivos cuál es el que se reproduce y siendo que todos los individuos esperan un tiempo exponencial de igual parámetro, se elige al individuo $j$ también con probabilidad $\frac{\xi_j(T_1)}{T_1}$.

Ambos procesos continúan con este mecanismo, con lo cual la cantidad de individuos de tipo $j$ en la $n$-ésima reproducción es igual a la cantidad de veces que la cadena $V$ pasó por el estado $j$ a tiempo $T_n - 1$. Notamos para cada $n \in \mathbb{N}$, $\tau_n$ el tiempo de la $n$-ésima reproducción, luego

$$Z_j(\tau_n) = \xi_j(T_n).$$

El proceso de ramificación $\mathbf{Z}$ tiene como matriz de medias a

$$\mathbb{E}(L_{ij}) = \mathbb{E}_i(\sum_{k=0}^{T_1-1} \mathbf{1}\{X^{(i)}(k) = j\}) = \mathbb{E}_i(\sum_{k=0}^{\infty} \mathbf{1}\{X(k) = j, T_1 > k\}) = \sum_{k=0}^{\infty} (P^k)_{ij}. \tag{4.1}$$

Siendo que $P$ es una matriz subestocástica y $R$-positiva, resulta $R > 1$ y por lo tanto $m_{ij} = \sum_{k=0}^{\infty} (P^k)_{ij}$. Más aún la matriz $M$ resulta un operador acotado. En efecto, sea $\epsilon > 0$ tal que $1/R + \epsilon < 1$, y $\|\cdot\|$ una norma de operadores. Como $\lim_{k\to\infty} \|P^k\|^{1/k} = 1/R$ (Teorema de Gelfand [Lax02, Teorema 4, p.195 ]), existe $k_0$ tal que

$$\|M\| \leq \sum_{k=0}^{k_0} \|P\|^k + \sum_{k=k_0+1}^{\infty} \|P^k\| \leq \sum_{k=0}^{k_0} \|P\|^k + \sum_{k=k_0+1}^{\infty} (\frac{1}{R} + \epsilon)^k < \infty. \tag{4.2}$$

Además, $M$ es $\frac{R-1}{R}$-positiva con los mismos autovectores de la matriz $P$ asociados a $1/R$. Por la Proposición 1.32, la esperanza de la cantidad de individuos a tiempo $t$ viene dada por la matriz

$$\mathbb{E}(\mathbf{Z}(t)) = e^{t(M-Id)},$$

y resulta $\left(e^{-t(\frac{1}{R-1})}\right)$-positiva. Además $(\mathbb{E}(\mathbf{Z}(t)))_{ij}$ es finito para todo $i, j \in \mathcal{S}$ y por lo tanto el proceso de ramificación tiene una cantidad finita de individuos para todo $t > 0$, es decir el proceso no explota y $\tau_n \overset{n\to\infty}{\longrightarrow} \infty$.

□

El siguiente resultado es la prueba de que el método de AFP converge cuando se considera espacios de estados finitos. Daremos su demostración ya que será de utilidad cuando tratemos el método de AFP en espacios de estados numerables.

**Teorema 4.6.** *[AFP88, Teorema 3] Sea $X = \{X(n), n \in \mathbb{N}\}$ un proceso de Markov definido en $\mathcal{S} \cup \{0\}$, con $\mathcal{S}$ un espacio finito, 0 un estado absorbente y $P_0$ la matriz de transición. Suponemos que la matriz $P$ es irreducible y aperiódica. Entonces el proceso $V$ (Definición 4.2) verifica que cuando $n \to \infty$*

$$\frac{\sum_{k=0}^{n} \mathbf{1}\{V(k) = j\}}{n} \xrightarrow{cs} \nu_j \quad \textit{para todo} \quad j \in \mathcal{S}$$

*donde $\nu$ es la única QSD de $X$.*

*Demostración.* La matriz $P$ es irreducible y aperiódica, por el Teorema 3.9 existe una única QSD $\nu$ que verifica

$$\sum_{i \in \mathcal{S}} \nu_i p_{ij} = \frac{1}{R} \nu_i, \quad \nu_i > 0 \quad \text{y} \quad \sum_{i \in \mathcal{S}} \nu_i = 1, \tag{4.3}$$

donde $1/R$ es el autovalor mas grande de la matriz $P$ (Teorema 3.8).

Sea $\mathbf{Z}(t)$ el proceso construido en el Lema 4.5. La matriz $\mathbb{E}(Z(t))$ es irreducible, aperiódica y tiene como mayor autovalor a $e^{t(\frac{1}{R-1}-1)}$, con autovectores a derecha e izquierda $\mu$ y $\nu$ respectivamente. En [GB03, Teorema 2.1] se prueba que bajo estas condiciones se verifica que si $t \to \infty$,

$$\frac{Z_i(t)}{|\mathbf{Z}(t)|} \xrightarrow{cs} \nu_i \quad \text{para cada} \quad i \in \mathcal{S}.$$

Luego, como $\tau_n \to \infty$

$$\lim_{n \to \infty} \frac{Z_i(\tau_n)}{|\mathbf{Z}(\tau_n)|} = \nu_i,$$

de donde

$$\lim_{n \to \infty} \frac{\xi_i(T_n)}{\sum_i \xi_i(T_n)} = \lim_{n \to \infty} \frac{\xi_i(T_n)}{T_n} = \nu_i.$$

Esta es la convergencia buscada pero en lugar de para todo $n$ solo en los tiempos de absorción de la cadena. Veamos que $\frac{T_{n+1}}{T_n} \to 1$. Dados $\epsilon > 0$ e $i \in \mathcal{S}$

$$\mathbb{P}_i\left(|\frac{T_{n+1}}{T_n} - 1| > \epsilon\right) \leq \mathbb{E}_i\left(\frac{(\frac{T_{n+1}-T_n}{T_n})^2}{\epsilon^2}\right) \leq \mathbb{E}_i\left(\frac{(T_{n+1} - T_n)^2}{n^2\epsilon^2}\right) = \frac{\mathbb{E}_\xi({T_1}^2)}{n^2\epsilon^2}, \tag{4.4}$$

donde $\xi = \frac{\xi(n)}{n}$. Como el espacio de estados es finito $\mathbb{E}_i({T_1}^2)$ es finito para todo $i \in \mathcal{S}$, y

$$\mathbb{E}_\xi({T_1}^2) = \sum_{i \in \mathcal{S}} \mathbb{E}_i({T_1}^2)\xi_i \leq \max_{i \in \mathcal{S}} \mathbb{E}_i({T_1}^2) < \infty.$$

Luego obtenemos que si $n \to \infty$,

$$\frac{T_{n+1}}{T_n} \xrightarrow{cs} 1. \tag{4.5}$$

Sea $m$ tal que $T_n \leq m \leq T_{n+1}$, como $\xi(T_n) \leq \xi(m) \leq \xi(T_{n+1})$ coordenada a coordenada obtenemos

$$\frac{\xi(T_n)}{T_{n+1}} \leq \frac{\xi(m)}{m} \leq \frac{\xi(T_{n+1})}{T_n}. \tag{4.6}$$

De donde para cada $i \in \mathcal{S}$, cuando $n \to \infty$

$$\frac{\xi_i(n)}{n} \xrightarrow{cs} \nu_i.$$

Por lo tanto la distribución empírica de la cadena $V$ converge casi seguramente a la distribución cuasiestacionaria de $X$. □

## 4.2. Método AFP en espacios de estados numerables.

Dada una cadena de Markov con estado absorbente, siempre es posible construir el proceso $V$ (Definición 4.2) generado por AFP y construir $\mathbf{Z}(t)$ el proceso de ramificación multitipo como en el Lema 4.5. Para garantizar la convergencia del método AFP es necesario que la cantidad de individuos de cada tipo sobre el total de la población converja cuando $t \to \infty$ a una medida de probabilidad. Para lograr esto para una cadena con espacio de estados numerable utilizaremos el Teorema 1.35 donde se prueba la convergencia en probabilidad de la proporción de individuos, y así se obtiene en ciertos casos, la convergencia en probabilidad del método AFP. La prueba de convergencia del método para casos más generales es un problema abierto.

**Teorema 4.7.** *Sea $X = \{X(n), n \in \mathbb{N}\}$ un proceso de Markov definido en $\mathcal{S} \cup \{0\}$, con $\mathcal{S}$ numerable, 0 un estado absorbente y $P_0$ la matriz de transición. Suponemos que la matriz $P$ es irreducible, aperiódica y cumple la condición $\alpha > C$ (Definición 3.18). Entonces el proceso $V$ (Definición 4.2) verifica que cuando $n \to \infty$*

$$\frac{\sum_{k=0}^{n} \mathbf{1}\{V(k) = j\}}{n} \xrightarrow{P} \nu_j \quad \textit{para todo} \quad j \in \mathcal{S}$$

*donde $\nu$ es la única QSD de $X$.*

*Demostración.* Por el Teorema 3.19, la matriz $P$ es $R$-positiva y existen autovectores a derecha e izquierda $\mu$ y $\nu$ asociados a $1/R \in (0,1)$ que verifican

$$\sum_{j \in \mathcal{S}} \nu_j = 1 \quad \text{y} \quad \tilde{c} \leq \mu_i \leq \tilde{C} \quad \forall i \in \mathcal{S} \tag{4.7}$$

para constantes positivas $\tilde{c}, \tilde{C}$, donde $\nu$ es la única QSD.

Por el Lema 4.5 existe el proceso de ramificación $\mathbf{Z}$ con matriz de medias $M = \{m_{ij}, i, j \in \mathcal{S}\}$ con $m_{ij} = \mathbb{E}(L_{ij}) = \sum_{k=0}^{\infty} (P^k)_{ij}$. La matriz $M$ resulta no negativa, irreducible, aperiódica, $\gamma$-positiva con

$$\gamma = \frac{R-1}{R} \in (0,1) \tag{4.8}$$

y con los mismos autovectores $\mu$ y $\nu$ asociados al autovalor $1/R$ de la matriz $P$. Además para cada $t \geq 0$ la matriz

$$C(t) := \mathbb{E}(Z(t)) = e^{t(M-Id)}$$

es irreducible, $\gamma_t$-positiva ($\gamma_t = e^{-t(\frac{R}{R-1}-1)} \in (0,1)$) y verifica que $(C^n(t))_{ij}$ es finito para todo $n \in \mathbb{N}, i, j \in \mathcal{S}$. Los autovectores asociados a $1/\gamma_t$ son los autovectores de la matriz $M$ asociados a $\gamma$.

Notemos que existe un estado $l \in \mathcal{S}$ tal que $\alpha(l) > 0$, pues la matriz $P$ cumple la condición $\alpha > C$. Por la irreducibilidad de $P$ existe $k \in \mathbb{N}$ tal que $(P^k)_{l0} > 0$. Consideramos una variable aleatoria $\hat{L}$ con distribución geométrica de parámetro $\alpha(l)(P^k)_{l0} > 0$. Sea $T_1^{(i)}$ la variable aleatoria que cuenta la cantidad de estados por los que pasa la cadena $X$ antes de ser absorbida, suponiendo $X(0) = i$. Luego la variable $T_1^{(i)}$ la podemos acotar estocásticamente por la variable $L := (k+1)\hat{L}$. Cada intento de éxito de la variable $\hat{L}$ equivale a $k+1$ pasos del proceso $X$. Para todo estado inicial $i \in \mathcal{S}$ obtenemos

$$T_1^{(i)} \preceq L. \tag{4.9}$$

La cantidad total de descendientes que tiene un individuo de tipo $i$ es $\sum_{j\in\mathcal{S}} L_{ij}$. Por como definimos estas variables aleatorias tenemos

$$\sum_{j\in\mathcal{S}} L_{ij} = \sum_{j\in\mathcal{S}} \sum_{k=0}^{T_1-1} \mathbf{1}\{X^{(i)}(k) = j\} = T_1^{(i)}.$$

Luego, por (4.9) se tiene que para todo $i \in \mathcal{S}$

$$\mathbb{E}(\sum_{j\in\mathcal{S}} L_{ij}) = \mathbb{E}_i(T_1) \leq \mathbb{E}(L) < \infty.$$

Siendo que la variable aleatoria $L$ tiene segundo momento finito y los autovectores $\mu$ y $\nu$ verifican (4.7), por la Proposición 1.37 vale que dado $\delta > 0$

$$\sum_{j\in\mathcal{S}} \mathbb{E}_j\left((\mathbf{Z}(\delta)\mu)^2\right)\nu_j < \infty. \tag{4.10}$$

Veamos que existe un tipo de individuo para el cual la cantidad de individuos tiende a infinito, para ello observemos que sin importar qué tipo de individuo se reproduce hay probabilidad al menos $\alpha(l) > 0$ de tener un individuo de tipo $l$. Como ya mencionamos, cuando un individuo se reproduce decimos que este no se muere ya que deja siempre un descendiente de su tipo. Con lo cual, diremos que en la generación $n$ hay un nuevo individuo de tipo $l$ cuando se reproduce un individuo de otro tipo y este deja al menos un descendiente de tipo $l$ o se reproduce un individuo de tipo $l$ y este (además de continuar vivo) deja otro individuo de tipo $l$. Definimos para cada $n \in \mathbb{N}$ los eventos

$$A_n := \{ \text{ hay al menos un nuevo individuo de tipo } l \text{ en la generación } n\}$$

y la $\sigma$-álgebra generada por los individuos que existieron hasta la generación $n$,

$$\mathcal{F}_n := \sigma(\{X^k, \quad 1 \leq k \leq n\}).$$

Claramente $A_n \in \mathcal{F}_n$, notemos que

$$\begin{aligned}\mathbb{P}(A_n|\mathcal{F}_{n-1}) &= \mathbb{P}(\text{algún individuo de la generación } n-1 \text{ tiene un descendiente de tipo } l)\\ &\geq \alpha(l).\end{aligned}$$

Luego $\sum_{n\in\mathbb{N}} \mathbb{P}(A_n|\mathcal{F}_{n-1}) = \infty$ y por el Teorema de Borel Cantelli condicionado [Dur96, Teorema 5.3.2], se obtiene que con probabilidad 1 el evento $A_n$ ocurre infinitas veces, es decir, en infinitas generaciones habrá un nuevo individuo de tipo $l$, siendo que en este proceso los individuos no se mueren, se verifica que cuando $t \to \infty$,

$$\mathbb{P}_i(Z_l(t) \to \infty) = 1.$$

Consideramos el proceso discretizado $\mathbf{Y} = (\mathbf{Y}(n))_{n\in\mathbb{N}}$ asociado al proceso $\mathbf{Z}$ (ver Lema 1.33). Este proceso verifica las hipótesis del Teorema 1.27, y además $\mathbb{P}_i(Y_l(n) \to \infty) = 1$, con lo cual por la Proposición 1.30

$$\mathbb{P}_i(W = 0) = 0 \text{ para todo } i \in \mathcal{S}. \tag{4.11}$$

Finalmente el proceso $\mathbf{Z}$ verifica (4.7), (4.8), (4.10), (4.11) y por el Teorema 1.35 cuando $t \to \infty$ se tiene

$$\frac{\mathbf{Z}(t)}{|\mathbf{Z}(t)|} \xrightarrow{P} \nu$$

y cuando $n \to \infty$

$$\frac{\mathbf{Z}(\tau_n)}{|\mathbf{Z}(\tau_n)|} \xrightarrow{P} \nu \quad \text{pues} \quad \tau_n \xrightarrow{n\to\infty} \infty.$$

Siendo que $\mathbf{Z}(\tau_n) = \xi(T_n)$

$$\lim_{n\to\infty} \frac{\xi_i(T_n)}{\sum_i \xi_i(T_n)} = \lim_{n\to\infty} \frac{\xi_i(T_n)}{T_n} = \nu_i.$$

Siendo que $T_1^{(i)} \preceq L$, se verifica que para cualquier estado inicial $i$

$$\mathbb{E}_i(T_1^2) \leq \mathbb{E}(L^2) < \infty,$$

es decir, $\sup_{i\in\mathcal{S}} \mathbb{E}_i(T_i^2)$ es finito. De donde de forma análoga a lo probado en el método de AFP para espacios finitos, de (4.4), (4.5) y (4.6) se obtiene que toda la sucesión converge. Es decir, para todo $i \in \mathcal{S}$ cuando $n \to \infty$

$$\frac{\xi_i(n)}{n} = \frac{\sum_{k=0}^{n-1} \mathbf{1}\{V(k) = i\}}{n} \; oversetP{\longrightarrow}\nu_i.$$

Por lo tanto la distribución empírica de la cadena $V$ converge en probabilidad a la única distribución cuasiestacionaria de $X$. $\square$

## 4.3. Procesos de Galton-Watson

En esta sección daremos un nuevo método, similar a AFP, para aproximar QSD y lo aplicamos a procesos de Galton-Watson subcríticos. Cabe destacar que las pruebas de convergencia dadas en las secciones anteriores no contienen a estos procesos dado que en el proceso de ramificación multitipo asociado la cantidad de hijos que puede tener cada individuo tiende a infinito con el tipo. Para este nuevo método probaremos la convergencia del mismo para el caso de Galton-Waston.

La principal diferencia de este nuevo método consiste en que el proceso de ramificación multitipo asociado es un proceso a tiempo discreto en lugar de ser a tiempo continuo. Para este último podemos aplicar el Teorema 1.29.

**Definición 4.8.** *Dada una cadena de Markov $X$ con estado absorbente, construimos un proceso de ramificación multitipo a tiempo discreto (ver Definición 1.12), $\mathbf{Y}$, de la siguiente manera.*

*La cantidad de veces que la cadena $X$ inicializada en $i$ pasa por el estado $j$ es exactamente la cantidad de descendientes de tipo $j$ que deja un individuo de tipo $i$. Es decir, el individuo de tipo $i$ tiene $L_{ij}$ descendientes de tipo $j$, donde $L_{ij}$ tiene la misma distribución que $\sum_{k=0}^{T-1} \mathbf{1}\{X^{(i)}(k) = j\}$. Por lo tanto,*

$$m_{ij} := \mathbb{E}(L_{ij}) = \mathbb{E}_i(\sum_{k=0}^{T_1-1} \mathbf{1}\{X^{(i)}(k) = j\}) = \mathbb{E}_i(\sum_{k=0}^{\infty} \mathbf{1}\{X(k) = j, T_1 > k\}) = \sum_{k=0}^{\infty}(P^k)_{ij}. \tag{4.12}$$

*Para cada individuo se corre una cadena $X$ independiente, por lo que cada individuo tiene descendientes independientemente del resto de los individuos lo que implica que las variables $L_{ij}$ son independientes entre sí. Notemos que de la ecuación (4.2), $M$ resulta un operador acotado y $\|M^n\| \leq \|M\|^n < \infty$.*

**Proposición 4.9.** *Sea $X$ una cadena de Markov con estado absorbente y matriz de transición $P_0$ con $P$ $R$-positiva. Sea $\mathbf{Y}$ el proceso de ramificación construido en la Definición 4.8. Si el proceso $\mathbf{Y}$ verifica que cuando $n \to \infty$ $\frac{\mathbf{Y}(n)}{|\mathbf{Y}(n)|} \xrightarrow{P} \nu$. Entonces $\nu$ es la QSD minimal de $X$.*

*Demostración.* Sea $\nu'$ el autovector a izquierda de la matriz $P$ asociado a $1/R$, de la Observación 3.17 $\nu'$ es la QSD minimal de $X$. Por construcción $\mathbf{Y}$ es $\frac{R-1}{R}$-positiva, con los mismos autovectores de $P$. Luego $\nu = \nu'$. □

Apliquemos este método a los procesos de ramificación llamados Galton-Watson. En la siguiente definición fijaremos la notación que usaremos en lo que resta de la sección. Siendo que hay varios procesos de ramificación involucrados, cuando nos estemos refiriendo al proceso de Galton-Watson hablaremos de partículas así podremos distinguirlos de los individuos de los procesos de ramificación multitipo asociado al método de simulación.

**Definición 4.10.** *Definimos $X = \{X(n), n \in \mathbb{N}\}$ un proceso de Galton-Watson (Definición 1.1). Suponemos que el número de ramificaciones que tiene una partícula es una* *variable aleatoria $L$ con distribución $(p_i)_{i\in\mathbb{N}_0}$, donde $p_i$ es la probabilidad de que una partícula se ramifique en $i$ partículas, con $i \in \mathbb{N}_0$. Para cada $j \in \mathbb{N}$, $k \in \mathbb{N}_0$, $L_j^k$ representa el número de ramificaciones de la $j$-ésimo partícula de la $k$-ésima generación. Asumimos que $\{L_j^k,\ k \geq 0,\ j \geq 1\}$ son variables aleatorias i.i.d., todas con la misma distribución que $L$, donde $\mathbb{P}(L = i) = p_i \quad \forall i \in \mathbb{N}_0$. Suponemos que $p_0 + p_1 < 1$.*

*Consideramos el caso subcrítico que es cuando el proceso se extingue con probabilidad 1 (Lema 1.4, Teorema 1.5), para ello suponemos $m := \mathbb{E}(L) < 1$. Sea $P_0$ la matriz de transición, y $P$ la matriz de transición sin la fila y la columna correspondientes al 0,*

$$p_{ij} = \mathbb{P}(X(n+1) = j | X(n) = i), \text{ para todo } i, j \in \mathbb{N}.$$

*Sea $T$ el tiempo de absorción del proceso,*

$$T := \min_{n\in\mathbb{N}}\{X(n) = 0\}.$$

*Sea $H$ la cantidad total de partículas creadas,*

$$H := \sum_{n=0}^{T} X(n).$$

*Sea $H_i$ la cantidad de veces que el proceso se encontró en el estado $i$,*

$$H_i := \sum_{n=0}^{T} \mathbf{1}\{X(n) = i\}.$$

Notemos que

$$H = \sum_{j\in\mathbb{N}} jH_j.$$

Nos interesa estudiar el total de partículas $H$ del proceso, para ello consideramos su función generadora

$$f_H(s) := \mathbb{E}(s^H).$$

Damos un resultado que relaciona la función generadora de $H$ con la función generadora de $L$.

**Teorema 4.11** (Teorema 3.2, [VDH16])**.** *Sean $X$ un proceso de Galton-Watson como en la Definición 4.10 y $H = \sum_{n\in\mathbb{N}_0} X(n)$. Si denotamos $f_L(s)$ y $f_H(s)$ a las funciones generadoras de $L$ y $H$ respectivamente, se obtiene la siguiente relación*

$$f_H(s) = sf_L(f_H(s)) \qquad s \in [0, 1]. \tag{4.13}$$

*Demostración.* Primero notemos que como el proceso Galton-Watson comienza con una sola partícula $f(s) = \mathbb{E}(s^X) = \mathbb{E}(s^L)$. Si la partícula inicial tiene $j$ descendientes, cada una de estas $j$ partículas tiene asociada una variable aleatoria $H^{(l)}, l \in \{1, \ldots, j\}$ que cuenta la cantidad total de partículas que deja el $l$-ésimo descendiente de la partícula inicial, estas variables son i.i.d. con la misma distribución que $H$. Notemos que en este caso $H = 1 + \sum_{l=1}^{j} H^{(l)}$. Luego

$$f_H(s) = \mathbb{E}(s^H) = \sum_{l=0} p_l \mathbb{E}(s^{1+\sum_{l=1}^{j} H^{(l)}} | X(1) = l) = s \sum_{l=0} p_l \mathbb{E}(s^H)^l = s f_L(f_H(s)).$$

□

**Corolario 4.12.** *Bajo las condiciones del teorema anterior se obtiene*

$$\mathbb{E}(H) = \frac{1}{1-m} \qquad \mathbb{E}(H^2) = \frac{1+m}{(1-m)^2} + \frac{\mathbb{E}(L^2) - m}{(1-m)^3}.$$

*Demostración.* Los resultados se obtienen de derivar la ecuación (4.13) y evaluar en $s = 1$. La primera igualdad

$$f'_H(s) = f_L(f_H(s)) + s f'_L(f_H(s)) f'_H(s) \Rightarrow \mathbb{E}(H) = 1 + m\mathbb{E}(H).$$

La segunda igualdad

$$f''_H(s) = f'_L(f_H(s)) f'_H(s) + s f'_L(f_H(s)) f''_H(s) + f'_H(s) f'_L(f_H(s)) + s f''_L(f_H(s))(f'_H(s))^2 \Rightarrow$$

$$\mathbb{E}(H(H-1)) = 2\mathbb{E}(H)m + m\mathbb{E}(H(H-1)) + \mathbb{E}^2(H)\mathbb{E}(L(L-1)) \Rightarrow$$

$$\mathbb{E}(H^2) - \mathbb{E}(H) = \mathbb{E}(H)m + m\mathbb{E}(H^2) + \mathbb{E}^2(H)\mathbb{E}(L(L-1)) \Rightarrow$$

$$\mathbb{E}(H^2) = \frac{1+m}{(1-m)^2} + \frac{\mathbb{E}(L^2) - m}{(1-m)^3}$$

□

**Lema 4.13.** *Sea $P_0$ la matriz de transición de un proceso de Galton-Watson, y $p_i$ la probabilidad de que un individuo tenga i descendientes. La matriz $P$ es irreducible si y sólo si $p_0 > 0$, $p_1 > 0$ y $p_j > 0$ para algún $j > 1$. Bajo esta condición también es aperiódica.*

*Demostración.* El resultado se deduce de la definición de la matriz $P$, $p_{ij} = \mathbb{P}(X(1) = j | X(0) = i)$. □

**Lema 4.14.** *[SVJ66, Lema iii]Sean $P_0$ la matriz de transición de un proceso de Galton-Watson, $p_i$ la probabilidad de que un individuo tenga i descendientes, y $L$ una variable aleatoria con distribución $(p_i)_{i\in\mathbb{N}_0}$. Si $P$ es irreducible y $\mathbb{E}(L) < 1$, entonces $P$ es $(1/\mathbb{E}(L))$-positiva si y sólo si $\sum_{j\in\mathbb{N}} j \ln(j) p_j < \infty$.*

**Observación 4.15.** *Sea $X$ un proceso de Galton-Watson con matriz de transición $P$. El autovector a derecha asociado a $\mathbb{E}(L)$ es $\mu_i = i$,*

$$(P\mu)_i = \sum_{j\in\mathbb{N}} p_{ij} j = \mathbb{E}(X(1)|X(0) = i) = i\mathbb{E}(X(1)|X(0) = 1) = i\mathbb{E}(L).$$

**Corolario 4.16.** *Si la matriz de transición $P$ es irreducible y la variable $L$ con distribución $(p_i)_{i\in\mathbb{N}_0}$ verifica que $\mathbb{E}(L) < 1$ y $\sum_{j\in\mathbb{N}} j\ln(j)p_j < \infty$, entonces $\sum_{j\in\mathbb{N}} j\nu_j < \infty$ con $\nu$ el autovector a izquierda de $P$ asociado a $\mathbb{E}(L)$.*

**Teorema 4.17.** *[ST85, Teorema 3] Si $P$ es irreducible y aperiódica, para $k \geq 2$ se obtiene*

$$\mathbb{E}(L^k) < \infty \quad \textit{si y sólo si} \quad \sum_j \nu_j j^k < \infty$$

*donde $\nu$ es la QSD minimal del proceso $X$.*

A continuación damos el resultado principal de esta sección, mostramos que para ciertos procesos de Galton-Watson el proceso de ramificación construido en la Definición 4.8, verifica que la proporción de individuos converge a la QSD minimal del proceso de Galton-Watson.

**Teorema 4.18.** *Sea $X = \{X(n), n \in \mathbb{N}\}$ el proceso de Galton-Watson de la Definición 4.10, suponemos $p_0, p_1$, $p_l > 0$ para algún $l > 1$ y tal que $\mathbb{E}(L^2) < \infty$. Sea $\mathbf{Y}$ el proceso de ramificación multitipo de la Definición 4.8, entonces cuando $n \to \infty$*

$$\frac{Y_i(n)}{|\mathbf{Y}(n)|} \xrightarrow{P} \nu_i,$$

*donde $\nu$ es la QSD minimal del proceso $X$.*

*Demostración.* El proceso de Galton-Watson es subcrítico con lo cual $\mathbb{E}(L) < 1$, además la matriz $P$ es irreducible y aperiódica con lo cual resulta $1/\mathbb{E}(L)$-positiva, con $\nu$ y $\mu = (1, 2, \ldots)$ autovectores a izquierda y derecha asociados al autovalor $\mathbb{E}(L)$ respectivamente. Siendo que $\mathbb{E}(L^2) < \infty$, del Corolario 4.12 obtenemos que $\mathbb{E}(H^2) < \infty$.

Veamos que el proceso $\mathbf{Y}$ cumple las hipótesis del Teorema 1.29. Notemos que la variable $Y_i(1)$ cuenta la cantidad de individuos de tipo $i$ en la primera generación que es igual a contar la cantidad de veces que el proceso $X$ estuvo en el estado $i$. Con lo cual la variable aleatoria $Y_i(1)$ tiene la misma distribución que $H_i$,

$$Y_i(1) \sim H_i, \qquad \forall i \in \mathbb{N}.$$

Por otro lado, si el proceso $X$ comienza con $j$ partículas, por la desigualdad de Holder se obtiene

$$\mathbb{E}_j(H^2) = \mathbb{E}_1\left( (\sum_{l=1}^{j} H^{(l)})^2 \right) \leq j\mathbb{E}_1(H^2)$$

donde $H^{(l)}$ son v.a.i.i.d. con la misma distribución que $H$. Notemos que

$$\mathbb{E}_j\left[ \left( \sum_{i\in\mathbb{N}} Y_i(1)i \right)^2 \right] = \mathbb{E}_j\left[ \left( \sum_{i\in\mathbb{N}} iH_i \right)^2 \right] = \mathbb{E}_j(H^2).$$

Luego, se verifica la condición (1.7):

$$\sum_{j\in\mathbb{N}} \mathbb{E}_j \left[ \left( \sum_{i\in\mathbb{N}} Y_i(1) i \right)^2 \right] \nu_j \leq \sum_{j\in\mathbb{N}} j^2 \mathbb{E}_1(H^2)\nu_j < \infty,$$

y el proceso $\mathbf{Y}$ esta bajo las condiciones del Teorema 1.27.

Definimos para cada $n \in \mathbb{N}$ los eventos

$$A_n := \{ \text{ hay al menos un nuevo individuo de tipo 1 en la generación } n \}$$

y la $\sigma$-álgebra generada por los individuos que existieron hasta la generación $n$,

$$\mathcal{F}_n := \sigma(\{\mathbf{Y}(k), \quad 1 \leq k \leq n\}).$$

Al igual que en el método de AFP consideramos que los individuos que se reproducen no se mueren, con lo cual decimos que un individuo de tipo 1 tiene un nuevo individuo de tipo 1, cuando deja otro descendiente de su tipo. Claramente $A_n \in \mathcal{F}_n$ y

$$\mathbb{P}(A_n|\mathcal{F}_{n-1}) = \mathbb{P}\begin{pmatrix} \text{un individuo de la generación } n-1 \\ \text{tiene un nuevo descendiente de tipo 1} \end{pmatrix} \geq p_1 > 0,$$

pues el individuo inicial de tipo 1 nunca se muere y como consideramos un proceso de ramificación a tiempo discreto todos los individuos se reproducen en todas las generaciones, en particular el individuo inicial de tipo 1 se reproduce en todas las generaciones. Luego $\sum_{n\in\mathbb{N}} \mathbb{P}(A_n|\mathcal{F}_{n-1}) = \infty$ y por el Teorema de Borel Cantelli condicionado [Dur96, Teorema 5.3.2], se obtiene que con probabilidad 1 el evento $A_n$ ocurre infinitas veces, es decir, en infinitas generaciones habrá un nuevo individuo de tipo 1. Siendo que los individuos no se mueren, se verifica que cuando $n \to \infty$

$$\mathbb{P}_1(Y_1(n) \to \infty) = 1.$$

Por la Proposición 1.30 se obtiene

$$\mathbb{P}_i(W = 0) = 0 \text{ para todo } i \in \mathbb{N}.$$

Finalmente el Teorema 1.29 implica que cuando $n \to \infty$

$$\frac{\mathbf{Y}(n)}{|\mathbf{Y}(n)|} \xrightarrow{P} \nu.$$

$\square$

# Capítulo 5

# Aproximación de QSD vía desigualdad de Holley

Técnicas de acoplamiento y dominación estocástica han sido ampliamente explotadas para tratar diversos problemas [GHM01, Capítulo 4]. La desigualdad de Holley [Hol74], es una condición suficiente para garantizar dominación estocástica en espacios producto. En este capítulo la utilizaremos, adaptada al espacio de trayectorias, [FR15] para probar convergencia monótona de distribuciones cuasiestacionarias.

La idea de esta versión es considerar procesos con estados absorbentes y obtener condiciones suficientes bajo las cuales se pueda garantizar la dominación estocástica de las trayectorias condicionadas.

Los procesos de nacimiento y muerte son cadenas de Markov definidas en subconjuntos de $\mathbb{N}$ en los que sólo se producen transiciones de un estado a si mismo y a sus contiguos (anterior y posterior). Trabajaremos con procesos de nacimiento y muerte $X^K$ definidos en espacios finitos $\{0, \ldots, K\}$ con $K \in \mathbb{N}$ donde supondremos que el 0 es el estado absorbente y que hay probabilidad positiva de ir de $K$ al 0. Bajo ciertas hipótesis, [DS65], estos procesos tienen una única QSD que la denotaremos $\nu^K$.

También consideraremos procesos de nacimiento y muerte $X$ definidos en $\mathbb{N}_0$. Estos procesos de nacimiento y muerte pueden tener infinitas distribuciones cuasiestacionarias, una o ninguna [Cav78, FMP91]. Denotaremos $\nu_{\text{mín}}$ a la QSD minimal.

El resultado principal del capítulo es el Teorema 5.19 que prueba (bajo ciertas hipótesis sobre los parámetros) que las QSD de los procesos definidos en $\{0, \ldots, K\}$ convergen a la QSD minimal del proceso de nacimiento y muerte en $\mathbb{N}$. Si bien varias de estas convergencias pueden probarse con un cálculo explícito de autovectores y autovalores, lo relevante de este capítulo son los métodos utilizados basados en la dominación estocástica que permiten extender los resultados a casos donde este tipo de cálculo no está disponible.

En [BH00] se analizan procesos de Markov a tiempo continuo $R$-positivos en espacios infinitos y se muestra que las QSD en espacios finitos aproximan a la QSD minimal de estos procesos. Los procesos de nacimiento y muerte considerados en este capítulo no son $R$-positivos, con lo cual, los resultados de [BH00] no son aplicables en nuestro contexto.

Llamemos $\delta_1$ a la medida inicial que concentra toda su masa en 1 y denotemos $\varphi^{K,\delta_1}(n)$

a la evolución condicionada a la no absorción del proceso finito en $\{0,\dots,K\}$ a tiempo $n$. Análogamente, denotemos $\varphi^{\delta_1}(n)$ a la evolución condicionada a la no absorción en el proceso en $\mathbb{N}_0$. La idea principal del capítulo se puede resumir con las siguientes convergencias,

$$\begin{tikzcd} \varphi^{K,\delta_1}(n) \arrow[r] \arrow[d, "5.16"'] & \varphi^{\delta_1}(n) \arrow[d, "5.18"] \\ \nu^K \arrow[r, "5.19"'] & \nu_{min}. \end{tikzcd}$$

La convergencia $\lim_{K\to\infty}\varphi^{K,\delta_1}(n) = \varphi^{\delta_1}(n)$ se deduce de que el proceso $X^K$ converge a $X$ cuando $K\to\infty$. De [DS65] se obtiene que el proceso $X^K$ tiene una única QSD y $\lim_{n\to\infty}\varphi^{K,\delta_1}(n) = \nu^K$. Esta convergencia la demostramos nuevamente usando dominación estocástica en el Lema 5.16. Por otro lado el proceso $X$ (Definición 5.24) tiene infinitas QSD y además $\lim_{n\to\infty}\varphi^{\delta_1}(n) = \nu_{\text{mín}}$ (ver en [FR15] o 5.18). El resultado principal de este capítulo es el Teorema 5.19 donde se obtiene que la sucesión $(\nu^K)_{K\in\mathbb{N}}$ es monótona creciente (con el orden de la dominación estocástica) y se prueba que $\lim_{K\to\infty}\nu^K = \nu_{min}$.

Comenzamos dando una introducción a la dominación estocástica y la desigualdad de Holley. Definimos el espacio de las trayectorias de las cadenas y mostramos la versión de la desigualdad de Holley adaptada para el espacio de las trayectorias.

Primero consideramos procesos de nacimiento y muerte en espacios finitos y luego en $\mathbb{N}$, ambos aperiódicos. Finalmente consideramos el caso periódico donde volvemos a probar las convergencias.

## 5.1. Dominación estocástica

En esta sección trabajaremos con dominación estocástica enfocado a medidas de probabilidad.

**Definición 5.1.** *Sean $X, X'$ dos variables aleatorias definidas en $\mathcal{S}$, con distribución $\mu$ y $\mu'$ respectivamente. Si $A$ es un evento de $\mathcal{S}$, denotamos $\mu_A := \mathbb{P}(X\in A)$ y $\mu'_A := \mathbb{P}(X'\in A)$.*

*Un acoplamiento $P$ de $X, X'$ o de $\mu,\mu'$ es una medida en $\mathcal{S}\times\mathcal{S}$ con marginales $\mu$ y $\mu'$, que para todo evento $A\subset\mathcal{S}$*

$$\mathbb{P}((\xi,\xi') : \xi\in A) = \mu(A)$$

*y*

$$\mathbb{P}((\xi,\xi') : \xi'\in A) = \mu'(A).$$

Pensamos en un acoplamiento como una redefinición de las variables aleatorias en un nuevo espacio de probabilidad común, de modo que preservan sus distribuciones.

**Definición 5.2.** *Supongamos que $\mathcal{S}$ es un subconjunto cerrado de $\mathbb{R}$. Definimos el orden parcial en el espacio producto $\mathcal{S}^{\mathbb{Z}}$*

1. *notamos* $\eta \preceq \eta'$ *si* $\eta(x) \leq \eta'(x)$ *para todo* $x \in \mathbb{Z}$,

2. *sea* $f : \mathcal{S}^{\mathbb{Z}} \to \mathbb{R}$, *la función* $f$ *se dice creciente si:* $\eta \preceq \eta'$ *implica que* $f(\eta) \leq f(\eta')$,

3. *un evento* $A$ *se dice creciente si* $\mathbf{1}\{A\}$ *es una función creciente.*

**Definición 5.3.** *Dadas dos medidas* $\mu$ *y* $\mu'$ *en* $\mathcal{S}^{\mathbb{Z}}$ *decimos que* $\mu$ *está dominada estocásticamente por* $\mu'$, *denotamos* $\mu \preceq \mu'$ *si* $\mu(f) \leq \mu'(f)$ *para toda* $f$ *creciente.*

**Observación 5.4.** *Notemos que en una dimensión* ($\mathcal{S} \subset \mathbb{R}$), *la Definición 5.3 es equivalente a*

$$\mu([r, +\infty)) \leq \mu'([r, +\infty)) \qquad \forall \quad r \in \mathbb{R}.$$

El siguiente resultado caracteriza a la dominación estocástica en términos de acoplamientos.

**Teorema 5.5.** *[Str65] Para dos medidas* $\mu$ *y* $\mu'$ *definidas en* $\mathcal{S}^{\mathbb{Z}}$ *son equivalentes:*

1. $\mu \preceq \mu'$,

2. *para cualquier función creciente, continua y acotada* $f$, $\mu(f) \leq \mu'(f)$,

3. *existe un acoplamiento que verifica que* $\mathbb{P}(X \preceq X') = 1$.

*Demostración.* Puede verse un esquema de la demostración en [GHM01, Teorema 4.6]. □

**Definición 5.6.** *Dados dos elementos de probabilidad positiva en* $\mathcal{S}^{\mathbb{Z}}$, *decimos que los elementos se conectan si sólo difieren en una coordenada. Una medida de probabilidad* $\mu$ *en* $\mathcal{S}^{\mathbb{Z}}$ *se dice irreducible, si el conjunto* $\{\eta \in \mathcal{S}^{\mathbb{Z}} : \mu(\eta) > 0\}$ *es conexo, es decir dados dos elementos con probabilidad positiva se conectan a través de cambios de coordenadas sin pasar por elementos con probabilidad 0.*

Damos a continuación la desigualdad de Holley que es una condición suficiente para obtener dominación estocástica. En la siguiente sección mencionamos otra versión de esta desigualdad adaptada para el contexto en el que trabajaremos.

**Teorema 5.7** (Desigualdad de Holley, [Hol74])**.** *Sean* $\mathcal{S}$ *y* $\lambda$ *subconjuntos finitos de* $\mathbb{R}$. *Sean* $\mu$ *y* $\mu'$ *medidas en* $\mathcal{S}^{\lambda}$. *Supongamos que* $\mu'$ *es una medida irreducible en* $\mathcal{S}^{\lambda}$ *que asigna probabilidad positiva al elemento maximal de* $\mathcal{S}^{\lambda}$ *(respecto a* $\preceq$*). Si*

$$\mu(X(j) \geq a | X = \xi, \text{ fuera de } j) \leq \mu'(X'(j) \geq a | X = \eta, \text{ fuera de } j)$$

*para todo* $j \in \lambda$, $a \in \mathcal{S}$ *y* $\eta, \xi \in \mathcal{S}^{\lambda - \{j\}}$ *tal que* $\xi \preceq \eta$, $\mu(X = \xi \text{ fuera de } j) > 0$ *y* $\mu'(X' = \eta \text{ fuera de } j) > 0$. *Entonces*

$$\mu \preceq \mu'.$$

### 5.1.1. Espacio de trayectorias

Sea $X = \{X(n), n \in \mathbb{N}\}$ una cadena de Markov en el espacio $\mathcal{S}\cup\{0\}$ finito o numerable, con 0 el único estado absorbente. Consideramos la matriz de transición $P = \{p_{ij}, i, j \in \mathcal{S} \cup \{0\}\}$.

Para un proceso con distribución inicial $\mu$, consideramos la evolución del proceso condicionada a no ser absorbido

$$\varphi_j^{\mu}(n) = \mathbb{P}_{\mu}(X(n) = j|X(n) \neq 0).$$

Recordemos que una QSD es una medida invariante para el proceso condicionado.

Damos algunas definiciones para trabajar en el espacio de trayectorias.

**Definición 5.8.**

- *Una cadena de Markov en $\mathcal{S} \cup \{0\}$ con matriz de transición $P$ y distribución inicial $\nu$ en $\mathcal{S}$ tiene trayectorias irreducibles en $\mathcal{S}$ si para cada $n \in \mathbb{N}$ la medida $\mu$ en $\mathcal{S}^{n+1}$ definida por $\mu(i_0, i_1, \ldots i_n) := \nu(i_0)p_{i_0,i_1} \ldots p_{i_{n-1},i_n}$ es irreducible.*

- *Dados dos números naturales $n < m$, definimos $\chi_n^m$ al conjunto de posibles trayectorias $i_n^m = (i_n, .., i_m)$ de una cadena con matriz de transición $P$ en el intervalo de tiempo $[n, m]$ sin ser absorbida.*

- *Sea $\nu$ una medida en $\mathcal{S}$, definimos la medida $\mu_n^m(\nu, P)$ en $\chi_n^m$ dada por*

$$\mu_n^m(\nu, P)(i_n, .., i_m) = \frac{\nu_{i_n} p_{i_n,i_{n+1}} \cdots p_{i_{m-1},i_m}}{1 - \sum_{i\in\mathcal{S}} \nu_i (P^{m-n})_{i0}}$$

**Observación 5.9.** *La medida $\mu_n^m(\nu, P)$ es la distribución de $(X(n), \ldots, X(m))$ con distribución inicial $\nu$ y transiciones dadas por $P$, condicional a no ser absorbida, con lo cual*

$$\varphi_j^{\nu}(m - n) = \sum_{(j_n,\ldots,j_{m-1})\in\chi_n^{m-1}} \mu_n^m(\nu, P)(j_n, \ldots, j_{m-1}, j), \qquad \forall j \in \mathcal{S}.$$

*De donde la última marginal del vector con distribución $\mu_n^m(\nu, P)$ tiene la distribución de $\varphi^{\nu}(m - n)$.*

### 5.1.2. Límite de Yaglom vía desigualdad de Holley

Introducimos la siguiente notación.

**Definición 5.10.**

- *Dados dos vectores $\mu$ y $\nu$ en $\mathcal{S}$, definimos el vector producto $\mu*\nu$,*

$$(\mu * \nu)_i := \mu_i \nu_i, \forall i \in \mathcal{S}.$$

- *Dada una matriz $P$ definimos $P_{i,\cdot}$ a la fila $i$-ésima de la matriz $P$,*

$$P_{i,\cdot} := (p_{i1}, p_{i2}, \ldots)$$

*análogamente consideramos $P_{\cdot,j}$ a la columna $j$-ésima.*

Damos a continuación otra versión de la desigualdad de Holley, que fue introducida en [FR15], la cual es una versión adaptada para el espacio de trayectorias.

**Proposición 5.11.** *[FR15, Proposición 2] Sean $\nu$ y $\nu'$ probabilidades en $\mathcal{S}$, $P$ y $P'$ matrices de transición en $\mathcal{S} \cup \{0\}$ con 0 un estado absorbente. Notamos $\mu = \mu_n^m(\nu, P)$ y $\mu' = \mu_n^m(\nu', P')$ la distribución de las trayectorias condicionadas.*

*Asumimos que $\mu$ es una medida irreducible en el espacio de las trayectorias $\chi_n^m$. Si para $j, j', l, l' \in \mathcal{S}$ tal que $j \leq j'$ y $l \leq l'$, siempre que los denominadores sean positivos,*

$$\frac{(\nu * P_{\cdot,l})}{\sum_{i\in\mathcal{S}} \nu_i p_{i,l}} \preceq \frac{(\nu' * P'_{\cdot,l'})}{\sum_{i\in\mathcal{S}} \nu'_i p'_{i,l}} \tag{5.1}$$

$$\frac{(P_{j,\cdot} * P_{\cdot,l})}{(P^2)_{j,l}} \preceq \frac{(P'_{j',\cdot} * P'_{\cdot,l'})}{(P'^2)_{j',l'}} \tag{5.2}$$

$$\frac{P_{j,\cdot}}{1 - p_{j0}} \preceq \frac{P'_{j',\cdot}}{1 - p_{j'0}} \tag{5.3}$$

*(como medidas sobre $\mathcal{S}$) entonces*

$$\mu \preceq \mu'.$$

El siguiente teorema es una aplicación de la desigualdad de Holley en el espacio de las trayectorias finitas de la cadena. Notamos $\delta_i$ la medida de probabilidad en $\mathcal{S}$ que concentra toda su masa en el estado $i \in \mathcal{S}$.

**Teorema 5.12.** *[FR15, Teorema 1] Sea $\mathcal{S}$ un espacio numerable con un elemento minimal al que llamamos 1. Sea $P$ la matriz de transición de una cadena de Markov aperiódica en $\mathcal{S} \cup \{0\}$. Asumimos que la cadena con distribución inicial $\delta_1$ tiene trayectorias irreducibles en $\mathcal{S}$. Si para todo $i, i', l, l' \in \mathcal{S}$ con $i \leq i'$, $l \leq l'$, cuando los denominadores son positivos se verifica*

$$\frac{(P_{i,\cdot} * P_{\cdot,l})}{(P^2)_{i,l}} \preceq \frac{(P_{i',\cdot} * P_{\cdot,l'})}{(P^2)_{i',l'}}, \tag{5.4}$$

$$\frac{P_{i,\cdot}}{1 - p_{i0}} \preceq \frac{P_{i',\cdot}}{1 - p_{i'0}}, \tag{5.5}$$

*como medidas en $\mathcal{S}$, entonces*

1. *La sucesión $(\varphi^{\delta_1}(n))_{n\geq 0}$ es monótona: $\varphi^{\delta_1}(n) \preceq \varphi^{\delta_1}(n+1)$ para todo $n \geq 0$.*

2. *Para cualquier probabilidad $\nu$ en $\mathcal{S}$, $\varphi^{\delta_1}(n) \preceq \varphi^{\nu}(n)$.*

3. *En particular, si $\nu$ es una QSD, entonces $\varphi^{\delta_1}(n) \preceq \nu$, para todo $n \geq 0$.*

4. *Si hay una QSD para $P$, entonces el límite de Yaglom de $\delta_1$ converge. La distribución del límite $\underline{\nu} := \lim_{n\to\infty} \varphi^{\delta_1}(n)$ es la QSD minimal.*

## 5.2. Proceso de nacimiento y muerte perezoso

**Definición 5.13.** *Para cada $K \in \mathbb{N}$, consideramos una cadena de Markov*

$$X^K := \{X^K(n), n \in \mathbb{N}\}$$

*definida en el espacio de estados $\{0, 1, .., K\}$, con matriz de transición $P^K$ definida de la siguiente manera:*

$$q, r, p > 0 \qquad q + r + p = 1$$

$$p^K_{i,i-1} = q, \quad p^K_{i,i} = r, \quad p^K_{i,i+1} = p \quad \text{para } i \in \{1, .., K-1\}$$

$$p^K_{0,0} = 1, \quad p^K_{K,0} = p, \quad p^K_{K,K-1} = q, \quad p^K_{K,K} = r.$$

*Dado que estaremos trabajando con varios procesos, utilizaremos la siguiente notación,*

$$\varphi^{K,\nu}_j(n) := \mathbb{P}_\nu(X^K(n) = j | X^K(n) \neq 0).$$

**Observación 5.14.** *Podríamos considerar el proceso en $\{0, \ldots, K+1\}$ con 0 y $K+1$ estados absorbentes. Por simplicidad unimos los estados absorbentes y consideramos los procesos de $\{0, \ldots, K\}$ con 0 el único estado absorbente. Esto no representa ninguna restricción.*

Cuando el superíndice de la matriz de transición sea $K$ nos estaremos refiriendo a la matriz de transición del proceso de nacimiento y muerte definido en $\{0, 1, \ldots, K\}$. Cualquier otro superíndice de una matriz es simplemente elevar la matriz a esa potencia.

Observemos que para cada $K \in \mathbb{N}$ el espacio de estados del proceso $X^K$ es finito y por lo tanto, por el Teorema 3.9 tiene una única QSD que la denotamos $\nu^K = (\nu^K_1, \nu^K_2, \ldots, \nu^K_K)$ y además es el límite de Yaglom

$$\lim_{n\to\infty} \varphi^{K,\delta_i}_j(n) = \nu^K_j \qquad \forall i, j \in \{1, .., K\}. \tag{5.6}$$

**Observación 5.15.** *En este caso hay una expresión explícita para las medidas $\nu^K$. Puede verse en [Yue05] que para cada $K \in \mathbb{N}$*

$$\nu^K_j = \nu^K_1 \left(\sqrt{\frac{p}{q}}\right)^{j-1} \sin\left(\frac{jK\pi}{(K+1)}\right) \qquad j = 2, \ldots, K,$$

*pudiendo elegir $\nu^K_1$ de tal forma que $\nu^K$ sea una probabilidad.*

A continuación mostraremos que a los procesos $X^K$ se les puede aplicar el Teorema 5.12. Para estos procesos vimos que el límite de Yaglom converge a la única QSD, sin restricciones sobre los parámetros $p, r$ y $q$. Sin embargo haremos su prueba ya que esto nos permitirá simplificar la prueba del Teorema 5.19.

**Proposición 5.16.** *Para cada $K \in \mathbb{N}$, sea $X^K$ el proceso de la Definición 5.13. Si $pq \leq r^2$ entonces el proceso $X^K$ verifica las hipótesis del Teorema 5.12.*

*Demostración.* Deben verificarse las condiciones (5.4) y (5.5). Si denotamos

$$b_K((i,l),j) := \frac{\sum_{w\geq j} p_{iw}^K p_{wl}^K}{(P^K)_{il}^2}, \qquad c_K(i,j) := \frac{\sum_{w\geq j} p_{iw}^K}{1-p_{i0}^K}, \tag{5.7}$$

la condición (5.4) se reescribe $b_K((i,l),j) \leq b_K((\hat{i},\hat{l}),j)$ para $i \leq \hat{i}$ y $l \leq \hat{l}$. Observemos que los valores de $b_K((i,l),j)$ relevantes ocurren cuando $|j-i| \leq 1$ o $|j-l| \leq 1$ y también cuando $|i-l| \leq 2$. Luego, la condición (5.4) es equivalente, para $j \geq 2$ a

$$b_K((j-1,j-1),j) \leq b_K((j,j-1),j) \leq b_K((j,j),j),$$

$$b_K((j-1,j-1),j) \leq b_K((j-1,j),j) \leq b_K((j,j),j).$$

En la siguiente tabla mostramos los valores $b_K((i,l),j)$.

| $(i,l)$ \ $j$ | 1 | 2 | $3 \leq j \leq K-1$ | $K$ |
|---|---|---|---|---|
| $(1,1)$ | 1 | $\frac{pq}{pq+r^2}$ | 0 | 0 |
| $(1,2)$ | 1 | 1/2 | 0 | 0 |
| $(j-1,j-1)$ | $-$ | $\frac{pq}{pq+r^2}$ | $\frac{pq}{2pq+r^2}$ | $\frac{pq}{2pq+r^2}$ |
| $(j,j-1)$ | $-$ | 1/2 | 1/2 | 1/2 |
| $(j-1,j)$ | $-$ | 1/2 | 1/2 | 1/2 |
| $(j,j)$ | 1 | $\frac{pq+r^2}{2pq+r^2}$ | $\frac{pq+r^2}{2pq+r^2}$ | $\frac{r^2}{pq+r^2}$ |
| $(K,K-1)$ | 1 | 1 | 1 | 1/2 |
| $(K,K)$ | 1 | 1 | 1 | $\frac{r^2}{pq+r^2}$ |

De la tabla se ve que solo hay restricciones en los siguientes casos:

$$b_K((1,1),2) \leq b_K((1,2),2) \text{ que se cumple si y sólo si } pq \leq r^2,$$

$$b_K((K,K-1),K) \leq b_K((K,K),K) \text{ que se cumple si y sólo si } pq \leq r^2.$$

El resto de las desigualdades simplemente se verifican al ser $p, q$ y $r$ probabilidades. Con lo cual, si $pq \leq r^2$ se verifica (5.4).

Pasemos a analizar los valores de $c_K(i,j)$. La condición (5.5) la podemos reescribir como

$$c_K(i,j) \leq c_K(\hat{i},j) \text{ con } i \leq \hat{i}.$$

Los valores relevantes ocurren para $|i-j| \leq 1$ haciendo que (5.5) sea equivalente a

$$c_K(j-1,j) \leq c_K(j,j), \quad j \geq 2.$$

En la siguiente tabla mostramos los valores $c_K(i,j)$.

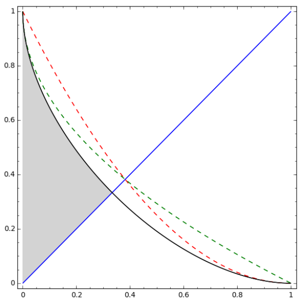
| $i \backslash j$ | $1$ | $2$ | $3 \leq j \leq K-1$ | $K$ |
|---|---|---|---|---|
| $1$ | $1$ | $\frac{p}{1-q}$ | $0$ | $0$ |
| $j-1$ | $-$ | $\frac{p}{1-q}$ | $p$ | $p$ |
| $j$ | $1$ | $p+r$ | $p+r$ | $\frac{r}{1-p}$ |
| $K-1$ | $1$ | $1$ | $1$ | $p$ |
| $K$ | $1$ | $1$ | $1$ | $\frac{r}{1-p}$ |

Las únicas restricciones relevantes son

$$c(1,2) \leq c(2,2) \text{ que se cumple si y sólo si } p \leq (p+r)^2.$$

$$c(K-1,K) \leq c(K,K) \text{ que se cumple si y sólo si } p(1-p) \leq r.$$

Si $pq \leq r^2$ se verifican ambas desigualdades. En el siguiente gráfico (en función de $p$ y $q$), ilustra que si $pq \leq r^2$, se cumplen todas las desigualdades,




Las curvas punteadas corresponden a las ecuaciones $p = (p+r)^2$ y $p(1-p) = r$ con $r = 1-p-q$. La curva sólida corresponde a $pq = r^2$ y la región sombreada es $pq \leq r^2$ y $p \leq q$. Finalmente, para cada $K \in \mathbb{N}$ el proceso $X^K$ verifica las hipótesis del Teorema 5.12. □

A continuación probaremos que las QSD $(\nu^K)_{K\in\mathbb{N}}$ son medidas crecientes.

**Lema 5.17.** *Consideremos los procesos $(X^K)_{K\geq 1}$ de la Definición 5.13. Si $pq \leq r^2$ y $q \geq p$, entonces*

$$\nu^K \preceq \nu^{K+1} \quad \forall \quad K \in \mathbb{N},$$

*donde $\nu^K$ es la única QSD de $X^K$.*

*Demostración.* Veamos que para la medida $\delta_1$ y las matrices de transición $P^K$ y $P^{K+1}$ vistas en $\mathbb{N} \times \mathbb{N}$ se cumplen las hipótesis de la desigualdad de Holley (Proposición 5.11).

Deben verificarse las condiciones (5.1), (5.2) y (5.3). Si denotamos

$$a_K(\nu|i,j) := \frac{\sum_{w\geq j} \nu_w p_{wi}^K}{(\nu P^K)_i}. \tag{5.8}$$

$$b_K((i,l),j) := \frac{\sum_{w\geq j} p_{iw}^K p_{wl}^K}{(P^K)_{il}^2}, \qquad c_K(i,j) := \frac{\sum_{w\geq j} p_{iw}^K}{1-p_{i0}^K},$$

las condiciones (5.1),(5.2) y (5.3) se pueden reescribir como

$$a_K(\nu|i,j) \leq a_{K+1}(\nu|\hat{\imath},j) \text{ con } i \leq \hat{\imath},\ j \in \mathbb{N}.$$

$$b_K((i,l),j) \leq b_{K+1}((\hat{\imath},\hat{l}),j) \text{ con } i \leq \hat{\imath}, l \leq \hat{l},\ j \in \mathbb{N}.$$

$$c_K(i,j) \leq c_{K+1}(\hat{\imath},j) \text{ con } i \leq \hat{\imath}.$$

Analicemos cada una de las desigualdades.

- Siendo que vamos a aplicar la desigualdad de Holley para la medida $\nu = \delta_1$, obtenemos que sin importar si consideramos la matriz $P^K$ o $P^{K+1}$ se obtiene que el denominador de $a_K(\delta_1|i,j)$ es positivo solo para valores de $i \in \{0,1,2\}$,

  $$a_K(\delta_1|i,1) = \frac{p_{1i}}{p_{1i}} = 1, \qquad a_K(\delta_1|i,j) = 0 \text{ si } j \geq 2, \qquad i \in \{0,1,2\}.$$

  Siendo que no depende del $K$ obtenemos que se verifica la desigualdad (5.1) para la medida $\delta_1$ en $\mathcal{S}$.

- Utilizando la Proposición 5.16, la ecuación (5.2) es equivalente a

  $$b_K((i,l),j) \leq b_{K+1}((i,l),j), \qquad \forall i,l,j \in \{1,\ldots,K\}$$

  y la única condición no trivial es

  $$b_K((K,K),K) \leq b_{K+1}((K,K),K) \text{ que se verifica si y sólo si } 0 \leq (pq)^2.$$

- Utilizando el Lema 5.16, la ecuación (5.3) es equivalente a

  $$c_K(i,j) \leq c_{K+1}(i,j), \qquad \forall i,j \in \{1,\ldots,K\}.$$

  La única condición no trivial es

  $$c_K(K,K) \leq c_{K+1}(K,K) \text{ que se verifica si y sólo si } 0 \leq pq.$$

Finalmente, bajo las condiciones $pq \leq r^2$ y $q \geq p$, se cumple la Proposición 5.11 de donde

$$\mu^0_{-n}(\delta_1, P^K) \preceq \mu^0_{-n}(\delta_1, P^{K+1}),$$

y de la Observación 5.9

$$\varphi^{K,\delta_1}(n) \preceq \varphi^{K+1,\delta_1}(n), \qquad \forall n \in \mathbb{N}.$$

Tomando límite $n \to \infty$ obtenemos

$$\nu^K \preceq \nu^{K+1}.$$

□

En lo que sigue pasaremos a trabajar con procesos de nacimiento y muerte definidos en el espacio de estados $\mathbb{N}$.

El siguiente ejemplo puede verse en [FR15, Sección 3.1], es una aplicación del Teorema 5.12. Sea $X = \{X(n), n \in \mathbb{N}\}$ un proceso de nacimiento y muerte que toma valores en $\mathbb{N}_0$. La matriz de transición $P$ es tal que

$$q, r, p > 0, \quad q+r+p = 1, \quad p_{0,0} = 1, \quad p_{i,i-1} = q, \quad p_{i,i} = r, \quad p_{i,i+1} = p, \qquad i \in \mathbb{N}. \tag{5.9}$$

Una medida $\nu$ en $\mathbb{N}$ es una QSD si y sólo si se verifica la ecuación (3.2), que en este caso es

$$\nu_{i+1}q + \nu_{i-1}p + \nu_i(q\nu_1 - (p+q)) = 0, \quad i \in \mathbb{N}, \tag{5.10}$$

con la convención que $\nu_0 = 0$. En [Cav78] se prueba que si $p < q$ hay una familia de QSD indexadas por $\nu_1$ con $\nu_1 \in (0, (1-\sqrt{\lambda})^2)$, donde $\lambda = p/q$.

Dado que la probabilidad de absorción es una QSD que verifica $(\nu P)_0 = q\nu_1$, la QSD con máxima $\nu_1$ es la QSD minimal $\nu_{\text{mín}}$, una binomial negativa con parámetros 2 y $\lambda$:

$$(\nu_{\text{mín}})_i = (1-\sqrt{\lambda})^2 i(\sqrt{\lambda})^{i-1}, \quad i \in \mathbb{N}. \tag{5.11}$$

**Corolario 5.18.** *[FR15, Corolario 7] Supongamos que $X$ es un proceso de nacimiento y muerte, con matriz de transición $P$ que verifica 5.9. Si $q > p$ y $pq \leq r^2$, el proceso $X$ cumple el Teorema 5.12 con $\underline{\nu} = \nu_{\text{mín}}$ dada por 5.11. Además el límite de Yaglom es $\nu_{min}$.*

*Demostración.* Su demostración es similar a la del Lema 5.16. □

El siguiente teorema, resultado principal de este capítulo, prueba que las QSD de los procesos de nacimiento y muerte definidos en el espacio $\{0, \ldots, K\}$ convergen cuando $K \to \infty$ a la QSD del proceso de nacimiento y muerte en los $\mathbb{N}$.

**Teorema 5.19.** *Sean $\nu^K$ las QSD de los procesos definidos en la Definición 5.13. Si $q > p$ y $pq \leq r^2$ existe $\nu$ tal que*

$$\nu^K \xrightarrow{K\to\infty} \nu.$$

*Además $\nu = \nu_{\text{mín}}$, con $\nu_{\text{mín}}$ dada por (5.11).*

*Demostración.* Vamos a probar que dado $K \in \mathbb{N}$, los procesos $X^K$ y $X$ (ambos vistos en $\mathbb{N}$) cumplen la desigualdad de Holley para la medida $\delta_1$. Es decir, la medida $\delta_1$ y las matrices de transición $P^K$ y $P$ (vistas en $\mathbb{N} \times \mathbb{N}$) cumplen las hipótesis de la Proposición 5.11. La matriz $P$ esta definida en 5.9 y $P^K$ en la Definición 5.13 considerando $p^K_{i,j} := 0$ si $i, j > K$.

Utilizaremos las notaciones de la ecuación (5.7) y (5.8) para la matriz $P^K$ y denotaremos para la matriz $P$ respectivamente $b((i,l),j), c(i,j)$ y $a(\nu|i,j)$.

Utilizando esta notación, el Lema 5.16 y el Corolario 5.18, las condiciones (5.1),(5.2) y (5.3) se pueden reescribir como

$$a_K(\delta_1|i,j) \leq a(\delta_1|i,j) \text{ con } i, j \in \mathbb{N}.$$

$$b_K((i,l),j) \leq b((i,l),j) \text{ con } i,l,j \in \mathbb{N}.$$

$$c_K(i,j) \leq c(i,j) \text{ con } i,j \in \mathbb{N}.$$

Analicemos cada una de las desigualdades.

- Notemos que solo hay que verificar la desigualdad (5.1) para $i \in \{0,1,2\}$

$$a_K(\delta_1|i,1) = \frac{p_{1i}}{p_{1i}} = a(\delta_1|i,1), \quad a_K(\delta_1|i,j) = 0 = a(\delta_1|i,j) \text{ si } j \geq 2.$$

- Notemos que $b_K((i,l),j)$ para $i > K$ ó $j > K$ no está definido pues el denominador es 0. Con lo cual resta verificar la condición

$$b_K((K,K),j) \leq b((K,K),j), \quad j \in \mathbb{N}.$$

  El caso $j \leq K-1$ se obtiene $1 = 1$. Para $j = K$,

$$\frac{r^2}{r^2+pq} = b_K((K,K),K) \leq b((K,K),K) = \frac{pq+r^2}{2pq+r^2} \iff 0 \leq (pq)^2.$$

- La condición que resta chequear es

$$c_K(K,j) \leq c(K,j), \quad j \in \mathbb{N}.$$

  Si $j > K$ o $j < K-1$ las desigualdades son triviales, luego restan chequear

$$\frac{r+q}{1-p} = c_K(K,K-1) \leq c(K,K-1) = 1.$$

$$\frac{r}{1-p} = c_K(K,K) \leq c(K,K) = r+p \iff 0 \leq pq.$$

Bajo las condiciones $pq \leq r^2$ y $q > p$, se cumple la Proposición 5.11 de donde

$$\mu^0_{-n}(\delta_1, P^K) \preceq \mu^0_{-n}(\delta_1, P),$$

y por la Observación 5.9

$$\varphi^{K,\delta_1}(n) \preceq \varphi^{\delta_1}(n).$$

Por el Corolario 5.18 $\varphi^{\delta_1}(n) \preceq \varphi^{\nu_{min}}(n)$, por lo tanto

$$\varphi^{K,\delta_1}(n) \preceq \varphi^{\nu_{min}}(n) = \nu_{min}. \tag{5.12}$$

Tomando límite $n \to \infty$ obtenemos

$$\nu^K \preceq \nu_{min} \quad \text{para todo } K \in \mathbb{N}.$$

Dado que $\nu^K$ es una sucesión de medidas crecientes y acotadas, existe el límite y lo llamamos $\nu$. Siendo que $\nu^K$ es la QSD asociada a la matriz $P^K$, por definición

$$\nu_j^K = \frac{\sum_i \nu_i^K p_{ij}^K}{1 - \sum_i \nu_i^K p_{i0}^K} = \begin{cases} \frac{\nu_{j-1}^K p + \nu_j^K r + \nu_{j+1}^K q}{1 - \nu_1^K q - \nu_K^K p} & j \in \{1, \ldots, K-1\} \\ \\ \frac{\nu_{K-1}^K p + \nu_K^K r}{1 - \nu_1^K q - \nu_K^K p} & j = K \end{cases}$$

con $\nu_0 = 0$. Cuando $K \to \infty$ se obtiene

$$\nu_j = \frac{\nu_{j-1}p + \nu_j r + \nu_{j+1}q}{1 - \nu_1 q} = \frac{(\nu P)_j}{1 - \nu_1 q} \qquad \forall j \in \mathbb{N}.$$

Por lo tanto $\nu$ es una QSD del proceso $X$ y como $\nu \preceq \nu_{\text{mín}}$ es la minimal. $\square$

**Corolario 5.20.** *Se obtiene que $\nu^K \overset{K\to\infty}{\longrightarrow} \nu_{\text{mín}}$ sin la restricción $pq \leq r^2$.*

*Demostración.* Sean $p, q, r$ los parámetros asociados a los procesos de nacimiento y muerte $X^K$ y $X$ tales que $p + r + q = 1$ y $q > p > 0$. Para cada $K$, el proceso $X^K$ tiene una única QSD $\nu^K$ que verifica la ecuación

$$0 = \nu_{j+1}^K q(1 - \mathbf{1}\{K = j\}) + \nu_{j-1}^K p + \nu_j^K (q\nu_1^K + p\nu_K^K - (p+q)). \tag{5.13}$$

Por otro lado, el proceso $X$ tiene infinitas QSD $\nu$ que verifican

$$\nu_{j+1}q + \nu_{j-1}p + \nu_j(q\nu_1 - (p+q)) = 0, \quad j \in \mathbb{N}, \tag{5.14}$$

denotamos $\nu_{\text{mín}}$ a la QSD minimal.

Consideramos otros procesos de nacimiento y muerte $\hat{X}^K$ y $\hat{X}$ con parámetros $\hat{p} = \alpha p$, $\hat{q} = \alpha q$ , $\hat{r} = 1 - \hat{p} - \hat{q}$ tales que $\alpha > 0$, $\hat{p}\hat{q} \leq \hat{r}^2$. La existencia de $\alpha$ se deduce de que $\lim_{\alpha\to 0} \hat{p}\hat{q} = 0$ y $\lim_{\alpha\to 0} \hat{r}^2 = 1$. Estos procesos tienen QSD $\hat{\nu}^K$ y $\hat{\nu}$ que verifican las mismas ecuaciones que $\nu^K$ y $\nu$ (5.13 y 5.14) respectivamente. Luego $\hat{\nu}^K = \nu^K$ y $\hat{\nu}_{\text{mín}} = \nu_{\text{mín}}$ y por el Teorema 5.19 se deduce que $\nu^K \to \nu_{\text{mín}}$. $\square$

Si bien la convergencia de las distribuciones cuasiestacionarias para el caso periódico de los procesos de nacimiento y muerte está incluida en el corolario anterior, no ocurre lo mismo con los procesos condicionados. Este comportamiento tampoco se deduce del cómputo explícito. En la siguiente sección damos dicho resultado utilizando técnicas de dominación estocástica para procesos periódicos.

### 5.2.1. Caso periódico

**Definición 5.21.** *Sea $P$ una matriz. El período de un estado $j$ se define como*

$$d(j) := mcd\{n \geq 1/(P^n)_{jj} > 0\}.$$

*Decimos que la matriz $P$ (o el proceso con matriz de transición $P$) tiene periodo $d$ si todos los estados tienen periodo $d$.*

Asumamos que la matriz de transición $P$ es irreducible en $\mathcal{S}$ y que tiene periodo $d \geq 2$. Sean $\mathcal{S}_1, \ldots, \mathcal{S}_d \subset \mathcal{S}$, las subclases cíclicas. Denotamos

$$i \sim j \qquad \Longleftrightarrow \qquad (P^{dn})_{il} > 0 \text{ para algún } n \in \mathbb{N}.$$

Suponemos que las clases son de tal forma que si $i \in \mathcal{S}_i$ y $P_{ij} > 0$ entonces $j \in \mathcal{S}_{i+1}$, denotando $\mathcal{S}_{d+1} = \mathcal{S}_1$.

Damos a continuación una versión del Teorema 5.12 adaptado para procesos periódicos.

**Teorema 5.22.** *[FR15, Teorema 9] Sea $\mathcal{S}$ un espacio numerable con un elemento minimal al que llamamos 1. Sea $P$ la matriz de transición de una cadena de Markov en $\mathcal{S} \cup \{0\}$ con período $d$. Sean $\mathcal{S}_1, \ldots, \mathcal{S}_d \subset \mathcal{S}$, elegidos tal que $1 \in \mathcal{S}_1$. Asumimos que la cadena con distribución inicial $\delta_1$ tiene trayectorias irreducibles.*

*Si para todo $i, i', l, l' \in \mathcal{S}$ con $i \leq i'$, $l \leq l'$, $i, i'$ en la misma clase y las desigualdades estocásticas (5.4) y (5.5) se satisfacen cuando los denominadores son positivos, entonces*

1. *Para todo $n \geq 0$ $\varphi^{\delta_1}(n) \preceq \varphi^{\delta_1}(n+d)$.*

2. *Para cualquier probabilidad $\nu$ en $\mathcal{S}_1$, $\varphi^{\delta_1}(n) \preceq \varphi^{\nu}(n)$.*

3. *En particular, si $\nu$ es una QSD, entonces $\varphi^{\delta_1}(dn+l-1) \preceq \nu_{\cdot|\mathcal{S}_l}$, para todo $n \geq 0$.*

4. *Si hay una QSD para $P$, entonces el límite de Yaglom de $\delta_1$ a través de subsucesiones de período $d$, esta dado por*

$$\lim_{n\to\infty} \varphi^{\delta_1}(dn+l-1) = \hat{\nu}_{\cdot|\mathcal{S}_l}.$$

*Además, para cualquier otra QSD $\nu$, se tiene $\hat{\nu}_{\cdot|\mathcal{S}_l} \preceq \nu_{\cdot|\mathcal{S}_l}$ para todo $l$.*

Vamos a considerar los procesos de nacimiento y muerte periódicos, definidos en espacio finito y en espacio numerable. Al igual que en el caso aperiódico probaremos que las QSD de los procesos definidos en espacios finitos convergen a la QSD minimal del proceso en $\mathbb{N}$.

**Definición 5.23.** *Para cada $K \in \mathbb{N}$, consideramos una cadena de Markov $X^K = \{X^K(n), n \in \mathbb{N}\}$ definida en el espacio de estados $\{0, 1, .., K\}$, con matriz de transición $P^K$ definida de la siguiente manera:*

$$p, q > 0 \qquad p + q = 1$$
$$p^K_{i,i-1} = q, \quad p^K_{i,i+1} = p \qquad \textit{para} \quad i \in \{1, .., K-1\}$$
$$p^K_{0,0} = 1, \quad p^K_{K,0} = p, \quad p^K_{K,K-1} = q.$$

Notemos que para cada $K \in \mathbb{N}$ el espacio de estados del proceso $X^K$ es finito y por lo tanto, por el Teorema 3.9 tiene una única QSD que la denotamos $\nu^K = (\nu^K_1, \nu^K_2, \ldots, \nu^K_K)$ y además es el límite de Yaglom

$$\lim_{n\to\infty} \varphi^{K,\delta_i}_j(n) = \nu^K_j \qquad \forall i, j \in \{1, .., K\}. \tag{5.15}$$

**Definición 5.24.** *Consideramos el proceso de nacimiento y muerte $X$ en $\mathbb{N}$. Sea $P$ la matriz de transición definida de la siguiente manera:*

$$q > p > 0 \qquad p + q = 1.$$

$$p^K_{0,0} = 1, \quad p_{i,i-1} = q, \quad p_{i,i+1} = p \quad \text{para } \; i \in \mathbb{N},$$

Notemos que las QSD del proceso $X$ verifican (5.10) y por lo tanto la QSD minimal está dada por (5.11).

**Teorema 5.25.** *Para cada $K \in \mathbb{N}$ consideremos el proceso $X^K$ de la Definición 5.23, y $\nu^K$ la única QSD de $X^K$ . Si $q > p$, existe $\nu$ tal que*

$$\nu^K \overset{K\to\infty}{\longrightarrow} \nu.$$

*Además $\nu = \nu_{\text{mín}}$, con $\nu_{\text{mín}}$ dada por (5.11).*

*Demostración.* El resultado se obtiene de las siguientes afirmaciones:

1. El proceso $X^K$ cumple las hipótesis del Teorema 5.22, $K \in \mathbb{N}$.

2. Las distribuciones $\nu^K$ verifican

$$\nu^K \preceq \nu^{K+1} \quad \forall \;\; K \in \mathbb{N}.$$

3. El proceso $X$ de la Definición 5.24 cumple

$$\lim_{n\to\infty} \varphi^{\delta_1}(2n) = (\nu_{\text{mín}})_{\cdot|\mathcal{S}_1}, \quad \lim_{n\to\infty} \varphi^{\delta_1}(2n+1) = (\nu_{\text{mín}})_{\cdot|\mathcal{S}_2}.$$

Estos resultados son similares a los Lemas 5.16 y 5.17 y al Corolario 5.18 pero en este caso hay que chequear las condiciones (5.1), (5.2),(5.3), (5.4) y (5.5) para $i, i'$ ambos pares o ambos impares. La demostración es análoga a la del caso aperiódico, pero en este caso con los siguientes los valores de $b_K((i,l), j)$ y $c_K(i,j)$.

| $(i,l)$ \ $j$ | 1 | 2 | $3 \le j < K-1$ | $K-1$ |
|---|---|---|---|---|
| $(1,1)$ | 1 | 1 | 0 | 0 |
| $(j,j)$ | 1 | 1/2 | 1/2 | 1/2 |
| $(j+1,j-1)$ | $-$ | 1 | 1 | 1 |
| $(j+1,j+1)$ | 1 | 1 | 1 | 1 |
| $(K,K-2)$ | 1 | 1 | 1 | 1 |
| $(K,K)$ | 1 | 1 | 1 | 1 |

| $i$ \ $j$ | 1 | 2 | $3 \le j \le K-2$ | $K-1$ | $K$ |
|---|---|---|---|---|---|
| 1 | 1 | 1 | 0 | 0 | 0 |
| $j-2$ | $-$ | $-$ | 0 | 0 | 0 |
| $j-1$ | $-$ | $p$ | $p$ | $p$ | $p$ |
| $j$ | 1 | $p$ | $p$ | $p$ | 0 |
| $K-1$ | 1 | 1 | 1 | $p$ | $p$ |
| $K$ | 1 | 1 | 1 | 1 | 0 |

□

# Bibliografía


[AEGR15] Enrique Andjel, François Ezanno, Pablo Groisman, and Leonardo T. Rolla. Subcritical contact process seen from the edge: convergence to quasi-equilibrium. *Electron. J. Probab.*, 20:no. 32, 16, 2015.

[AFGJ16] Amine Asselah, Pablo A. Ferrari, Pablo Groisman, and Matthieu Jonckheere. Fleming-Viot selects the minimal quasi-stationary distribution: the Galton-Watson case. *Ann. Inst. Henri Poincaré Probab. Stat.*, 52(2):647–668, 2016.

[AFP88] David Aldous, Barry Flannery, and José Luis Palacios. Two applications of urn processes the fringe analysis of search trees and the simulation of quasi-stationary distributions of markov chains. *Probability in the Engineering and Informational Sciences*, 2:293–307, 7 1988.

[AK68] Krishna B. Athreya and Samuel Karlin. Embedding of urn schemes into continuous time Markov branching processes and related limit theorems. *Ann. Math. Statist.*, 39:1801–1817, 1968.

[AN04] K. B. Athreya and P. E. Ney. *Branching processes.* Dover Publications, Inc., Mineola, NY, 2004. Reprint of the 1972 original [Springer, New York; MR0373040].

[Bar49] MS Bartlett. Some evolutionary stochastic processes. *Journal of the Royal Statistical Society. Series B (Methodological)*, 11(2):211–229, 1949.

[BE84] H. Brézis and J.R. Esteban. *Análisis funcional: teoría y aplicaciones.* Alianza universidad textos. Alianza, 1984.

[BH00] L. A. Breyer and A. G. Hart. Approximations of quasi-stationary distributions for Markov chains. *Math. Comput. Modelling*, 31(10-12):69–79, 2000. Stochastic models in engineering, technology, and management (Gold Coast, 1996).

[BHM00] Krzysztof Burdzy, Robert Hołyst, and Peter March. A Fleming-Viot particle representation of the Dirichlet Laplacian. *Comm. Math. Phys.*, 214(3):679–703, 2000.

[Bie74] Bienaymé. Sur une question de probabilités. *Bull. Soc. Math. France*, 2:153–154, 1873/74.

[Bré99] Pierre Brémaud. *Markov chains*, volume 31 of *Texts in Applied Mathematics*. Springer-Verlag, New York, 1999. Gibbs fields, Monte Carlo simulation, and queues.

[BT14] Antar Bandyopadhyay and Debleena Thacker. Rate of convergence and large deviation for the infinite color Pólya urn schemes. *Statist. Probab. Lett.*, 92:232–240, 2014.

[BZ14] Daniela Bertacchi and Fabio Zucca. Strong local survival of branching random walks is not monotone. *Adv. in Appl. Probab.*, 46(2):400–421, 2014.

[BZ17] Daniela Bertacchi and Fabio Zucca. A generating function approach to branching random walks. *Braz. J. Probab. Stat.*, 31(2):229–253, 2017.

[Cav78] James A. Cavender. Quasi-stationary distributions of birth-and-death processes. *Adv. Appl. Probab.*, 10(3):570–586, 1978.

[CMSM13] Pierre Collet, Servet Martínez, and Jaime San Martín. *Quasi-stationary distributions*. Probability and its Applications (New York). Springer, Heidelberg, 2013. Markov chains, diffusions and dynamical systems.

[CV16] Nicolas Champagnat and Denis Villemonais. Exponential convergence to quasi-stationary distribution and $Q$-process. *Probab. Theory Related Fields*, 164(1-2):243–283, 2016.

[DS65] J. N. Darroch and E. Seneta. On quasi-stationary distributions in absorbing discrete-time finite Markov chains. *J. Appl. Probability*, 2:88–100, 1965.

[Dur96] Richard Durrett. *Probability: theory and examples*. Duxbury Press, Belmont, CA, second edition, 1996.

[FKM96] P. A. Ferrari, H. Kesten, and S. Martínez. $R$-positivity, quasi-stationary distributions and ratio limit theorems for a class of probabilistic automata. *Ann. Appl. Probab.*, 6(2):577–616, 1996.

[FKMP95] P. A. Ferrari, H. Kesten, S. Martinez, and P. Picco. Existence of quasi-stationary distributions. A renewal dynamical approach. *Ann. Probab.*, 23(2):501–521, 1995.

[FM07] Pablo A. Ferrari and Nevena Marić. Quasi stationary distributions and Fleming-Viot processes in countable spaces. *Electron. J. Probab.*, 12:no. 24, 684–702, 2007.

[FMP91] P. A. Ferrari, S. Martinez, and P. Picco. Some properties of quasi-stationary distributions in the birth and death chains: a dynamical approach. In *Instabilities and nonequilibrium structures, III (Valparaíso, 1989)*, volume 64 of *Math. Appl.*, pages 177–187. Kluwer Acad. Publ., Dordrecht, 1991.

[FR15] Pablo A. Ferrari and Leonardo T. Rolla. Yaglom limit via Holley inequality. *Braz. J. Probab. Stat.*, 29(2):413–426, 2015.

[Fri49] Bernard Friedman. A simple urn model. *Comm. Pure Appl. Math.*, 2:59–70, 1949.

[FV79] Wendell H. Fleming and Michel Viot. Some measure-valued Markov processes in population genetics theory. *Indiana Univ. Math. J.*, 28(5):817–843, 1979.

[GB03] Hans-Otto Georgii and Ellen Baake. Supercritical multitype branching processes: the ancestral types of typical individuals. *Adv. in Appl. Probab.*, 35(4):1090–1110, 2003.

[GHM01] Hans-Otto Georgii, Olle Häggström, and Christian Maes. The random geometry of equilibrium phases. In *Phase transitions and critical phenomena, Vol. 18*, volume 18 of *Phase Transit. Crit. Phenom.*, pages 1–142. Academic Press, San Diego, CA, 2001.

[GJ13] P. Groisman and M. Jonckheere. Simulation of quasi-stationary distributions on countable spaces. *Markov Process. Related Fields*, 19(3):521–542, 2013.

[Har63] Theodore E. Harris. *The theory of branching processes.* Die Grundlehren der Mathematischen Wissenschaften, Bd. 119. Springer-Verlag, Berlin; Prentice-Hall, Inc., Englewood Cliffs, N.J., 1963.

[Hol74] Richard Holley. Remarks on the FKG inequalities. *Comm. Math. Phys.*, 36:227–231, 1974.

[JR95] S. D. Jacka and G. O. Roberts. Weak convergence of conditioned processes on a countable state space. *J. Appl. Probab.*, 32(4):902–916, 1995.

[Kal02] Olav Kallenberg. *Foundations of modern probability.* Probability and its Applications (New York). Springer-Verlag, New York, second edition, 2002.

[KD47] A. N. Kolmogoroff and N. A. Dmitriev. Branching stochastic processes. *C. R. (Doklady) Acad. Sci. URSS (N.S.)*, 56:5–8, 1947.

[Kin63] J. F. C. Kingman. Ergodic properties of continuous-time Markov processes and their discrete skeletons. *Proc. London Math. Soc. (3)*, 13:593–604, 1963.

[KLPP97] Thomas Kurtz, Russell Lyons, Robin Pemantle, and Yuval Peres. A conceptual proof of the Kesten-Stigum theorem for multi-type branching processes. In *Classical and modern branching processes (Minneapolis, MN, 1994)*, volume 84 of *IMA Vol. Math. Appl.*, pages 181–185. Springer, New York, 1997.

[Kol38] A. N. Kolmogorov. Zur lösung einer biologischen aufgabe. *Comm. Math. Mech. Chebychev Univ.*, 2:1–6, 1938.

[Lax02] Peter D. Lax. *Functional analysis.* Pure and Applied Mathematics (New York). Wiley-Interscience [John Wiley & Sons], New York, 2002.

[Mar06] Andrei Andreevich Markov. Rasprostranenie zakona bolshih chisel na velichiny, zavisyaschie drug ot druga. *Izvestiya Fiziko-matematicheskogo obschestva pri Kazanskom universitete*, 15(135-156):18, 1906.

[Moy67] Shu-teh C. Moy. Extensions of a limit theorem of Everett, Ulam and Harris on multitype branching processes to a branching process with countably many types. *Ann. Math. Statist.*, 38:992–999, 1967.

[Mun75] James R. Munkres. *Topology: a first course.* Prentice-Hall, Inc., Englewood Cliffs, N.J., 1975.

[MV12] Sylvie Méléard and Denis Villemonais. Quasi-stationary distributions and population processes. *Probab. Surv.*, 9:340–410, 2012.

[Sen06] E. Seneta. *Non-negative matrices and Markov chains.* Springer Series in Statistics. Springer, New York, 2006. Revised reprint of the second (1981) edition [Springer-Verlag, New York; MR0719544].

[Sev51] B. A. Sevastyanov. The theory of branching random processes. *Uspehi Matem. Nauk (N.S.)*, 6(6(46)):47–99, 1951.

[ST85] E. Seneta and R. L. Tweedie. Moments for stationary and quasistationary distributions of Markov chains. *J. Appl. Probab.*, 22(1):148–155, 1985.

[Ste33] J.F. Steffensen. Deux problèmes du Calcul des Probabilités. *Ann. Inst. H. Poincaré*, 3(3):319–344, 1933.

[Str65] V. Strassen. The existence of probability measures with given marginals. *Ann. Math. Statist.*, 36:423–439, 1965.

[SVJ66] E. Seneta and D. Vere-Jones. On quasi-stationary distributions in discrete-time Markov chains with a denumerable infinity of states. *J. Appl. Probability*, 3:403–434, 1966.

[VDH16] Remco Van Der Hofstad. *Random graphs and complex networks.* Cambridge University Press, 2016.

[VJ67] D. Vere-Jones. Ergodic properties of nonnegative matrices. I. *Pacific J. Math.*, 22:361–386, 1967.

[WG75] H W Watson and Francis Galton. On the probability of the extinction of families. *Journal of the Anthropological Institute of Great Britain*, 4:138–144, 1875.

[Yag47] A. M. Yaglom. The ergodic principle for Markov processes with stationary distributions. *Doklady Akad. Nauk SSSR (N.S.)*, 56:347–349, 1947.

[Yue05] Wen-Chyuan Yueh. Eigenvalues of several tridiagonal matrices. *Appl. Math. E-Notes*, 5:66–74, 2005.