



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

Tesis de Licenciatura

Sistemas de partículas con ramificación y selección

Santiago Durán

Director: Pablo Groisman

Febrero de 2015

Índice general

1. Cadenas de Markov	13
1.1. Un proceso aleatorio en el tiempo	13
1.2. Cadenas de Markov	14
1.3. Irreducibilidad y Aperiodicidad	18
1.4. Medidas Invariantes	20
1.5. Teorema Ergódico	25
1.6. La versión en tiempo continuo	26
1.7. Procesos de Poisson	29
1.8. Tiempos de parada	29
2. La ecuación F-KPP	31
2.1. El problema planteado por Fisher, Kolmogorov, Petrovskii y Piskunov	31
2.2. Resolución del problema y ambigüedad en el resultado	34
2.3. Frente de propagación: falla en lo microscópico. El sistema de Brunet-Derrida	36
3. El sistema de Brunet-Derrida en los enteros	39
3.1. Construcción Gráfica	39
3.2. El caso de dos partículas	41
3.3. Irreducibilidad	43
3.4. Existencia y unicidad de medida invariante	44
3.5. Velocidad Asintótica	49
4. Sobre el contexto histórico	67
4.1. Estadística entre 1850 y 1950. Eugenesia y Darwinismo Social.	67
4.2. Fisher y los genes ventajosos	69
4.3. La ciencia como producción social y cultural	70
5. Apéndice	73
5.1. Exponenciales	73

Agradecimientos

Lejos de casa y con la tesis casi terminada, me parece un buen momento para mirar hacia atrás y pensar qué pasó hasta el día de hoy, cómo es que estoy recibíendome. Fueron unos 5 años transcurridos en este lugar, y sería una simplificación decir que fueron 5 años sólo cursando o estudiando en Ciudad Universitaria. La dinámica que tiene nuestra facultad -llamémosla *dinámica burbuja*- hace que toda la vida que nos construimos pase por acá. Comemos, trabajamos, militamos, dormimos..., y es por eso que los momentos vividos y la gente conocida son una pieza clave de nuestra historia. En mi caso personal, esa pieza clave ya es una parte hermosa de mí, y se vuelve un recuerdo imborrable (y mucho más que un recuerdo) que llevo conmigo de ahora en más.

En ese contexto, la facultad determinó mucho en mis últimos años, pero también lo hicieron los que estuvieron por fuera de ella en el día a día. Empezando por casa, quiero remarcar lo importante que fue para mí el acompañamiento que tuve de parte de mi familia desde el primer momento. No me refiero al acompañamiento económico, que por suerte estuvo y que, claro está, fue clave para poder estudiar en una facultad exigente como la nuestra. Me parece más importante lo más simple: mis viejos siempre estuvieron. Sin ponerme presiones y confiando en mi manera de estudiar, no hubo momento en que dejen de preocuparse por mi carrera y por las dudas que me surgían. Mi viejo pasó por momentos muy similares, y su ayuda y sus ganas de que yo pueda pasarlos igual o mejor fueron muy valiosos. Mamá por su parte le agregó oído y paciencia para aguantarme y orientarme en momentos difíciles. A los lindos momentos compartidos se sumaron también discusiones y diferencias, que al fin y al cabo son normales entre padres e hijos y creo que son sanas en este momento en que me toca vivirlas. Con el tiempo aprendimos a entendernos y escucharnos. Mis hermanos, Manu y la tota, también me tuvieron que aguantar, obvio que desde otro lado, y también les estoy muy agradecido porque estén día a día.

Durante estos años también estuvieron mis primos, mis tíos y mis abuelos. A uno de ellos, mi abuelo Jack, con gusto le recordé regularmente qué era lo que estudiaba, y si bien hoy ya no está con nosotros, hasta hace sólo unos meses sí, y se hubiese puesto contento de verme acá hoy. Por otro lado, y como no puedo negarlo ni ocultarlo, gran parte de mi familia hizo que la facultad sea un ámbito familiar para mí, en un sentido totalmente literal.

En un plano ya no familiar, hay un grupo al que tengo mucho que agradecer. Tengo la suerte de tener una banda amiga que me banca todos los días, con todo lo que eso significa. Con ellos compartimos muchísimas cosas: viajes, salidas, juntadas. En uno u otro caso fueron los primeros en responder cuando las papas quemaban, y la confianza que tengo con cada uno de ellos ya va más allá de cuánto tiempo efectivamente nos podemos encontrar hoy en día. Me refiero a cada uno de los pibes: Coco, Choco, Tomi, Mario, Frenk, Heypn, Fede, Besio y Maxi.

En estos años otra persona pasó a ser parte de mi vida; una persona que en sólo un tiempo me enseñó a compartir, a querer, a amar como nunca antes lo había hecho; que se volvió importantísima para mí y que con mucho amor y cariño me llegó a lo más profundo. Si bien Ale no está acá hoy, sí estuvo al lado mío durante todo el último tiempo, y su presencia fue muy significativa para que yo esté donde estoy. Ella estuvo ayudando, escuchando, acompañando, y hasta vio muy de cerca la evolución de esta tesis, y por eso le destino buena parte de los agradecimientos.

Pensando ya en los momentos que viví dentro de estas mismas paredes, y pensando en la gente que conocí, me surge una sensación de que la lista es larga y de que probablemente me olvide de alguien. Prefiero tomar ese riesgo y repasar la lista, como un buen ejercicio de memoria y retrospectiva que me va a traer lindos recuerdos.

Algunos están desde el primer día y no sabría muy bien por dónde empezar. Los primeros fueron los de la banda del Glorioso: Frank, Sebas, Maxi y Marto. Fiore también está desde el principio, y con todos ellos nos seguimos encontrando hoy en día, y nos hacemos un rato cuando es necesario. Al poco tiempo empecé a conocer a gente de la carrera, que son los que vi más seguido a medida que pasaron los años: Santi, con quien compartí casi literalmente toda la carrera, Carlo, un poco después Manu, Luz, Mer, Fede, Tincho, Rafa, Lupi. Con ellos armamos un grupo que también siento que es parte de mí, y que está en las buenas y en las malas. En estos años salimos, viajamos, hicimos asados tras asado; compartimos muchos momentos, y espero que el hecho de recibirnos no impida que sean muchos más.

La vida en la carrera contó también con otros viajes como las UMAs o el IMPA. Entre eso, las materias cursadas, y la gente que uno conoce de los pasillos, del bar, etc, son muchos los que conocí y que son parte de mi vida cotidiana: Mati, Mateo, Chebi, Nati, Sofi, Meli, Aye, Bruno, Jaz, Xava, Mati Z, Vir, Yami, Pau, Gise. Con la mayoría también nos encontramos organizando cosas interesantes para pensar y discutir la carrera, en ese espacio que llamamos Conjunto Abierto. También les estoy muy agradecido, porque estoy convencido de que la vida en la facultad sin ellos no sería tan interesante.

Mi lugar como estudiante me llevó a interesarme por los problemas de la facultad a un nivel más general, y eso me llevó a ser parte del FEM! desde que entré en el 2010. Como dice la frase, estudiante que lucha también aprende, y claramente la militancia fue una parte clave de mi aprendizaje en la facultad y de mi formación como científico. Si

la facultad no fomenta el debate, al menos está bueno poder juntarse con aquellos que piensan parecido, y que eso se materialice en la vida diaria.

El FEM! es un gran espacio porque es un hermoso grupo humano, donde pude conocer, aprender, discutir y valorar el trabajo colectivo, entre muchas otras cosas. Yo sé que me marcó para toda la vida; fui parte de él toda la carrera, y me voy a seguir sintiendo parte aunque ya no sea estudiante de la facultad.

Un espacio así lo hacen las personas que lo integran. Por eso, mis agradecimientos muy grandes son para todos ellos. Es difícil nombrar a todos porque, alegremente, los que lo formaron y forman parte son como más de treinta. Me limito a los que más tiempo compartieron conmigo, que fueron Lau, Nico, Manu, Euge, Lipa, Marce, Maxi, Cani, Flor, Fiore, Sebas, Axel, Eze, Agus, Tavo, Eric, Lupi y Ale.

También hay muchos que conocí mediante la militancia, pero que esta se volvió una excusa para vernos seguido, al punto de que nuestra relación es ya una amistad que no recordamos cuándo empezó. En esta categoría entran Lau, Pau, Hernán, el Mago, Mati C, Caro, Lupi y Safir. Mención aparte para Nico, con quien nuestro paso como docentes juntos me hizo aprender mucho, y que con muchas ganas e ideas logró que frecuentemente me confunda amistad y trabajo.

Por último, quería escribir unas líneas por Patu, cuya ayuda fue muy importante para mí en el último tiempo. Cuando fue mi profesor de proba me gustaron mucho sus clases, y después me convenció aún más como director. Fueron dos años de idas y venidas, donde yo mismo tuve altibajos. Patu siempre me acompañó entendiendo mis tiempos, aún cuando eso implicó hacer presentaciones a último momento y a las corridas. Tengo mucho por agradecerle; estuvo presente cada vez que lo necesité, me dio la contención necesaria en esos momentos de frustración que tiene la tesis, y por si eso fuera poco, también me ayudó mucho orientándome para el final de la carrera y para lo que se viene después. En ese sentido también estoy muy contento con el grupo de Proba, con el que me siento muy a gusto y donde encontré más gente a la que llenar de preguntas, como Santi, Anita, Matt y Juli.

Esto se volvió ya una larga enumeración de nombres, y aún así si de agradecimientos se trata podría llegar hasta Ortigoza... En fin, espero no haberme olvidado de nadie, y repito mis fuertes agradecimientos a todos los que están cerca mío cada día. Sin esa compañía no sería la persona que soy, ni estaría en el lugar donde estoy.

Introducción

A lo largo de este trabajo recorreremos los procesos estocásticos desde las definiciones más generales, para adentrarnos en las cadenas de Markov y en sistemas de partículas. En función de buscar propiedades y demostrar resultados con respecto al sistema de partículas en el que nos centramos, desarrollamos inicialmente un marco teórico sobre el cual basarnos.

Los procesos estocásticos son familias de variables aleatorias que cambian a lo largo del tiempo: para cada tiempo t se encuentran en un determinado estado. Si bien estudiar un proceso estocástico puede dificultarse según las características del mismo, darle ciertas condiciones de regularidad permite tener un mayor manejo. Es por eso que apelamos a las cadenas de Markov, un caso particular de los procesos estocásticos que cumple una propiedad que es útil e intuitiva. Su idea fundamental es que la distribución de lo que le puede ocurrir en un futuro, condicionado al presente, es independiente del pasado.

La manera más gráfica de imaginarse una cadena de Markov la tenemos si tomamos una a tiempo discreto, como si tuviésemos «pasos» indexados por los números naturales. En cada paso conocemos las probabilidades para saber cuál va a ser el siguiente sólo mirando dónde estamos parados, sin importar el recorrido hecho hasta ese momento. Para el caso a tiempo continuo nos encontramos con que es una generalización de este mismo concepto. De hecho, la construcción que damos en este trabajo consiste en tomar una cadena a tiempo discreto, agregándole entre paso y paso tiempos de espera continuos con variables aleatorias con «pérdida de memoria», como son las variables exponenciales.

Estudiar el comportamiento de una cadena de Markov en tiempos largos, y por lo tanto, la existencia y unicidad de un comportamiento asintótico, tiene vital importancia. Una condición que garantiza esta existencia y unicidad es la irreducibilidad, es decir, la probabilidad no nula de llegar de un estado a cualquier otro, y la recurrencia positiva, es decir, un tiempo esperado finito de retorno a cualquier estado. El comportamiento asintótico está dado por la medida estacionaria, que a su vez es la única que cumple la propiedad de que si la cadena comienza distribuida según ella, para todos los tiempos siguientes la distribución del proceso está dada por ella misma.

Otro teorema que estudia el comportamiento asintótico es el de los promedios ergódicos, o simplemente Teorema ergódico. Con él se demuestra cómo la proporción de tiempo transcurrido en cada estado converge a una cantidad determinada, asociada a la medida estacionaria. Este teorema y una generalización para casos recurrentes positivos que

damos acá son de utilidad para las propiedades sobre las que trabajamos.

La motivación de nuestro trabajo la encontramos en un problema concreto estudiado durante el siglo XX, donde se modela cómo se esparce un gen en una población a lo largo del tiempo. El planteo original, hecho por Ronald Fisher, busca la función que describe este proceso, y la escribe como solución a una determinada ecuación diferencial. Si bien dicha solución no se puede escribir explícitamente, Fisher busca una solución en el equilibrio que indique cómo «viaja» el gen, y ese es el rol que cumplen las «ondas viajeras». El modelado en ese punto tiene un problema, que es la existencia de infinitos equilibrios, determinados según con qué velocidad de propagación comienza el gen su esparcimiento en el espacio. Para salvar esa discusión se propone un estudio microscópico en el frente de propagación del gen, y ahí es donde toma sentido el trabajo con sistemas de partículas.

El tipo de sistemas de partículas con el que nosotros trabajamos es el conocido como ramificación-selección. En particular le prestamos atención al estudio que hacen Durrett y Remenik del sistema de Brunet-Derrida, que consiste en N partículas que se desplazan en la recta real. Cada cierto tiempo alguna de ellas se ramifica o duplica, generando una partícula nueva que ocupa un lugar en la recta elegido a partir de una distribución determinada. Luego, de las $N + 1$ partículas que quedan se seleccionan N , de modo tal de mantener constante la cantidad de partículas. En este caso seleccionamos las N de la derecha o, equivalentemente, eliminamos la que está más a la izquierda.

En [8], Durrett y Remenik estudian la diferencia de velocidad entre el modelo con y sin eliminación de partículas; prueban la convergencia de la medida empírica a una medida de probabilidad determinística ν_t para todo t cuya densidad es solución de un problema de frontera libre; demuestran que el proceso es ergódico, por lo que hay una velocidad v_N de la nube en equilibrio, y que estas velocidades convergen a la velocidad de la onda viajera minimal cuando $N \rightarrow \infty$ (hecho conocido como *Principio de selección*).

Nosotros, mediante el análisis de un caso similar con herramientas nuevas, aportamos a la elaboración de una introducción a estos temas, para facilitar su lectura a quienes hayan alcanzado un nivel avanzado en la Licenciatura en Ciencias Matemáticas. El caso similar consiste en una adaptación en \mathbb{Z} del sistema recién descrito, cuya construcción gráfica damos detalladamente. En esta adaptación, en cada paso alguna de las partículas elige entre desplazarse, con probabilidad $1/2$ a izquierda y $1/2$ a derecha, y ramificarse, generando una partícula nueva que se queda en el mismo lugar que la original.

Dado este sistema, estudiamos algunos resultados que nos ayudaron a describir su comportamiento. Primero probamos la irreducibilidad, basándonos en que los desplazamientos de las partículas pueden darse en cualquier sentido. Llamamos 0 al estado en que todas las partículas se encuentran en la misma ubicación, y vimos que de cualquier estado se puede llegar al 0 , y del 0 se puede llegar a cualquier estado en una cantidad finita de pasos. Luego probamos la recurrencia positiva, viendo que el tiempo esperado de retorno al 0 tiene una esperanza finita, gracias al hecho de que si se repite N veces la ramificación

de la partícula que se encuentra más a la derecha, las N que quedan se ubican todas en la misma posición. De este hecho concluimos la existencia y unicidad de una medida invariante. Finalmente probamos la existencia de una velocidad asintótica para la primera partícula con un método de demostración distinto al utilizado por Durrett y Remenik, basándonos fuertemente en el Teorema ergódico.

Para la demostración que realizamos para este último punto, escribimos a la velocidad como un promedio de los saltos de la primera partícula, y luego categorizamos a esos saltos según los tres distintos movimientos que puede tener. Para probar la convergencia, nos definimos una sucesión de tiempos de parada $(\tau_n)_{n \in \mathbb{N}}$, que sirve para distinguir cuándo el sistema realiza ramificaciones y cuándo realiza desplazamientos.

Para cerrar la tesis, incluimos un último capítulo a modo de comentario, donde ponemos en discusión las motivaciones que encontró Fisher para estudiar este problema, y por lo tanto, la motivación detrás de lo estudiado hasta el día de hoy en el área. Fisher realiza el estudio de genes «ventajosos» en el contexto del auge de la eugenesia, una corriente que busca explicar científicamente los contrastes sociales como diferencias biológicas dentro de la especie. Esta corriente vino acompañada de teorías más generales (como el conocido «darwinismo social»), y tuvo consecuencias muy fuertes en el aspecto político con repercusiones mundialmente conocidas, como fueron los movimientos racistas de la primera mitad del siglo XX.

Este aspecto no anula lo estudiado, y en esta tesis se deja en claro por qué el autor cree que el estudio del modelo de Fisher es de gran utilidad, dado que sirve para otros problemas reales que tienen una aplicación necesaria. Sin embargo, este eje es tomado porque evidencia que las motivaciones científicas no son inocentes; por el contrario, se ve que la ciencia defiende determinados intereses, como toda producción cultural, y es por eso que no se la puede considerar neutral. Hecha esta observación, es tomada como pauta para hacer una mirada crítica del sistema científico, y expresar la voluntad del autor de que las investigaciones se orienten en búsqueda de una sociedad más justa e igualitaria.

Con todo esto, la tesis está estructurada de la siguiente manera:

En el Capítulo 1 tratamos las nociones básicas de las cadenas de Markov, junto a las definiciones y las propiedades que utilizamos.

En el Capítulo 2 describimos el problema que disparó el paper de Fisher, y las preguntas que al día de hoy se intentan responder en los trabajos de Durrett y Remenik.

En el Capítulo 3 profundizamos en el caso estudiado por nosotros. Desarrollamos los resultados de la irreducibilidad, de la existencia y unicidad de la medida invariante, y demostramos la existencia de una velocidad asintótica para la primera partícula.

En el Capítulo 4 disertamos en torno a las investigaciones científicas, la motivación social que tiene la eugenesia, y el rol al interior de cada sociedad.

En el Capítulo 5 proporcionamos un Apéndice con resultados necesarios pero no centrales en el desarrollo de la tesis.

Capítulo 1

Cadenas de Markov

En este capítulo vamos a dar el marco teórico que utilizamos a lo largo de la tesis, de modo tal de tener una base clara para lo siguiente. Este marco cuenta con definiciones y con resultados de vital importancia en el área; algunos de ellos los demostramos y otros sólo los enunciamos, en ambos casos explicitando las referencias utilizadas.

En líneas generales, tomamos referencias de los libros Markov Chains and Mixing Times, de D. Levin, Y. Peres y E. Wilmer ([18]), Markov Chains, de J. R. Norris ([22]), y de la tesis de licenciatura Distribuciones cuasiestacionarias y procesos de ramificación multitipo, de Analía Ferrari ([9]).

1.1. Un proceso aleatorio en el tiempo

Para comenzar introducimos a los procesos estocásticos:

Definición 1.1.1. Un proceso estocástico es una familia de variables aleatorias $(X_t)_{t \in I}$ definidas todas sobre un mismo espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$, e indexadas por un parámetro t que representará al tiempo. Si $X_t : \Omega \rightarrow \mathcal{S}$, \mathcal{S} es el espacio de estados que toma el proceso.

La utilización de dichos procesos es muy razonable en tanto nos permiten modelar situaciones concretas que dependen de factores aleatorios y que varían a través del tiempo. Si, por ejemplo, cada día hábil queremos ir a la Facultad de Ciencias Exactas y Naturales de la UBA, y tenemos que elegir entre tomarnos la línea 107 cuyo recorrido termina ahí en Ciudad Universitaria, o caminar desde nuestra casa, podemos tomar la decisión tirando una moneda equilibrada, independientemente de qué pasó en los demás días. Así, tenemos un proceso estocástico con el espacio de estados $\mathcal{S} = \{107, \text{caminar}\}$, indexado por el conjunto de tiempos $I = \{\text{días hábiles}\}$, y variables aleatorias independientes $(X_n)_{n \in I}$ con

$$\mathbb{P}(X_n = 107) = \mathbb{P}(X_n = \text{caminar}) = 1/2 \quad \forall n \in I.$$

Es decir, $X_n \sim Be(1/2) \quad \forall n \in I$, y tenemos una familia de variables independientes e idénticamente distribuidas (de ahora en más *iid*). En otros casos nos vamos a encontrar variables que sí van a depender entre ellas. Por ejemplo, si queremos modelar el caso en que después de tres días seguidos de esperar más de media hora el colectivo decidimos ya ni tirar la moneda y caminar, el resultado de cada variable se verá afectado por la información de variables anteriores.

Con esta definición amplia de proceso estocástico nos vamos a encontrar con muchas variantes, donde el tiempo es discreto o continuo ($I = \mathbb{N}$, $I = [0, +\infty)$, etc.), el espacio de estados es un conjunto numérico o no, y también puede ser discreto o continuo ($\mathcal{S} = \mathbb{Z}$, $\mathcal{S} = \mathbb{Z}_d = \frac{\mathbb{Z}}{d\mathbb{Z}}$, $\mathcal{S} = \mathbb{R}$, etc.).

1.2. Cadenas de Markov

Los procesos estocásticos son de muchos tipos distintos, y si no les ponemos restricciones resultan inmanejables. Nosotros vamos a restringir los espacios sobre los que vamos a trabajar, y nos vamos a centrar en una clase de procesos que cumple con determinadas propiedades que le dan una mayor flexibilidad para estudiarlos.

Antes de proceder a dar las definiciones correspondientes a la sección, vamos a hacer algunas consideraciones previas.

Vamos a considerar, hasta que señalemos lo contrario, que \mathcal{S} es un espacio finito o numerable y que I es un conjunto de tiempos a lo sumo numerable y discreto. Así, a cada elemento de I lo tomamos como un paso: a medida que recorremos los elementos de I , el proceso $(X_n)_{n \in I}$ va saltando de un paso a otro. Por ejemplo, si $I = \{\text{días hábiles}\}$ o $I = \mathbb{N}_0$. Para hablar de un caso general vamos a usar $I = \mathbb{N}$.

También vamos a ponerle un nombre al «comienzo del proceso»: nuestros procesos pueden empezar en un estado determinado, elegido arbitrariamente (es decir, si pedimos que $X_0 = s_j$ para algún $s_j \in \mathcal{S}$), o más en general, pueden estar determinados aleatoriamente a partir de alguna distribución en \mathcal{S} . Por eso definimos:

Definición 1.2.1. Sea $(X_n)_{n \geq 0}$ un proceso estocástico en $\mathcal{S} = \{s_1, s_2, \dots, s_k, \dots\}$ numerable. Llamamos μ o $\mu^{(0)}$ a la distribución inicial del proceso,

$$\mu^{(0)} = (\mu_1^{(0)}, \mu_2^{(0)}, \dots, \mu_k^{(0)}, \dots) = (\mathbb{P}(X_0 = s_1), \dots, \mathbb{P}(X_0 = s_k), \dots).$$

Cuando elijamos empezar el proceso en un estado determinístico s_j , la distribución inicial va a ser δ_{s_j} :

$$\mu^{(0)} = \delta_{s_j} := \begin{cases} 1 & \text{en la coordenada } j \\ 0 & \text{en las demás} \end{cases} \quad (1.1)$$

Denotamos como \mathbb{P}_i y \mathbb{E}_i a la probabilidad y a la esperanza inducidas por condicionar a que $X_0 = s_i$, respectivamente.

Ahora sí, una cadena de Markov es un proceso estocástico donde el estado en un determinado paso nos da la información suficiente para saber cuáles son las probabilidades para el siguiente paso. Es decir, como una propiedad de pérdida de memoria, no nos importa cómo llegamos a un estado, qué pasó en los pasos anteriores, para saber qué es lo que sigue. A esta propiedad la llamamos propiedad de Markov.

Más precisamente:

Definición 1.2.2. Sea $(X_n)_{n \in \mathbb{N}}$ un proceso estocástico sobre el espacio de estados \mathcal{S} . Decimos que $(X_n)_{n \in \mathbb{N}}$ es una *cadena de Markov* si

$$\begin{aligned} \mathbb{P}(X_{n+1} = s_{i_{n+1}} | X_n = s_{i_n}, \dots, X_0 = s_{i_0}) &= \mathbb{P}(X_{n+1} = s_{i_{n+1}} | X_n = s_{i_n}) \\ \forall n \in \mathbb{N}, \forall s_{i_0}, \dots, s_{i_{n+1}} \in \mathcal{S}, \text{ tal que } \mathbb{P}(X_0 = s_{i_0}, \dots, X_n = s_{i_n}) &> 0. \end{aligned}$$

Entonces la probabilidad de que una cadena de Markov esté en un estado en el paso $n + 1$ sólo depende del estado en que se encuentra en el paso n . Si además esas probabilidades sólo dependen de dichos estados y no de quién es n decimos que la cadena es *homogénea*.

Definición 1.2.3. Una cadena de Markov $(X_n)_{n \in \mathbb{N}}$ es *homogénea* si para todo $i, j \in \mathbb{N}$ existe un coeficiente p_{ij} que cumple que para todo $n \in \mathbb{N}$ y para toda tira de estados $s_{i_0}, \dots, s_{i_{n-1}}, s_i, s_j \in \mathcal{S}$

$$\mathbb{P}(X_{n+1} = s_j | X_n = s_i, \dots, X_0 = s_{i_0}) = \mathbb{P}(X_{n+1} = s_j | X_n = s_i) =: p_{ij}. \quad (1.2)$$

En general cuando hablemos de cadenas vamos a asumir que son cadenas homogéneas. Por esta razón, a una cadena de Markov le asignamos los coeficientes $(P)_{ij} = p_{ij}$ como en la definición anterior, siendo esta «la probabilidad de pasar del estado s_i al estado s_j » en un paso, y por eso definimos lo siguiente:

Definición 1.2.4. Sea J a lo sumo numerable, y P una matriz con entradas p_{ij} con $i, j \in J$. P se dice *matriz de transición* si verifica:

- $p_{ij} \geq 0$ para todo $i, j \in J$.
- $\sum_{j \in J} p_{ij} = 1$ para todo $i \in J$.

No es difícil ver que estas dos propiedades las cumplen los coeficientes p_{ij} de los que hablamos en (1.2). Por lo tanto, si P es una matriz cuyas entradas p_{ij} cumplen que $\mathbb{P}(X_{n+1} = s_j | X_n = s_i) = p_{ij} \forall n \in \mathbb{N}$, entonces P es una matriz de transición, y decimos que P es la *matriz de transición de la cadena* $(X_n)_{n \in \mathbb{N}}$.

Dada $(X_n)_{n \in \mathbb{N}}$ una cadena de Markov con distribución inicial μ y matriz de transición P , conocemos

$$\mathbb{P}(X_0 = s_i) = \mu_i.$$

$$\mathbb{P}(X_{n+1} = s_j | X_n = s_i) = p_{ij}.$$

y con esta información podemos dar dos caracterizaciones distintas para la cadena.

1. Podemos calcular la probabilidad de que la cadena pase por una determinada tira de estados hasta el paso n , condicionando paso a paso:

$$\begin{aligned} \mathbb{P}(X_0 = s_{i_0}, \dots, X_n = s_{i_n}) &= \mathbb{P}(X_0 = s_{i_0}, \dots, X_{n-1} = s_{i_{n-1}}) \cdot p_{i_{n-1}i_n} \\ &\vdots \\ &= \mu^{(0)}(i_0) p_{i_0i_1} p_{i_1i_2} \cdots p_{i_{n-1}i_n}. \end{aligned}$$

2. Para cada n llamemos $\mu^{(n)}$ al siguiente vector:

$$\mu^{(n)} = (\mu_1^{(n)}, \mu_2^{(n)}, \dots) = (\mathbb{P}(X_n = s_1), \mathbb{P}(X_n = s_2), \dots).$$

Este vector describe a la cadena a tiempo n , y también podemos calcular su distribución a partir de la distribución inicial y de la matriz de transición. Para eso, veamos que

- a) sabemos quién es cada uno de ellos si conocemos a la anterior y a la matriz de transición:

$$\mu^{(n)} = \mu^{(n-1)} P.$$

Demostración.

$$\begin{aligned} \mu_j^{(n)} &= \mathbb{P}(X_n = s_j) \\ &= \sum_{i \in \mathcal{S}} \mathbb{P}(X_n = s_j | X_{n-1} = s_i) \mathbb{P}(X_{n-1} = s_i) \\ &= \sum_{i \in \mathcal{S}} \mu_i^{(n-1)} p_{ij} \\ &= (\mu^{(n-1)} P)_j. \end{aligned}$$

□

- b) Generalizando la fórmula anterior, sabemos quién es $\mu^{(n)}$ sólo a partir de $\mu^{(0)}$ y la matriz de transición. Para cada n , la distribución $\mu^{(n)}$ satisface

$$\mu^{(n)} = \mu^{(0)} P^n.$$

Demostración. Veamos que la igualdad es cierta haciendo inducción en n .

Para $n = 1$ y $s_j \in \mathcal{S}$

$$\begin{aligned}\mu_j^{(1)} &= \mathbb{P}(X_1 = s_j) \\ &= \sum_{i \in \mathcal{S}} \mathbb{P}(X_1 = s_j | X_0 = s_i) \mathbb{P}(X_0 = s_i) \\ &= \sum_{i \in \mathcal{S}} \mu_i^{(0)} p_{ij} \\ &= (\mu^{(0)} P)_j.\end{aligned}$$

De donde,

$$\mu^{(1)} = \mu^{(0)} P.$$

Supongamos que vale para n , veamos que es cierto para $n + 1$

Por lo que afirmamos en (a)

$$\mu^{(n+1)} = \mu^{(n)} P.$$

Pero por la hipótesis inductiva, $\mu^{(n)} = \mu^{(0)} P^n$, entonces

$$\mu^{(n+1)} = \mu^{(0)} P^{n+1}.$$

□

Ejemplo 1.2.5. Un ejemplo clásico de cadena de Markov es un paseo al azar, donde tenemos una partícula en algún espacio, que en cada paso pega un salto aleatorio hacia una nueva ubicación.

Tomamos, por ejemplo, el espacio como \mathbb{Z}_d , Z_0 una variable con alguna distribución en \mathbb{Z}_d , la familia de variables aleatorias $(\xi_i)_{i \in \mathbb{N}}$ *iid* y la variable Z_n con

$$\mathbb{P}(\xi_i = +1) = \mathbb{P}(\xi_i = -1) = 1/2,$$

$$Z_n = Z_0 + \sum_{i=1}^n \xi_i$$

para $n \geq 1$ (tomando la respectiva congruencia módulo n).

En las próximas líneas estudiamos Z_n en detalle para ver que es una cadena de Markov homogénea.

Sea $j_0, \dots, j_n \in \mathbb{Z}_d$ tales que $\mathbb{P}(Z_n = j_n, \dots, Z_0 = j_0) > 0$. Entonces, utilizando la independencia de las variables $(\xi_i)_{i \in \mathbb{N}}$,

$$\begin{aligned}
\mathbb{P}(Z_n = j_n | Z_{n-1} = j_{n-1}, \dots, Z_0 = j_0) &= \frac{\mathbb{P}(Z_n = j_n, \dots, Z_0 = j_0)}{\mathbb{P}(Z_{n-1} = j_{n-1}, \dots, Z_0 = j_0)} \\
&= \frac{\mathbb{P}(Z_{n-1} + \xi_n = j_n, Z_{n-1} = j_{n-1}, \dots, Z_0 = j_0)}{\mathbb{P}(Z_{n-1} = j_{n-1}, \dots, Z_0 = j_0)} \\
&= \frac{\mathbb{P}(\xi_n = j_n - j_{n-1}, Z_{n-1} = j_{n-1}, \dots, Z_0 = j_0)}{\mathbb{P}(Z_{n-1} = j_{n-1}, \dots, Z_0 = j_0)} \\
&= \frac{\mathbb{P}(\xi_n = j_n - j_{n-1}) \cdot \mathbb{P}(Z_{n-1} = j_{n-1}, \dots, Z_0 = j_0)}{\mathbb{P}(Z_{n-1} = j_{n-1}, \dots, Z_0 = j_0)} \\
&= \mathbb{P}(\xi_n = j_n - j_{n-1}) \\
&= \frac{\mathbb{P}(\xi_n = j_n - j_{n-1}) \cdot \mathbb{P}(Z_{n-1} = j_{n-1})}{\mathbb{P}(Z_{n-1} = j_{n-1})} \\
&= \frac{\mathbb{P}(\xi_n = j_n - j_{n-1}, Z_{n-1} = j_{n-1})}{\mathbb{P}(Z_{n-1} = j_{n-1})} \\
&= \frac{\mathbb{P}(Z_{n-1} + \xi_n = j_n, Z_{n-1} = j_{n-1})}{\mathbb{P}(Z_{n-1} = j_{n-1})} \\
&= \mathbb{P}(Z_{n-1} + \xi_n = j_n | Z_{n-1} = j_{n-1}) \\
&= \mathbb{P}(Z_n = j_n | Z_{n-1} = j_{n-1}).
\end{aligned}$$

Nótese que la independencia de las $(\xi_i)_{i \in \mathbb{N}}$ la utilizamos en dos ocasiones. Como la cuenta no depende de n , la cadena es homogénea. Además, en la demostración vimos que

$$\mathbb{P}(Z_n = j_n | Z_{n-1} = j_{n-1}) = \mathbb{P}(\xi_n = j_n - j_{n-1}),$$

y ξ_n sólo puede ser 1 o -1 con probabilidad $1/2$ cada uno. Entonces la matriz de transición nos queda

$$\begin{aligned}
p_{s(s+1)} &= \mathbb{P}(Z_{n+1} = s+1 | Z_n = s) = 1/2 \\
p_{s(s-1)} &= \mathbb{P}(Z_{n+1} = s-1 | Z_n = s) = 1/2 \\
&(\forall s \in \mathbb{Z}_d).
\end{aligned}$$

En la Figura 1.1 mostramos el comportamiento del paseo.

1.3. Irreducibilidad y Aperiodicidad

En las cadenas de Markov nos va a interesar estudiar cómo se desplaza la cadena entre los diferentes estados, la frecuencia con la que retorna a ellos, los caminos que puede realizar, etc. En ese sentido, introducimos nuevos conceptos.

Definición 1.3.1. Una cadena de Markov $(X_n)_{n \in \mathbb{N}}$ en \mathcal{S} , con matriz de transición P se dice irreducible si $\forall s_i, s_j \in \mathcal{S}$ existe $n \in \mathbb{N}$ tal que $(P^n)_{ij} > 0$. Es decir, si hay probabilidad positiva de llegar de un estado a cualquier otro en alguna cantidad finita de pasos.

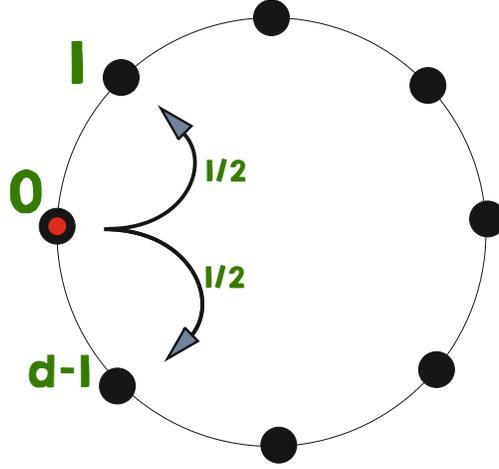


Figura 1.1: Paseo al azar en \mathbb{Z}^d

En [22] se considera una definición más amplia para la irreducibilidad. Allí el autor define una relación \rightarrow en el conjunto de espacios $\mathcal{S} \times \mathcal{S}$:

$$i \rightarrow j \text{ si } \exists n \in \mathbb{N}_0 / (P^n)_{ij} > 0.$$

En palabras, $i \rightarrow j$ si se puede llegar de s_i a s_j en alguna cantidad finita de pasos. Luego define $i \leftrightarrow j$ si también vale el recíproco de s_j a s_i . De esta forma, el espacio de estados \mathcal{S} se puede particionar en clases de equivalencia, y la irreducibilidad se define para cuando el espacio \mathcal{S} tiene una única clase. A estas clases de equivalencia Norris las llama clases de comunicación.

Esta definición es equivalente a la que dimos nosotros: tener una única clase significa que todo par de estados $s_i, s_j \in \mathcal{S}$ están comunicados: ($i \leftrightarrow j$), que es lo mismo que decir que siempre se puede llegar de un estado a otro en una cantidad finita de pasos.

Definición 1.3.2. El período de un estado $s_i \in \mathcal{S}$ es $\text{mcd}\mathcal{T}(i)$, donde mcd es el máximo común divisor y

$$\mathcal{T}(i) = \{n \geq 1 / (P^n)_{ii} > 0\}.$$

Definición 1.3.3. Si $\text{mcd}\mathcal{T}(i) = 1 \forall s_i$ decimos que la cadena es aperiódica.

Lema 1.3.4. Si la cadena $(X_n)_{n \in \mathbb{N}}$ es irreducible, entonces

$$\text{mcd}\mathcal{T}(i) = \text{mcd}\mathcal{T}(j) \forall s_i, s_j \in \mathcal{S}.$$

Demostración. Dados dos estados s_i y s_j , por la irreducibilidad sabemos que hay dos números naturales r y k tales que $(P^r)_{ij} > 0$ y $(P^k)_{ji} > 0$. Sea $l = r + k$, tenemos:

- $l \in \mathcal{T}(i)$ y $l \in \mathcal{T}(j)$
- $\mathcal{T}(i) + m \subseteq \mathcal{T}(j)$ o, lo que es equivalente, $\mathcal{T}(i) \subseteq \mathcal{T}(j) - l$. Veámoslo:

Sea $t_0 \in \mathcal{T}(i)$: la cadena puede salir y volver del estado s_i en t_0 pasos. Si la cadena empieza en el estado s_j , en k pasos llega al estado s_i , en t_0 pasos vuelve al estado s_i , y en r pasos más va del estado s_i al estado s_j . Entonces $k + t_0 + r = t_0 + l \in \mathcal{T}(j)$.

- Con estos dos puntos anteriores, el $\text{mcd}\mathcal{T}(j)$ divide a todos los elementos de $\mathcal{T}(i)$:

Sea $t_0 \in \mathcal{T}(i)$. Por el último resultado, $t_0 = t_1 - l$, con $t_1 \in \mathcal{T}(j)$. Como también $l \in \mathcal{T}(j)$, el $\text{mcd}\mathcal{T}(j)$ divide a la resta.

De esta forma concluimos que $\text{mcd}\mathcal{T}(j) \leq \text{mcd}\mathcal{T}(i)$. Análogamente $\text{mcd}\mathcal{T}(i) \leq \text{mcd}\mathcal{T}(j)$. \square

Este lema es muy útil, porque cada vez que tengamos una cadena irreducible, alcanza con ver que $\text{mcd}\mathcal{T}(i) = 1$ para algún estado $s_i \in \mathcal{S}$ para comprobar su aperiodicidad.

Ejemplo 1.3.5. Para entender cuándo una cadena es aperiódica y cuándo no, un buen ejemplo es el paseo al azar en \mathbb{Z}_d con d un número par. Con las mismas definiciones que en el Ejemplo 1.2.5, en el caso en que d es par Z_n mantiene la paridad, es decir: Z_n es par para n par, y es impar para n impar.

Por eso mismo, para cualquier estado la cadena sólo puede retornar en tiempos pares y por lo tanto $\text{mcd}\mathcal{T}(i) \geq 2$ para cualquier $s_i \in \mathcal{S}$.

Como la aperiodicidad es una propiedad que queremos tener en las cadenas, para solucionar estos casos se utiliza la siguiente modificación:

Definición 1.3.6. Dado un paseo al azar como explicamos en el Ejemplo 1.2.5, se le llama paseo al azar «lazy» al paseo que en cada paso se queda en su lugar con probabilidad $\frac{1}{2}$.

Utilizando las mismas variables aleatorias, la distribución de las $(\xi_i)_{i \in \mathbb{N}}$ quedaría $\mathbb{P}(\xi_i = +1) = \mathbb{P}(\xi_i = -1) = 1/4$, y $\mathbb{P}(\xi_i = 0) = \frac{1}{2}$. Notar que en este caso la cadena puede volver a cualquier estado en 1 paso, y eso la hace aperiódica.

1.4. Medidas Invariantes

Dado que las cadenas de Markov modelan procesos que se extienden en el tiempo, se vuelve interesante estudiar su comportamiento para tiempos grandes, y ver en ese sentido qué tipo de convergencias nos podemos encontrar. Consideremos una cadena de Markov

$(X_n)_{n \in \mathbb{N}}$ que toma valores en \mathcal{S} , con distribución inicial μ y matriz de transición P , y tomemos \mathcal{S} finito. Como dijimos anteriormente en esta sección, tenemos una forma de describir a la distribución de la cadena en cada paso a partir de la distribución anterior:

$$\mu^{(n)} = \mu^{(n-1)}P$$

En caso de que exista un límite π para $\mu^{(n)}$ cuando $n \rightarrow \infty$, tomando límite de los dos lados de la igualdad encontramos que tiene que cumplir que:

$$\pi = \pi P \tag{1.3}$$

Como \mathcal{S} es finito, la convergencia del producto del lado derecho se puede ver coordenada a coordenada, dado que cada una de ellas es una suma finita.

Volvemos a considerar \mathcal{S} a lo sumo numerable y extendemos esta noción.

Definición 1.4.1. Si tenemos una medida π que satisface (1.3), la llamamos *medida invariante* o *medida estacionaria*. Si además π es una distribución (i.e. si $\sum_{s \in \mathcal{S}} \pi(s) = 1$), la llamamos *distribución invariante* o *distribución estacionaria*.

¿Por qué ese nombre? Las distribuciones invariantes cumplen esencialmente dos propiedades. La primera, que ya comentamos para el caso finito, es que son el límite para las $\mu^{(n)}$ cuando este existe. Pero además cumplen que si la cadena en algún paso está distribuida según alguna de ellas, la distribución para el paso siguiente es la misma, porque

$$\mu^{(n+1)} = \mu^{(n)}P = \pi P = \pi.$$

Razonando inductivamente concluimos que si la cadena toma una distribución invariante, se mantiene con esta para siempre.

En lo que sigue, estudiamos condiciones que resultan suficientes para que una cadena de Markov tenga una única distribución invariante. Para eso estudiamos previamente los llamados estados «recurrentes».

Consideremos un tiempo τ_i que cuente la cantidad de pasos que tarda la cadena en volver al estado s_i si ella comenzó allí. Es decir,

$$\tau_i = \inf \{n \geq 1 : X_n = s_i\}.$$

Definición 1.4.2. El estado $s_i \in \mathcal{S}$ se dice *recurrente* si $\mathbb{P}_i(\tau_i < \infty) = 1$.

La propiedad de recurrencia se mantiene para una clase de comunicación (como definimos en la Sección 1.3): o todos los estados en una clase son recurrentes o ninguno lo es. Por eso hablamos de si una clase es recurrente o no. Más aún, si una cadena X_n es irreducible, recordemos que tiene una sola clase, y por eso podemos hablar de si la cadena es o no recurrente.

Una demostración de esta afirmación se puede leer en [22, p. 24 a 27].

Definición 1.4.3. Si un estado recurrente además cumple que $m_i = \mathbb{E}_i(\tau_i) < \infty$, decimos que el estado es recurrente positivo.

Un estado recurrente positivo es recurrente, pero se puede ver que no vale la vuelta.

Sea ahora un estado $s_k \in \mathcal{S}$. Para él y para cada i , consideramos la cantidad de visitas a s_i antes de volver a s_k :

$$\gamma_i^k = \mathbb{E}_k \left(\sum_{n=0}^{\tau_k-1} \mathbf{1}_{\{X_n=s_i\}} \right).$$

Notemos que si contamos la cantidad de visitas a todos los estados antes de volver a s_k , contamos exactamente el tiempo que tarda en volver a s_k . Es decir,

$$\sum_{j \in J} \gamma_j^k = m_k.$$

Dadas las definiciones, vemos algunos resultados al respecto.

Proposición 1.4.4. Si la cadena X_n es irreducible y recurrente, entonces $\forall k \in I$

1. $\gamma_k^k = 1$.
2. $\gamma^k = (\gamma_i^k)_{i \in I}$ cumple que $\gamma^k P = \gamma^k$.
3. $0 < \gamma_i^k < \infty$.

Demostración. 1. Por definición, τ_k es el tiempo del primer retorno, por lo que la cadena no pasa por k más que en el paso inicial.

2. Como el evento $\tau_k \geq n$ depende de las variables X_1, \dots, X_{n-1} , podemos usar la propiedad de Markov para calcular la probabilidad del siguiente evento:

$$\begin{aligned} \mathbb{P}_k(X_{n-1} = s_i, X_n = s_j \text{ y } \tau_k \geq n) &= \\ &= \mathbb{P}_k(X_n = s_j | X_{n-1} = s_i \text{ y } \tau_k \geq n) \cdot \mathbb{P}_k(X_{n-1} = s_i \text{ y } \tau_k \geq n) \\ &= \mathbb{P}_k(X_n = s_j | X_{n-1} = s_i) \cdot \mathbb{P}_k(X_{n-1} = s_i \text{ y } \tau_k \geq n) \\ &= \mathbb{P}_k(X_{n-1} = s_i \text{ y } \tau_k \geq n) p_{ij}. \end{aligned} \tag{1.4}$$

Como X_n es recurrente, $\mathbb{P}_k(\tau_k < \infty) = 1$, y además $\mathbb{P}_k(X_0 = X_{\tau_k} = s_k) = 1$, se

sigue que si $j \neq k$:

$$\begin{aligned}
\gamma_j^k &= \mathbb{E}_k \left(\sum_{n=1}^{\tau_k} \mathbf{1}_{\{X_n = s_j\}} \right) = \mathbb{E}_k \left(\sum_{n=1}^{\infty} \mathbf{1}_{\{X_n = s_j \text{ y } \tau_k \geq n\}} \right) \\
&= \sum_{n=1}^{\infty} \mathbb{P}_k (X_n = s_j \text{ y } \tau_k \geq n) \\
&= \sum_{i \in I} \sum_{n=1}^{\infty} \mathbb{P}_k (X_{n-1} = s_i, X_n = s_j \text{ y } \tau_k \geq n) \\
&= \sum_{i \in I} p_{ij} \sum_{n=1}^{\infty} \mathbb{P}_k (X_{n-1} = s_i \text{ y } \tau_k \geq n) \\
&= \sum_{i \in I} p_{ij} \mathbb{E}_k \left(\sum_{m=0}^{\infty} \mathbf{1}_{\{X_m = s_i \text{ y } \tau_k - 1 \geq m\}} \right) \\
&= \sum_{i \in I} p_{ij} \mathbb{E}_k \left(\sum_{m=0}^{\tau_k - 1} \mathbf{1}_{\{X_m = s_i\}} \right) = \sum_{i \in I} \gamma_i^k p_{ij}.
\end{aligned} \tag{1.5}$$

3. Como X_n es irreducible, existen $n, m \geq 0$ que cumplen que $p_{ik}^{(n)}, p_{ki}^{(m)} > 0$. Entonces por (1) y por (2):

- $\gamma_i^k \geq \gamma_k^k p_{ki}^{(m)} = p_{ki}^{(m)} > 0$.
- $\gamma_i^k p_{ik}^{(n)} \leq \gamma_k^k = 1 \Rightarrow \gamma_i^k < \infty$.

□

Notemos que de esta proposición concluimos que si una cadena es irreducible y recurrente tiene una medida invariante, pero esta no necesariamente es una distribución. Esta sutileza sólo la podremos salvar si la cadena es recurrente positiva, como veremos en las próximas líneas.

Proposición 1.4.5. Si la cadena $(X_n)_{n \in \mathbb{N}}$ es irreducible y $\pi = (\pi_i)_{i \in J}$ es una medida invariante con $\pi_k = 1$ para algún k , entonces para cada i , $\pi_i \geq \gamma_i^k$. Si $(X_n)_{n \in \mathbb{N}}$ es recurrente, cada desigualdad es de hecho una igualdad, y por lo tanto $\pi = \gamma^k$.

Demostración. π cumple la ecuación $\pi = \pi P$. Para cada j ,

$$\begin{aligned}
\pi_j &= \sum_{i_0 \in J} \pi_{i_0} p_{i_0 j} = \sum_{i_0 \neq k} \pi_{i_0} p_{i_0 j} + 1 \cdot p_{kj} \\
&= \sum_{i_0 \neq k} \left(\sum_{i_1 \in J} \pi_{i_1} p_{i_1 i_0} \right) p_{i_0 j} + p_{kj} \\
&= \sum_{i_0 \neq k} \left(\sum_{i_1 \neq k} \pi_{i_1} p_{i_1 i_0} \right) p_{i_0 j} + \sum_{i_0 \neq k} p_{ki_0} p_{i_0 j} + p_{kj} \\
&= \sum_{i_0, i_1 \neq k} \pi_{i_1} p_{i_1 i_0} p_{i_0 j} + \left(p_{kj} + \sum_{i_0 \neq k} p_{ki_0} p_{i_0 j} \right) \\
&\quad \vdots \\
&= \sum_{i_0, \dots, i_n \neq k} \pi_{i_n} p_{i_n i_{n-1}} \cdots p_{i_0 j} \\
&\quad + \left(p_{kj} + \sum_{i_0 \neq k} p_{ki_0} p_{i_0 j} + \cdots + \sum_{i_0, \dots, i_{n-1} \neq k} p_{ki_{n-1}} \cdots p_{i_0 j} \right) \\
&\geq \mathbb{P}_k(X_1 = j \text{ y } \tau_k \geq 1) + \mathbb{P}_k(X_2 = j \text{ y } \tau_k \geq 2) \\
&\quad + \cdots + \mathbb{P}_k(X_n = j \text{ y } \tau_k \geq n) \\
&\xrightarrow{n \rightarrow \infty} \gamma_j^k.
\end{aligned} \tag{1.6}$$

$$\Rightarrow \pi_j \geq \gamma_j^k.$$

En el caso en que $(X_n)_{n \in \mathbb{N}}$ es recurrente, llamamos $\mu = \pi - \gamma^k$, y tenemos que:

1. $\mu \geq 0$ por lo demostrado recién.
2. por la Proposición 1.4.4 γ^k es una medida invariante, y por lo tanto también lo es μ .
3. $\mu_k = 1 - 1 = 0$.

Dado cualquier $i \in J$, como $(X_n)_{n \in \mathbb{N}}$ es irreducible, existe un n tal que $p_{ik}^{(n)} > 0$.

Como $0 = \mu_k = \sum_{j \in J} \mu_j p_{jk}^{(n)} \geq \mu_i p_{ik}^{(n)}$, $\mu_i = 0$, y por lo tanto $\pi = \gamma^k$.

□

Proposición 1.4.6 (Existencia y unicidad de distribución invariante). Si el proceso X_n es irreducible, las siguientes propiedades son equivalentes:

1. Todos los estados son recurrentes positivos.
2. Algún estado es recurrente positivo.

3. La cadena tiene una distribución invariante π .

Cuando esto sucede, $\pi_i = \frac{1}{m_i}$, y de esto se deduce la unicidad.

Demostración. 1) \Rightarrow 2) es obvia.

2) \Rightarrow 3) Si i es recurrente positivo, en particular es recurrente, y por lo tanto $(X_n)_{n \in \mathbb{N}}$ lo es también. Por la Proposición 1.4.4, γ^i es una medida invariante. Pero además

$$\sum_{j \in J} \gamma_j^i = m_i < \infty.$$

Entonces $\pi = \gamma^i/m_i$ es una distribución invariante.

3) \Rightarrow 1) Fijemos un estado s_k . Como π es una distribución invariante, hay algún i para el cual $\pi_i > 0$, y como P es irreducible hay un n tal que $p_{ik}^{(n)} > 0$. Luego, $\pi_k = \sum_{j \in J} \pi_j p_{jk}^{(n)} \geq \pi_i p_{ik}^{(n)} > 0$. Para cada i definimos $\lambda_i = \pi_i/\pi_k$, que es una medida invariante que cumple que $\lambda_k = 1$. Por la Proposición 1.4.5, $\lambda_i \geq \gamma_i^k$. Entonces

$$m_k = \sum_{i \in J} \gamma_i^k \leq \sum_{i \in J} \frac{\pi_i}{\pi_k} = \frac{1}{\pi_k} < \infty. \quad (1.7)$$

y k es recurrente positivo.

Demostradas las equivalencias, para calcular explícitamente la distribución invariante retomamos el argumento utilizado en 3) \Rightarrow 1). Como ya sabemos que dadas estas condiciones $(X_n)_{n \in \mathbb{N}}$ es recurrente, nuevamente por la Proposición 1.4.5 vale que $\lambda = \gamma^k$, y eso implica que la desigualdad en (1.7) es en realidad una igualdad. \square

1.5. Teorema Ergódico

Además del estudio de las medidas invariantes, tenemos otras formas de estudiar las propiedades asintóticas de las cadenas de Markov. Si bien acá no tenemos independencia entre las variables aleatorias, sí tenemos en cambio otra propiedad que va en ese sentido: que condicionando al paso n lo que sigue es independiente de lo que pasó. Por eso, para estudiar límites de promedios de alguna función de la cadena en \mathbb{R} tenemos que adaptar la Ley de los Grandes Números a lo que esta última propiedad nos permite.

En el libro de Norris [22, p. 52] encontramos demostrado un teorema que estudia la proporción de tiempo transcurrido en cada estado cuando $n \rightarrow \infty$, y esto se construye a partir de los promedios de las indicadoras de dichos estados. En su demostración, la Ley de los Grandes Números cumple un rol fundamental, y a su vez, este teorema se usa de paso intermedio para demostrar una versión más general para el caso recurrente positivo, que es el que acá citamos.

Definición 1.5.1. Dada $(X_n)_{n \in \mathbb{N}}$ una cadena de Markov con distribución estacionaria π , $f : \mathcal{S} \rightarrow \mathbb{R}$ se dice integrable con respecto a π si $\mathbb{E}_\pi(f(X_1)) = \sum_{i \in I} f(s_i)\pi(s_i) < \infty$.

Teorema 1.5.2 (Teorema Ergódico). Sea $(X_n)_{n \in \mathbb{N}}$ una cadena de Markov irreducible y recurrente positiva con distribución estacionaria π , y sea $f : \mathcal{S} \rightarrow \mathbb{R}$ integrable con respecto a π . Entonces, para cualquier distribución inicial de la cadena,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=0}^{n-1} f(X_j) = \sum_{i \in I} f(s_i) \pi(s_i) = \mathbb{E}_\pi(f(X_1)).$$

El Teorema ergódico es una herramienta muy importante para los resultados que exponemos en el Capítulo 3.

1.6. La versión en tiempo continuo

Retomando el hilo principal del capítulo, nos interesa estudiar cadenas de Markov indexadas por un parámetro continuo. Vamos a continuar con espacios \mathcal{S} finitos o numerables, y por lo tanto la cadena todavía se va a manejar con «saltos» de un estado a otro. La diferencia va a ser que aquellos tiempos en los que la cadena decida hacer esos saltos (que los vamos a llamar *tiempos de espera*) van a ser elegidos aleatoriamente.

Para esto tomamos definiciones de la Sección 1.2, ahora con $I = [0, +\infty)$, donde I es el conjunto de índices correspondiente al proceso $(X_t)_{t \in I}$. Para hacer esta construcción hay muchas alternativas conocidas, como la de Ferrari y Galves [10] con acoplamientos de variables uniformes, la de Norris [22] con matrices de tazas Q o la de Seppalainen [24] construida con tiempos exponenciales sobre cadenas a tiempo discreto, aunque son todas equivalentes. Aquí vamos a desarrollar la última, que es la que más apropiada nos pareció para los casos que vamos a trabajar.

Notemos que ante este nuevo panorama, dado un estado inicial, la descripción del proceso tiene que indicar (1) cuáles son las distribuciones de los tiempos de espera aleatorios, y (2) cómo elegir el nuevo estado al que llega la cadena una vez que se realiza el salto. Daremos una motivación de cómo determinar ambos aspectos, para poder entender de dónde sale la definición que enunciaremos luego.

Nosotros queremos mantener la propiedad de Markov que dice que condicionando al pasado sólo nos interesa el estado presente para determinar el futuro de la cadena, como pasaba en la Definición 1.2.2. Por eso, los tiempos de los que hablamos en (1) sólo pueden depender del estado s_i donde se encuentra la cadena en cada momento, y no del tiempo ya transcurrido en s_i . Es decir, si la cadena está en un momento en s_i y T es el tiempo restante que se va a quedar ahí, T tiene que cumplir

$$\mathbb{P}(T > s + t | X > t) = \mathbb{P}(T > s) \quad \forall s, t \in \mathbb{R}_{\geq 0}.$$

Esta propiedad se la conoce como propiedad de falta de memoria, y como mostramos en el apéndice en el Teorema 5.1.1, las únicas variables continuas que cumplen con esto

son las exponenciales. Entonces, en cada estado $s_i \in \mathcal{S}$ el tiempo de espera va a ser una variable aleatoria con distribución exponencial y con esperanza $c(s_i)^{-1}$ (es decir, con parámetro $c(s_i)$).

Por otro lado, para definir lo comentado en (2) y mantener la misma propiedad de Markov, necesitamos determinar probabilidades de llegar a cada estado s_j sabiendo que salimos de un estado s_i , y las vamos a llamar $p(i, j)$. Este nombre ya lo utilizamos, y no es casual: es la matriz de transición con la que describíamos los saltos de una cadenas a tiempo discreto.

Con todas estas consideraciones, construimos una cadena $(X_t)_{t \in \mathbb{R}_{>0}}$ a tiempo continuo partiendo de lo siguiente: sea $(Y_n)_{n \in \mathbb{N}_0}$ una cadena de Markov a tiempo discreto, con distribución inicial μ y matriz de transición P , $(\tau_n)_{n \in \mathbb{N}_0}$ una familia de variables aleatorias *iid* con distribución exponencial de parámetro 1, e independiente de $(Y_n)_{n \in \mathbb{N}}$, y $c : \mathcal{S} \rightarrow \mathbb{R}_{>0}$ la función que va a determinar la esperanza de los tiempos de espera.

Como X_t va a visitar sucesivamente Y_0, Y_1, Y_2, \dots , definimos los tiempos de espera $S_n = c(Y_n)^{-1} \tau_n$ que son independientes de $(\mathcal{S}_k, Y_k)_{0 \leq k \leq n-1}$ y tienen distribución exponencial con esperanza $c(Y_n)^{-1}$, como se puede concluir de la Proposición 5.1.6 del Apéndice. Luego, definimos $J_0 = 0$ y $J_n = \sum_{k=0}^{n-1} S_k$ para $n \geq 1$, y para $n \in \mathbb{N}_0$:

$$X_t = Y_n \text{ para } J_n \leq t < J_{n+1}.$$

Esta definición sirve para todo t en $\mathbb{R}_{>0}$ en caso de que $\lim_{n \rightarrow \infty} J_n = \sum_{n \in \mathbb{N}} S_n = \infty$. Si esto no sucediera, se dice que la cadena tiene *explosión*, y ahí podemos determinar que la cadena está definida sólo hasta ese tiempo, o volver a hacer una construcción de este estilo para los tiempos siguientes. Con algunas condiciones simples las cadenas no van a tener explosión; si por ejemplo existe $C \in \mathbb{R}$ para el cual $c(x) \leq C$, entonces para todo n $\mathbb{E}(S_n) \geq C^{-1}$ y por lo tanto $\mathbb{E}(\sum_{n \in \mathbb{N}} S_n) = \sum_{n \in \mathbb{N}} \mathbb{E}(S_n) = \infty$ por convergencia monótona. En principio vamos a considerar cadenas con explosión, pero por cómo las vamos a ir construyendo, nos vamos a encontrar con que las más importantes que utilizamos en este trabajo no van a tener.

Hacemos la observación de que cuando no hay explosión, la función que asigna $t \mapsto X_t$ es continua a derecha y tiene límite a izquierda en todo t . Dentro de los procesos a tiempo continuo, los que cumplen con esto tienen propiedades cómodas de estudiar, y son conocidos como procesos *cádlág* (por su escritura en francés).

Teniendo ya una construcción apropiada, damos la definición correspondiente:

Definición 1.6.1. Sea $(X_t)_{t \in I}$ un proceso estocástico, con $I \subseteq \mathbb{R}_{>0}$ un intervalo real, $X_t : \Omega \rightarrow \mathcal{S}$, \mathcal{S} finito o numerable. Decimos que $(X_t)_{t \in I}$ es una *cadena de Markov a tiempo continuo* si existen $c : \mathcal{S} \rightarrow \mathbb{R}_{>0}$ y una cadena de Markov $(Y_n)_{n \in \mathbb{N}_0}$ tales que, si definimos $S_n = c(Y_n)^{-1} \tau_n$, $J_0 = 0$ y $J_n = \sum_{k=0}^{n-1} S_k$ para $n \geq 1$, cumplen que

$$X_t = Y_n \text{ para } J_n \leq t < J_{n+1}.$$

A la cadena Y_n se la conoce como *cadena de saltos* o *esqueleto* y a su matriz P como la matriz de saltos. De nuestra construcción se deduce que $Y_n = X_{J_n}$. Nos basaremos fuertemente en ella para trabajar con cadenas a tiempo continuo.

Para finalizar la construcción, vamos a enunciar una propiedad de Markov equivalente a la que dimos en la Definición 1.2.2. En ese sentido, definimos la *probabilidad de transición* de s_i a s_j en un tiempo t como $p_t(i, j) := \mathbb{P}_i(X_t = s_j)$, y en [23, Cap. 5] podemos encontrar demostrado lo siguiente.

Proposición 1.6.2. Dada $(X_t)_{t \in I}$ una cadena de Markov, para toda tira de tiempos $0 = t_0 < t_1 < \dots < t_n$ y estados s_{i_0}, \dots, s_{i_n} vale que

$$\mathbb{P}(X_{t_0} = s_{i_0}, \dots, X_{t_n} = s_{i_n}) = \mu(i_0) p_{t_1 - t_0}(i_0, i_1) \dots p_{t_n - t_{n-1}}(i_{n-1}, i_n).$$

Corolario 1.6.3 (Propiedad de Markov). Bajo las mismas hipótesis,

$$\mathbb{P}(X_{t_n} = s_{i_n} | X_{t_{n-1}} = s_{i_{n-1}}, \dots, X_{t_0} = s_{i_0}) = p_{t_n - t_{n-1}}(i_{n-1}, i_n).$$

En lo que sigue damos algunas definiciones y resultados para las cadenas en tiempo continuo, notando que en su mayoría son directas adaptaciones del caso discreto.

Definición 1.6.4. La cadena $(X_t)_{t \in \mathbb{R}_{>0}}$ tiene una distribución invariante π si esta cumple que

$$\mathbb{P}_\pi(X_t = \cdot) = \pi(\cdot) \quad \forall t$$

No es difícil ver que una cadena de Markov a tiempo continuo tiene invariante si y sólo si tiene su cadena de saltos.

Definición 1.6.5. Decimos que el estado i es recurrente si $\mathbb{P}(T_i < \infty) = 1$, donde

$$T_i = \inf \{t \geq J_1(\omega) : X_t(\omega) = i\}.$$

Definición 1.6.6. Si un estado recurrente además cumple que $m_i = \mathbb{E}_i(T_i) < \infty$, decimos que el estado es recurrente positivo.

Proposición 1.6.7. Un estado recurrente positivo para X_t si y sólo si lo es para la cadena de saltos Y_n

Definición 1.6.8. Una cadena de Markov $(X_t)_{t \in \mathbb{R}_{>0}}$ es *irreducible* si para todos $s_i, s_j \in \mathcal{S}$

$$\mathbb{P}_i(X_t = j \text{ para algún } t > 0) > 0.$$

Proposición 1.6.9. Si la cadena X_t es irreducible, las siguientes propiedades son equivalentes:

- a) Todos los estados son positivos recurrentes.
- b) Algún estado es recurrente positivo.
- c) la cadena no tiene explosión y tiene una distribución invariante μ .

Se pueden encontrar las demostraciones de estos resultados en [22, Cap. 3].

1.7. Procesos de Poisson

El ejemplo más usual de cadenas de Markov a tiempo continuo son los procesos de Poisson, que consisten en «tomar nota» cada vez que pasa un tiempo exponencial (con tiempos que tienen siempre el mismo parámetro λ). Es un proceso que sirve para modelar muchos problemas usuales, como la cola en un banco o la llegada de paquetes de información a un servidor. Más precisamente,

Definición 1.7.1. Un proceso estocástico $(X_t)_{t \geq 0}$ es un *proceso de Poisson de parámetro λ* si los tiempos de espera S_1, S_2, \dots son variables aleatorias exponenciales de parámetro λ , y la cadena de saltos asociada es $X_{J_n} = n$ (es decir, el proceso en cada salto suma 1). Vamos a notar $(X_t)_{t \geq 0} \sim PP(\lambda)$.

Una propiedad muy importante que tienen los procesos de Poisson es que dado un t fijo, $X_t \sim P(\lambda t)$ (es decir, tiene distribución de Poisson con parámetro λt).

Estas cadenas se suelen utilizar también como herramienta accesoria para construir otras cadenas a tiempo continuo de la siguiente manera: cada vez que en un proceso de Poisson ω aparece una marca, definimos X_t entre esa marca y la siguiente, como hicimos antes con los tiempos S_n . Esta construcción la vamos a utilizar para nuestro sistema de partículas en el Capítulo 3.

1.8. Tiempos de parada

Retomemos por un momento el ejemplo que comentamos en la sección 1.1, del proceso en que elegimos cómo llegar a la Facultad a partir de tirar una moneda cada día hábil, y modifiquémoslo para que sea una cadena de Markov. Supongamos que tenemos una moneda no equilibrada que sale cara con probabilidad $3/4$, y que cada vez que el resultado es cara, tomamos la misma acción que habíamos tomado el día anterior. Así nos queda una cadena homogénea con matriz de transición de tamaño 2×2 , con los siguientes valores para cualquier n :

$$\mathbb{P}(X_{n+1} = 107 | X_n = 107) = \mathbb{P}(X_{n+1} = \text{caminar} | X_n = \text{caminar}) = 3/4$$

$$\mathbb{P}(X_{n+1} = \text{caminar} | X_n = 107) = \mathbb{P}(X_{n+1} = 107 | X_n = \text{caminar}) = 1/4.$$

En general, a una cadena de Markov la tenemos definida en todo el espacio de tiempos I asociado a ella, pero en algunas ocasiones nos va a resultar más interesante (o más útil) mirarla sólo en algunos de estos tiempos, que pueden ser determinísticos o aleatorios.

Si en nuestro ejemplo nos interesa determinar *a)* cuándo viajamos tres días seguidos en 107 por primera vez, o *b)* en qué momento se da la mayor seguidilla de caminatas seguidas, tenemos dos *tiempos aleatorios* que son esencialmente distintos. Al primero lo podemos determinar observando la cadena paso a paso y deteniéndola cuando ocurre dicho evento,

y por eso lo vamos a llamar *tiempo de parada*. En el otro caso, en cambio, es imposible dar información repitiendo el mismo algoritmo: para afirmar que tuvimos la mayor seguidilla es necesario saber que las que vinieron después fueron más cortas.

Esta construcción intuitiva la tomamos de un ejemplo equivalente hecho con el valor de una acción en [18, Sección 6.2.1].

Formalizamos los conceptos:

Definición 1.8.1. Sea $(X_t)_{t \in I}$ una cadena de Markov en el espacio $(\Omega, \mathcal{F}, \mathbb{P})$, siendo $I = \mathbb{N}$ o $I = \mathbb{R}$.

- Un *tiempo aleatorio* T es una variable aleatoria, con valores en $I \cup \{+\infty\}$.
- Un tiempo aleatorio τ se dice *tiempo de parada* de la cadena de Markov si $\{\tau \leq t\} \in \mathcal{F}_t$, donde \mathcal{F}_t denota a la σ -álgebra generada por $\{X_s : s \leq t\}$.

Los tiempos de parada de las cadenas de Markov cumplen que si «reiniciamos» la cadena cuando suceden, el proceso resultante es de nuevo una cadena de Markov con la misma cadena de saltos. Se trata de una propiedad fundamental conocida como *Propiedad de Markov fuerte*, cuya demostración no vamos a dar en esta tesis. La misma se puede encontrar en [22, p. 227].

Proposición 1.8.2 (Propiedad de Markov fuerte). Dada una cadena de Markov $(X_t)_{t \in I}$ con distribución inicial μ y matriz de saltos P , y un tiempo de parada τ , si $\tau < \infty$ y $X_\tau = s_i$, entonces $(X_{\tau+t})_{t \in I}$ es una cadena de Markov con distribución inicial δ_i y cadena de saltos P , independiente de $(X_s : s \leq \tau)$.

Hasta acá nos quedan explicados con suficiente precisión todos los conceptos que utilizamos para el trabajo central desarrollado durante la tesis, que vamos a retomar en el Capítulo 3.

Capítulo 2

La ecuación F-KPP

2.1. El problema planteado por Fisher, Kolmogorov, Petrovskii y Piskunov

En el año 1937, R. A. Fisher publicó un artículo titulado «The way of advance of advantageous genes» [12], en donde proponía un modelo para estudiar el comportamiento de una mutación de un gen en una determinada población. Este artículo fue ampliamente citado a lo largo del siglo XX, y es al día de hoy que muchos trabajos científicos se siguen realizando en torno a los resultados obtenidos, y a los interrogantes que desde ese momento se abrieron. Cabe destacar que en disciplinas muy variadas interesa lo estudiado al respecto, con aplicaciones que van desde lo biológico hasta los campos de la probabilidad y las ecuaciones diferenciales en matemática. En paralelo a Fisher, y en otro polo del mundo de entreguerras, Kolmogorov, Petrovskii y Piskunov llegaban a conclusiones muy similares. Pasemos entonces a estudiar el modelo por ellos planteados y los resultados obtenidos.

Consideremos una población distribuida uniformemente a lo largo de un espacio que modelamos con la recta real, donde en un determinado momento se da una mutación genética ventajosa para la supervivencia de la población en el medio. Siguiendo la teoría desarrollada por Darwin en «El Origen de las Especies» ([4]), podemos esperar que con el paso del tiempo la selección natural dé lugar a la reproducción de ese gen, que se va reproduciendo y expandiendo en la población. Así, la proporción de individuos que cuentan con ese gen va aumentando y desplazándose en el espacio, acorde al rango de alcance que tiene el gen al reproducirse.

Llamemos p a la proporción con la que encontramos este gen ventajoso, y $q = 1 - p$ al gen que previamente ocupaba ese lugar, que asumimos que era el único. A p lo pensamos como una función que depende de la coordenada espacial x y la coordenada temporal t .

Sea m la intensidad con la que se selecciona el nuevo gen, que suponemos independiente de p , y notemos que por cómo lo definimos, $0 \leq p \leq 1$.

Identificamos dos movimientos por los cuales se expande la proporción de este nuevo gen. Por un lado vamos a tener una difusión del mismo, de los lugares donde está más concentrado a donde está menos, y por otro vamos a ver su crecimiento puntual acorde a la proporción presente en cada punto. Supongamos por un momento que modelamos cada proceso por separado, para estudiarlos así y después ver qué sucede cuando ocurren en simultáneo, siendo p_1 y p_2 las funciones que los modelan respectivamente.

Asumimos que al rango de difusión del gen podemos expresarlo como $-k \frac{\partial p_1}{\partial x}$, donde k es el coeficiente de difusión (como el utilizado para estudiar la difusión del calor). Es de destacar que el signo negativo lo tenemos porque la concentración sube de izquierda a derecha sólo si la concentración a izquierda es mayor que a derecha (pues la propagación del gen se da desde los lugares donde está mayormente concentrado). Es decir, en términos analíticos, si a la izquierda de un punto x_0 la concentración p es mayor que a la derecha, tenemos que $\frac{\partial p_1}{\partial x}(x_0, t)$ es negativa, y en ese sentido el signo menos es el que hace que su difusión sea positiva.

El crecimiento puntual lo pensamos proporcional a la cantidad que hay en dicho punto, pero con el limitante de la capacidad máxima que deja crecer el medio.

Veamos que bajo estas hipótesis, p satisface la siguiente ecuación diferencial:

$$\frac{\partial p}{\partial t} = k \frac{\partial^2 p}{\partial x^2} + mp(1 - p). \quad (2.1)$$

1) Difusión del gen

Consideremos $x_0 < x_1 \in \mathbb{R}$, y sea G la función que mide la cantidad de genes ventajosos en la región (x_0, x_1) . G es una función que depende de t , y se la puede expresar como

$$G(t) = \int_{x_0}^{x_1} p_1(x, t) dx.$$

Por lo tanto, por la regla de Leibniz para derivar con respecto a t ,

$$\frac{dG}{dt}(t) = \int_{x_0}^{x_1} \frac{\partial p_1}{\partial t}(x, t) dx. \quad (2.2)$$

Por otro lado, la velocidad de cambio de $G(t)$, dada por $\frac{dG}{dt}(t)$, se mide por la diferencia de la cantidad de genes que entran en x_0 con la cantidad que sale por x_1 , y eso ya asumimos que en un punto x se expresa como $-k \frac{\partial p_1}{\partial x}(x, t)$. Entonces

$$\begin{aligned}
\frac{dG}{dt}(t) &= -k \frac{\partial p_1}{\partial x}(x_0, t) - \left(-k \frac{\partial p_1}{\partial x}(x_1, t) \right) \\
&= k \left(\frac{\partial p_1}{\partial x}(x_1, t) - \frac{\partial p_1}{\partial x}(x_0, t) \right) \\
&= k \cdot \int_{x_0}^{x_1} \frac{\partial^2 p_1}{\partial x^2}(x, t) dx.
\end{aligned} \tag{2.3}$$

(en la última igualdad usamos el Teorema Fundamental del Cálculo para la función $\frac{\partial^2 p_1}{\partial x^2}(x, t)$).

Si igualamos (2.2) y (2.3):

$$\int_{x_0}^{x_1} \frac{\partial p_1}{\partial t}(x, t) dx = k \cdot \int_{x_0}^{x_1} \frac{\partial^2 p_1}{\partial x^2}(x, t) dx.$$

Derivando a ambos lados de la igualdad con respecto a x , nos queda la primera parte de la ecuación deseada.

$$\frac{\partial p_1}{\partial t} = k \frac{\partial^2 p_1}{\partial x^2}$$

2) Crecimiento puntual

El crecimiento puntual responde, como dijimos, a un crecimiento proporcional a la cantidad presente de genes nuevos (mp_2), pero está condicionada por la cantidad que resta por crecer: cuanto menos queda, más difícil se hace su difusión. Como la proporción de p puede llegar a ser 1, la expresión anterior también es proporcional a $1 - p_2$. Así, el crecimiento infinitesimal $\frac{\partial p_2}{\partial t}(x, t)$ queda modelado como

$$\frac{\partial p_2}{\partial t} = mp_2(1 - p_2)$$

Esta ecuación es conocida como la ecuación logística, que es una versión más precisa del modelo de crecimiento exponencial.

Ahora, para juntar ambos movimientos, retomamos la fórmula de Trotter, que la encontramos enunciada y demostrada en [20, p. 117]. La misma sirve para observar qué sucede en el límite cuando ocurren sucesivamente dos procesos en intervalos de tiempo $1/n$.

Teorema 2.1.1. Sean \mathcal{L}_1 , \mathcal{L}_2 y \mathcal{L} generadores de los semigrupos S_1 , S_2 y S respectivamente. Si $\mathcal{L} = \mathcal{L}_1 + \mathcal{L}_2$, entonces para cada f

$$S(t)f = \lim_{n \rightarrow \infty} \left[S_1 \left(\frac{t}{n} \right) S_2 \left(\frac{t}{n} \right) \right]^n f.$$

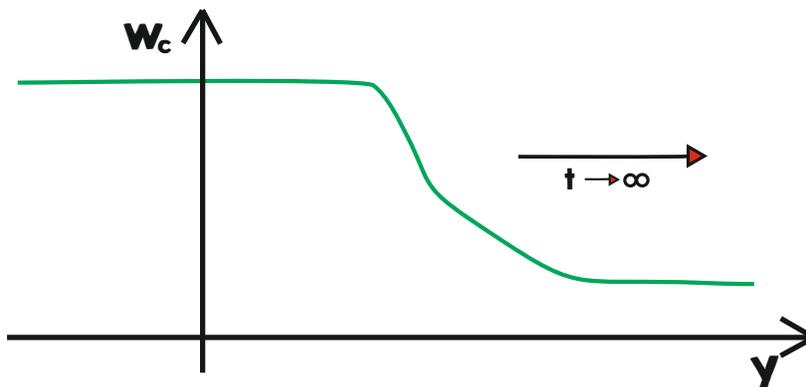


Figura 2.1: Onda viajera con velocidad c

(Para una explicación más completa, leer el capítulo sobre los procesos de Feller en [20, cap. 3])

Aplicando este teorema en nuestro caso, tenemos que si S_1 y S_2 son los espacios de soluciones asociados a p_1 y p_2 respectivamente, la suma de los generadores genera, a su vez, el espacio de soluciones asociado al límite de pequeños intervalos de ambos procesos. Por lo tanto, podemos resumir el problema en

$$\frac{\partial p}{\partial t} = k \frac{\partial^2 p}{\partial x^2} + mp(1 - p). \quad (2.4)$$

2.2. Resolución del problema y ambigüedad en el resultado

Como señalan Brunet y Derrida en [3], la ecuación (2.1) es una ecuación diferencial parcial no lineal y no hay forma de escribir una expresión explícita para $p(x, t)$. Sin embargo, si uno quiere llegar a una solución en el equilibrio, es razonable esperar que se mueva con el tiempo y no que sea una función estática, como se ve en la Figura 2.2. A este tipo de funciones las llamamos «ondas viajeras» (en inglés «traveling wave»).

Analíticamente, las ondas viajeras son funciones asociadas a una velocidad c dada, y que dependen de una sola variable. Llamémoslas w_c :

$$p(x, t) = w_c(x - ct). \quad (2.5)$$

(Si pensamos a w como función de $w(y)$, con el cambio de variable $x - ct$, podemos ver que dado un x fijo, a medida que avanza el tiempo t , el gráfico de la función se desplaza

hacia la derecha, modelando el hecho de que los genes se difunden apareciendo en lugares donde antes no estaban).

Así definida w , escribimos relaciones entre su derivada y las de p .

$$\begin{cases} \frac{\partial p}{\partial t} = -cw'_c \\ \frac{\partial^2 p}{\partial x^2} = w''_c. \end{cases}$$

Y luego reescribimos (2.1) como una ecuación de una sola variable:

$$kw''_c + cw'_c + mw_c(1 - w_c) = 0. \quad (2.6)$$

Para aproximar las soluciones se puede hacer lo siguiente:

1) Plantear

$$\begin{cases} w'_c = z \\ kz' = -cz - mw_c(1 - w_c) \end{cases}$$

2) Resolver el sistema homogéneo

$$\begin{cases} 0 = z \\ 0 = -cz - mw_c(1 - w_c) \end{cases} \quad (2.7)$$

De (2.7) sale que $w \equiv 0$ y $w \equiv 1$ son soluciones en el equilibrio. Para obtener el diagrama de fases se puede linealizar la ecuación, pero ni mediante este método podemos llegar a una fórmula explícita para w .

En su paper [12, p. 356], Fisher busca ondas viajeras directamente planteando

$$\frac{\partial p}{\partial t} = -c \frac{\partial p}{\partial x} \quad (2.8)$$

condición que en nuestro caso se puede deducir de (2.5). A él también le queda para resolver (2.6), y sin encontrar la solución explícita, demuestra que sólo hay solución para velocidades de propagación c que cumplan

$$c \geq 2\sqrt{km}$$

Lo impactante de esta respuesta que contempla infinitas soluciones, es que el problema tiene infinitos equilibrios. En palabras del mismo Fisher [12, p. 359], el comportamiento del sistema sería muy distinto si uno lo fuerza a empezar con un gradiente de propagación bajo (correspondiente a una velocidad de propagación alta) que si se lo hace con un gradiente alto (correspondiente a una velocidad baja). (Fisher llama gradiente a $-\frac{p_x}{p_t}$,

que a partir de (2.8) se deduce que es igual a $1/c$). Comenzando con un c grande la mutación se va a extender con una velocidad mayor que si comenzamos con un c chico, y esa diferencia de velocidades se va a mantener dado que ya desde el inicio tenemos las condiciones invariantes que les pedimos a las ondas viajeras.

En las antípodas de esta conclusión, el modelado físico esperaría que a la larga la velocidad de propagación se ajuste de modo tal que no dependa de las condiciones iniciales, tales como la magnitud del gradiente con las que se comienza.

2.3. Frente de propagación: falla en lo microscópico. El sistema de Brunet-Derrida

Para dar respuesta a esta situación, el mismo Fisher conjetura que el modelo es apropiado para estudiar qué sucede donde la densidad de genes (nuevos y viejos) es alta, pero no así en el frente y en la parte de atrás de la onda viajera, donde la presencia de estos genes es mucho más chica. El comportamiento de un conjunto finito de partículas no podría llevarse tan simplemente al límite, como asumimos que pasa si sólo nos dedicamos a estudiar las soluciones de (2.1).

Una vez planteada esta conjetura, muchos estudios se realizaron para ver qué sucede en esa región, donde lo microscópico parece tener un comportamiento distinto de lo macroscópico. En este trabajo nosotros le prestamos especial atención a una de las propuestas de Brunet y Derrida.

Para estudiar el comportamiento de los genes nuevos en el frente de propagación, los pensamos como partículas, y proponemos un sistema con una cantidad finita de ellas (digamos N) que interactúan entre sí en \mathbb{R} . Al movimiento de la difusión lo vamos a modelar con un paseo al azar, y al crecimiento puntual lo construimos como un proceso de ramificación-selección.

¿A qué llamamos ramificación? Consiste básicamente en una multiplicación de partículas, como si alguna de ellas en un momento determinado tuviese hijos. En función de la posición de la partícula que tiene hijos y de alguna variable aleatoria que queramos utilizar, le asignamos una posición a la(s) nueva(s) partícula(s) que surgen con este procedimiento.

Finalmente, como identificamos una capacidad total del medio que no puede ser superada mediante el crecimiento puntual, tenemos que plasmarla en el modelado de los sistemas de N partículas. Es por eso que hacemos una selección de partículas: cada vez que la cantidad total de ellas aumenta a causa de una ramificación, nos quedamos con un subconjunto de modo tal que la cantidad se mantenga en N . El criterio de esta selección lo hacemos acorde al movimiento que toma la onda viajera. Por ejemplo, si en la ramificación una partícula se duplicó, de las $N + 1$ totales podemos seleccionar las N que se encuentran más a la derecha.

Se publicaron muchos trabajos en el área a lo largo del siglo XX para poder caracterizar el fenómeno que describimos acá, como por ejemplo en [2, 3]. Durrett y Remenik ([8]) demuestran, entre otras cosas, que el proceso visto desde la partícula que está más a la izquierda es ergódico, lo que implica la existencia de una velocidad v_N a la cual viaja la nube de partículas en equilibrio. También demuestran que estas velocidades convergen en forma creciente a la velocidad de la onda viajera minimal. Este proceso es interpretado como un *Principio de selección*: el sistema microscópico tiene una única velocidad para cada N , y estas velocidades “seleccionan” la velocidad minimal cuando $N \rightarrow \infty$. Contrastando con la ecuación límite del caso macroscópico, esto permitiría dar una respuesta al problema descripto. De todas maneras, varios de los modelos propuestos tienen al día de hoy preguntas abiertas cuyas respuestas son de interés en distintas disciplinas.

Capítulo 3

El sistema de Brunet-Derrida en los enteros

A partir del problema planteado por Fisher ([12]), Brunet y Derrida ([2]) primero, y Durrett y Remenik adaptándolo después ([8]), estudiaron los sistemas de partículas de ramificación y selección de los que ya hablamos en el capítulo anterior. En este capítulo nosotros tomamos aquel que estudiaron Durrett y Remenik y lo llevamos a \mathbb{Z} para a partir de ahí analizar su comportamiento.

Consideremos un conjunto de N partículas distribuidas aleatoriamente en \mathbb{Z} según una distribución inicial. Cada partícula va a tener asociados dos procesos de Poisson, que los vamos a pensar como dos relojes que suenan en repetidas ocasiones, según una distribución exponencial. Cuando el primero de los relojes de una partícula suene, la partícula se desplazará a izquierda o a derecha con probabilidad $1/2$. Cuando el segundo suene, la partícula se «ramificará», generando una nueva partícula que ocupará su mismo lugar. En este último caso, para poder mantener la cantidad de partículas en N , nos quedaremos con las N que se encuentren más a la derecha, o lo que es equivalente, eliminaremos aquella que se encuentre más a la izquierda.

Para formalizar la definición del sistema realizamos la construcción gráfica de Harris, que consiste en construir al sistema como una función determinística de una familia de procesos de Poisson. Tomamos como modelo la construcción de sistemas similares en [11, 13, 19].

3.1. Construcción Gráfica

Vamos a construir el sistema con la familia $\{\eta_t, t \geq 0\}$, que nos va a indicar a tiempo t la ubicación de cada una de las partículas, ordenadas de menor a mayor. Es decir, $\eta_t \in \mathcal{S}$, donde

$$\mathcal{S} = \{(x_1, \dots, x_N) \in \mathbb{Z}^N / x_1 \leq \dots \leq x_N\}.$$

Para cada partícula $1 \leq i \leq N$, sean $\omega(i) = (\omega^D(i), \omega^R(i))$ Procesos de Poisson independientes de tasa $\lambda/2$, con $\lambda \in \mathbb{R}_{>0}$: para cada i tenemos dos procesos, y por lo tanto tenemos $2N$ en total, siendo ω^D el que nos indicará los desplazamientos y ω^R las ramificaciones.

También para cada i consideramos luego una sucesión de variables $(A_{i,k})_{k \in \mathbb{N}}$ *iid*, con distribución $Be(1/2)$. Si $A_{i,k} = -1$ el k -ésimo desplazamiento de la partícula i será a la izquierda y si $A_{i,k} = 1$ será a la derecha, eligiendo uno u otro con probabilidad $1/2$.

Sea $\omega = ((\omega^D(i), \omega^R(i)), i \in \{1, \dots, N\})$ la secuencia *iid* de procesos marcados asociados a las N partículas, y $w[0, t]$ la proyección de las marcas en el período de tiempo $[0, t]$.

Con estas herramientas, construimos el proceso $\{\eta_t, t \geq 0\}$ de manera tal que η_t es función de una configuración inicial η_0 y las marcas en $w[0, t]$, $t \geq 0$. Fijemos entonces una distribución inicial $\eta_0 \in \mathcal{S}$ y un tiempo $t > 0$. Los procesos de Poisson observados en un intervalo $[0, t]$ son una variable aleatoria con distribución de Poisson, y por lo tanto la cantidad de marcas es finita casi seguramente. Si K es la cantidad de marcas que tiene ω en $[0, t]$, consideramos $\{b_k, k \leq K\}$ las marcas ordenadas y construimos el proceso inductivamente a partir de ellas.

- A tiempo 0 la configuración es la de η_0 .
- Supongamos η_{b_k} ya definido. Para $t \in [b_k, b_{k+1})$ definimos $\eta_t = \eta_{b_k}$, y para $\eta_{b_{k+1}}$:
 - Si b_{k+1} es una marca de $\omega^D(i)$ sumamos $A_{i,k+1}$ a una determinada partícula de η_{b_k} (que puede ser restar o sumar 1, según cuánto vale $A_{i,k+1}$). Notemos que no podemos sumar $A_{i,k+1}$ directamente a la partícula i , porque $\eta_{b_{k+1}}$ tiene que tener sus coordenadas ordenadas. Este problema surge cuando hay más de una partícula en la misma posición que la i -ésima. Para solucionarlo indicamos que
 - si $A_{i,k+1} = 1$, sea $i_0 = \max\{j : \eta_{b_k}(j) = \eta_{b_k}(i)\}$, y definimos $\eta_{b_{k+1}}(i_0) = \eta_{b_k}(i) + A_{i,k+1}$. Es decir, de todas las partículas que están en el mismo lugar que la i -ésima, sumamos 1 a la que tiene mayor índice.
Para $j \neq i_0$, las partículas quedan en su lugar: $\eta_{b_{k+1}}(j) = \eta_{b_k}(j)$.
 - si $A_{i,k+1} = -1$, sea $i_0 = \min\{j : \eta_{b_k}(j) = \eta_{b_k}(i)\}$, y definimos $\eta_{b_{k+1}}(i_0) = \eta_{b_k}(i) + A_{i,k+1}$. Es decir, ahora restamos 1 a la que tiene menor índice.
Para $j \neq i_0$, las partículas quedan en su lugar: $\eta_{b_{k+1}}(j) = \eta_{b_k}(j)$.
 - Si b_{k+1} es una marca de $\omega^R(i)$, en $\eta_{b_{k+1}}$ hacemos que todas las partículas salten a la posición de la partícula con índice inmediato superior, hasta la i -ésima que hace nacer una nueva. Es decir:

$$\begin{aligned} \eta_{b_{k+1}}(j) &= \eta_{b_k}(j+1) \text{ para } j < i \\ \eta_{b_{k+1}}(j) &= \eta_{b_k}(j) \text{ para } j \geq i. \end{aligned}$$

Notar que $\eta_{b_{k+1}}(i-1) = \eta_{b_{k+1}}(i) = \eta_{b_k}(i)$, que sería la partícula que se ramificó.

Así definido η_t queda siempre en el espacio de estados \mathcal{S} , y por construcción nos queda una cadena de Markov.

La cadena de Markov que vamos a estudiar es una modificación de η_t :

$$\begin{aligned}\Delta_t &= (\Delta_t(1), \dots, \Delta_t(N)) \\ \Delta_t(j) &= \eta_t(j) - \eta_t(1).\end{aligned}$$

Notar que Δ_t indica la ubicación de cada partícula vista desde la que se encuentra más a la izquierda, y que $\Delta_t(1) = 0, \forall t$. No es difícil ver que es una cadena de Markov, basándose en el hecho de que η_t lo es, y de que en cada marca del proceso de Poisson sólo una de las partículas se ramifica o se desplaza, ya sea la primera o alguna de las siguientes.

Observación 3.1.1. Un enfoque equivalente para ver los $2N$ procesos de Poisson que conforman ω es verlos en simultáneo.

Como los $\omega^D(i)$ y los $\omega^R(i)$ son procesos de Poisson, unirlos (es decir, ver las marcas de todos ellos juntos) nos da un proceso de Poisson con la suma de los parámetros, como se puede deducir de la definición. De esta manera nos queda un proceso de Poisson de parámetro $N\lambda$, y si llamamos S_j a los tiempos que espera este proceso, estos tienen distribución $\epsilon(N\lambda)$.

Observación 3.1.2. Para la construcción que dimos la distribución de los tiempos que espera la cadena para dar un salto es siempre la misma. Esto hace que su comportamiento sea análogo al de una cadena a tiempo discreto. En particular, es de destacar que la cadena no tiene explosión, como conclusión de que sus tiempos de espera no son sumables (ver en la Sección 1.6).

3.2. El caso de dos partículas

Tomando el caso $N = 2$ tenemos un ejemplo gráfico, que nos deja a simple vista algunas de las propiedades del sistema, y nos permite hacer algunas cuentas que son complicadas cuando $N > 2$.

Como $\Delta_t(1) = 0$, la única coordenada que nos interesa estudiar es $\Delta_t(2) = \eta_t(2) - \eta_t(1)$. Incluso si llamamos $\alpha_t(i)$ a cada partícula durante todo el proceso, sin reordenarlas en cada paso, tenemos que $\Delta_t(2) = |\alpha_t(1) - \alpha_t(2)|$; a esta variable aleatoria la llamamos Z_t , y vamos a estudiar directamente la cadena de saltos Z_{J_n} . Para eso, calculamos su matriz de transición P .

Empecemos por el caso en que las dos partículas se encuentran juntas, es decir, cuando $Z_{J_n} = 0$. Allí si se da una ramificación Z_{J_n} se mantiene en 0. En caso de un desplazamiento, pasa a 1.

$$\begin{aligned}P(0, 0) &= \frac{1}{2} \\ P(0, 1) &= \frac{1}{2}\end{aligned}$$

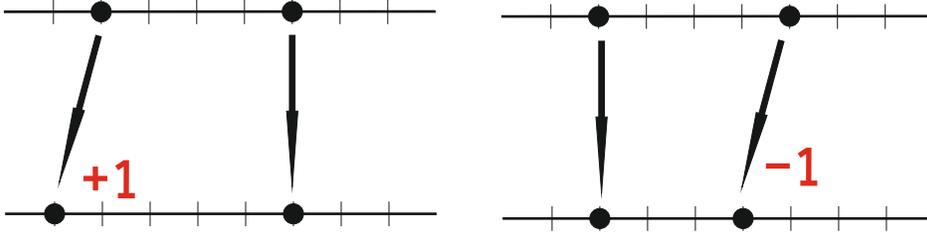


Figura 3.1: Desplazamiento de dos partículas

Por otro lado, en un estado $i \geq 1$, con un desplazamiento Z_{J_n} sólo puede variar en 1 como se puede ver en la Figura 3.1: si la partícula de la izquierda se mueve a la izquierda, o si la de la derecha se mueve a la derecha, Z_{J_n} aumenta en 1, y en los otros dos casos (la de la izquierda a la derecha, o la de la derecha a la izquierda), Z_{J_n} disminuye en 1. Si llamamos

$$G_k^D = \{J_k \text{ es una marca de } \omega^D\}$$

Entonces

$$\begin{aligned} \mathbb{P}(Z_{J_{n+1}} = i + 1, G_{k+1}^D | Z_{J_n} = i) &= \left(\frac{1}{2}\right)^3 + \left(\frac{1}{2}\right)^3 = \frac{1}{4} \\ \mathbb{P}(Z_{J_{n+1}} = i - 1, G_{k+1}^D | Z_{J_n} = i) &= \frac{1}{4} \end{aligned}$$

Si se da una ramificación, recordemos que luego se hace la selección de las dos que se encuentran a la derecha. Por lo tanto si se ramifica la partícula que está más a la izquierda el sistema vuelve al mismo estado, y si se ramifica la de la derecha, automáticamente $Z_{J_{n+1}} = 0$. Ambos eventos se pueden apreciar en la Figura 3.2. Si llamamos

$$G_k^R = \{J_k \text{ es una marca de } \omega^R\},$$

entonces

$$\begin{aligned} \mathbb{P}(Z_{J_{n+1}} = i, G_{k+1}^R | Z_{J_n} = i) &= \frac{1}{2} \frac{1}{2} = \frac{1}{4} \\ \mathbb{P}(Z_{J_{n+1}} = 0, G_{k+1}^R | Z_{J_n} = i) &= \frac{1}{2} \frac{1}{2} = \frac{1}{4} \end{aligned}$$

Es decir, si juntamos, toda la matriz nos queda para cada i :

$$\begin{aligned} P(i, i) &= \frac{1}{4} & P(i, i - 1) &= \frac{1}{4} \\ P(i, 0) &= \frac{1}{4} & P(i, i + 1) &= \frac{1}{4} \end{aligned}$$

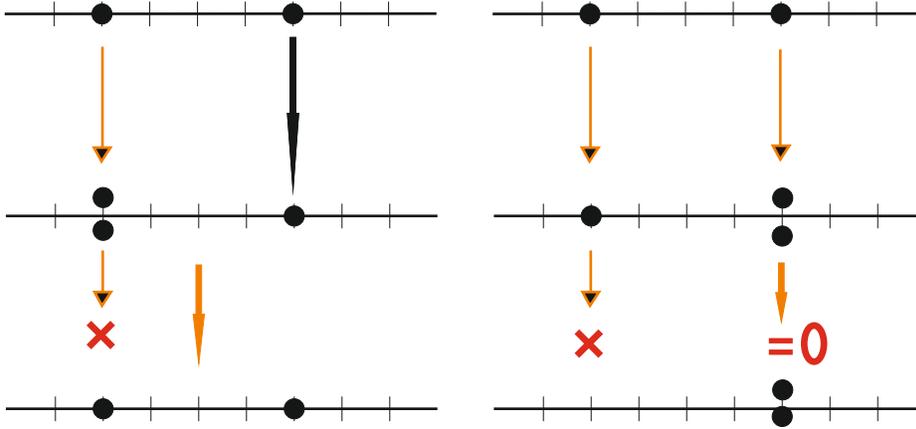


Figura 3.2: Ramificación de dos partículas

Notemos que desde cualquier estado hay probabilidad $1/4$ de volver al estado 0 . Esto nos dice que de alguna manera la variable aleatoria τ_0 que cuenta la cantidad de pasos que tarda el esqueleto en volver al 0 está acotada por una geométrica de parámetro $1/4$.

$$\tau_0 \leq G \quad \text{con } G \sim Ge(1/4)$$

Luego, como $\mathbb{E}(G) = 4 < \infty$, $\mathbb{E}(\tau_0) < \infty$, y eso nos dice que la cadena es recurrente positiva.

Vamos a generalizar y formalizar esta idea en las dos siguientes secciones.

3.3. Irreducibilidad

Volviendo al caso de N partículas, vamos a ver algunas de las propiedades más importantes que podemos encontrar en las cadenas de Markov; todas las definiciones que usemos las fuimos enumerando en el Capítulo 1.

Lema 3.3.1. El esqueleto de la cadena puede pasar de cualquier estado $s_{i_0} \in \mathcal{S}$ al estado $0 = (0, \dots, 0)$ en una cantidad finita pasos; es decir, si $\Delta_{J_i} = (0, \Delta_{J_i}(2), \dots, \Delta_{J_i}(N)) = s_{i_0}$ para algún $i \in \mathbb{N}$, entonces existe un $n \in \mathbb{N}$ tal que

$$\mathbb{P}(\Delta_{J_{i+n}} = 0) > 0.$$

Equivalentemente, por la homogeneidad de la cadena, si P_{ij} es la matriz de transición del esqueleto, entonces para cada i_0 existe un $n \in \mathbb{N}$ tal que

$$P_{i_0 0}^{(n)} > 0.$$

Demostración. Como nuestra cadena es homogénea, podemos asumir sin pérdida de generalidad que empezamos en s_{i_0} , es decir, que $\Delta_0 = s_{i_0}$.

Vamos a construir un algoritmo por el cual la cadena se mueve paso a paso, y en una cantidad finita cada coordenada llega al 0. Recordemos que η como lo definimos en la Sección 3.1 es un vector ordenado, i.e. si $i < j$, $\eta(i) \leq \eta(j)$, y por lo tanto $\Delta(i) \leq \Delta(j)$ también. Empezamos por $\Delta_0(2) \in \mathbb{Z}$, y suponemos que la segunda partícula se desplaza $\Delta_0(2)$ veces a la izquierda. Si se da esta secuencia de pasos, la segunda coordenada llega al 0 ($\Delta_{J_{\Delta_0(2)}} = 0$).

En términos de las marcas de los procesos de Poisson, sería encontrarse con $\Delta_0(2)$ marcas seguidas del tipo $\omega^D(2)$ con $A_{2,k} = -1$. Como el tiempo exponencial esperado por cada marca es el mismo, y como $\mathbb{P}(A_{2,k} = -1) = 1/2$, la probabilidad de que cualquier marca sea $\omega^D(2)$ con $A_{2,k} = -1$ es $\frac{1}{2N}$. Además las marcas entre ellas son independientes, por lo que

$$\begin{aligned} & \mathbb{P}\left(J_i \text{ sea una marca de } \omega^D(2) \quad \forall 1 \leq i \leq \Delta_0(2)\right) \\ &= \prod_{i=1}^{\Delta_0(2)} \mathbb{P}\left(J_i \text{ sea una marca de } \omega^R(N)\right) \\ &= \prod_{i=1}^{\Delta_0(2)} \frac{1}{2N} = \left(\frac{1}{2N}\right)^{\Delta_0(2)} > 0. \end{aligned} \tag{3.1}$$

Si repetimos este procedimiento para cada $3 \leq j \leq N$, cada $\Delta(j)$ en $\Delta_0(j)$ pasos llega al 0. Entonces toda la cadena en una cantidad de pasos igual a $\sum_{j=2}^N \Delta_0(j)$ llega al estado $0 = (0, \dots, 0)$. \square

Teorema 3.3.2. La cadena Δ_t es irreducible.

Demostración. Citando la construcción de Norris, en [22, p. 111] encontramos que el teorema se puede demostrar directamente sobre el esqueleto.

Sean dos estados s_{i_0} y s_{i_1} en \mathcal{S} . Por el Lema 3.3, el esqueleto llega al 0 en finitos pasos.

Luego, también en finitos pasos la cadena puede ir del 0 a cualquier estado: existe un $n(i_1) \in \mathbb{N}$ tal que $P_{0i_1}^{(n)} > 0$. La demostración es totalmente análoga al Lema 3.3, tomando desplazamientos hacia la derecha ($\omega^D(j)$ con $A_{i,k} = 1$) en cada salto. De esta manera se llega de s_{i_0} a s_{i_1} , y por lo tanto la cadena es irreducible. \square

3.4. Existencia y unicidad de medida invariante

Teorema 3.4.1. a) Para la cadena de Markov Δ_t , el estado $0 := (0, \dots, 0)$ cumple que $\mathbb{E}(\tau_0) < \infty$, donde $\tau_0 := \inf \{t \geq 0 : \Delta_t = 0\}$.

b) La cadena Δ_t tiene una única medida estacionaria π .

Demostración. a) Si comenzamos mirando la cadena en el esqueleto Δ_{J_n} donde J_n son las marcas ordenadas del proceso ω , podemos notar que en N pasos Δ_{J_n} tiene probabilidad

positiva de llegar al estado 0, independientemente del estado inicial. Esto sucede si, por ejemplo, se dan N ramificaciones seguidas sobre la última partícula. Lo que vamos a demostrar es que la cantidad de pasos que hay que esperar para que suceda esto tiene esperanza finita, y por lo tanto, τ_0 también. Separamos dicha demostración en varios pasos, de modo tal que podamos definir las variables aleatorias necesarias para acotar τ_0 , y así pasar del esqueleto a la cadena en tiempo continuo.

Paso I: el esqueleto puede llegar al estado 0 en un bloque de N pasos consecutivos.

Sea Δ_0 un estado inicial arbitrario. Supongamos que en el proceso ω encontramos N marcas consecutivas correspondientes a $\omega^R(N)$: esto indica que en N pasos seguidos la última partícula se ramifica. En cada uno de estos pasos, el proceso elimina a la partícula que se encuentra más a la izquierda, y en su lugar agrega una nueva partícula en la posición de la última, es decir, en $\eta_0(N)$. Al repetir N veces, nos quedan las N partículas acumuladas en ese último lugar, y por lo tanto $\Delta_{J_N} = 0$.

La probabilidad de este evento es positiva:

$$\begin{aligned} \mathbb{P}(\Delta_{J_N} = 0) &\geq \mathbb{P}\left(J_i \text{ sea una marca de } \omega^R(N) \quad \forall 1 \leq i \leq N\right) \\ &= \prod_{i=1}^N \mathbb{P}\left(J_i \text{ sea una marca de } \omega^R(N)\right) \\ &= \prod_{i=1}^N \frac{1}{2N} = \left(\frac{1}{2N}\right)^N > 0 \end{aligned} \tag{3.2}$$

Paso II: contamos cuántos bloques de N pasos consecutivos pasan hasta que el esqueleto realiza esta acción

Consideremos los eventos

$$\mathcal{G}_n = \{\text{todas las marcas de } \omega \text{ entre la } ((n-1)N+1)\text{-ésima y la } nN\text{-ésima son marcas de } \omega^R(N)\}. \tag{3.3}$$

Es decir, \mathcal{G}_n nos indica si hubo N ramificaciones seguidas entre las marcas ahí indicadas. Notar que miramos cada N marcas, dividiéndolas así en bloques disjuntos de tamaño N . Gracias a esto afirmamos que la familia de eventos $(\mathcal{G}_n)_{n \in \mathbb{N}}$ son todos independientes entre sí. Además, por lo que demostramos en el Paso I, $\mathbb{P}(\mathcal{G}_n) = \left(\frac{1}{2N}\right)^N \quad \forall n \in \mathbb{N}$.

Definimos ahora la variable aleatoria G como suma de indicadoras de conjuntos disjuntos:

$$G = \sum_{j=1}^{\infty} j \mathbf{1}_{\{\bigcap_{i=1}^{j-1} (\mathcal{G}_i)^c \cap \mathcal{G}_j\}} \tag{3.4}$$

G es la variable aleatoria que nos permitirá acotar la cantidad de pasos que demora el esqueleto en llegar al 0. Veamos primero que G es una variable aleatoria con distribución geométrica de parámetro $p = (1/2N)^N$.

Lema 3.4.2. $G \sim Ge(p)$

Demostración. Sea $v \in \Omega$

$$\{G = k\} \iff \bigcap_{i=1}^{k-1} (\mathcal{G}_i)^c \cap \mathcal{G}_k. \quad (3.5)$$

Por lo tanto, y utilizando la independencia y la probabilidad ya calculada de los \mathcal{G}_n ,

$$\begin{aligned} \mathbb{P}(G = k) &= \mathbb{P}\left(\bigcap_{i=1}^{k-1} (\mathcal{G}_i)^c \cap \mathcal{G}_k\right) \\ &= \left(\prod_{i=1}^{k-1} \mathbb{P}(\mathcal{G}_i)^c\right) \mathbb{P}(\mathcal{G}_k) \\ &= \left(1 - \left(\frac{1}{2N}\right)^N\right)^{k-1} \cdot \left(\frac{1}{2N}\right)^N \\ &= (1 - p)^{k-1} \cdot p. \end{aligned} \quad (3.6)$$

$\implies G \sim Ge(p)$

□

Caracterizada G , notemos que $G = k$ sii \mathcal{G}_k es el primero de los eventos $(\mathcal{G}_n)_{n \in \mathbb{N}}$ que sucedió, es decir, si recién entre las marcas $((k-1)N+1)$ -ésima y la kN -ésima encontramos N ramificaciones seguidas. Por lo tanto si $G = k$, el esqueleto se demoró a lo sumo Nk pasos en completar la acción descripta al principio (decimos «a lo sumo» por la salvedad que hacemos de separar las marcas de a N para utilizar la independencia, y por los otros caminos por los que la cadena puede llegar al 0).

Paso III: contamos el tiempo que demora el proceso, en su versión continua, en realizar esta misma acción. A este tiempo lo llamamos Z .

Recordemos que en la Observación 3.1.1 definimos los \bar{S}_j como cada uno de los tiempos esperados por el proceso ω . Como en el Paso II dividimos las marcas de a N , consideremos los tiempos que quedan de sumar de a N los \bar{S}_j : más precisamente, sean $\Gamma_i = \sum_{j=1}^N \bar{S}_{j+(i-1)N}$. Por el Teorema 5.1.2 del apéndice, $\Gamma_i \sim \Gamma(N, \lambda) \quad \forall i$.

$$\text{Sea } Z = \sum_{i=1}^G \Gamma_i.$$

Z efectivamente acota el tiempo que demora el proceso en realizar las N ramificaciones seguidas, porque suma de a N exponenciales hasta que sucede dicho evento. Damos más precisiones al respecto en el siguiente paso.

Paso IV: $\tau_0 \leq Z$ casi seguramente

Para eso notemos primero que

$$\bigcup_{k=1}^{\infty} \{G = k\} = \Omega.$$

Luego

$$\begin{aligned}
\{\tau_0 \leq Z\} &= \{\tau_0 \leq Z\} \cap \bigcup_{k=1}^{\infty} \{G = k\} \\
&= \bigcup_{k=1}^{\infty} \{\tau_0 \leq Z\} \cap \{G = k\} \\
&= \bigcup_{k=1}^{\infty} \left\{ \tau_0 \leq \sum_{i=1}^G \Gamma_i \right\} \cap \{G = k\} \\
&= \bigcup_{k=1}^{\infty} \left\{ \tau_0 \leq \sum_{i=1}^k \Gamma_i \right\} \cap \{G = k\} \\
&\stackrel{*}{=} \bigcup_{k=1}^{\infty} \{G = k\} \\
&= \Omega.
\end{aligned} \tag{3.7}$$

Veamos que $*$ es válido a partir del siguiente lema:

Lema 3.4.3. $\{G = k\} \subseteq \left\{ \tau_0 \leq \sum_{i=1}^k \Gamma_i \right\}$.

Demostración del lema. Llamamos $\bar{\tau}_0 = \min \{n \geq 1 : \Delta_{J_n} = 0\}$, y con él caracterizamos a τ_0 :

$$\tau_0 = \sum_{i=1}^{\bar{\tau}_0} \bar{S}_i$$

Supongamos que $\{G = k\}$. Por lo que comentamos al final del paso II, sabemos que hay N marcas seguidas del tipo $\omega^R(N)$ antes del paso kN , y por cómo elegimos esta secuencia de eventos, eso implica que $\bar{\tau}_0 \leq kN$.

Luego:

$$\begin{aligned}
\tau_0 &= \sum_{i=1}^{\bar{\tau}_0} \bar{S}_i \leq \sum_{i=1}^{kN} \bar{S}_i \\
&= \sum_{i=1}^k \sum_{j=1}^N \bar{S}_{j+(i-1)N} \\
&= \sum_{i=1}^k \Gamma_i
\end{aligned} \tag{3.8}$$

$$\implies \{G = k\} \subseteq \left\{ \tau_0 \leq \sum_{i=1}^k \Gamma_i \right\}$$

□

Paso V: $\mathbb{E}(\tau_0) < \infty$.

En el paso anterior demostramos que $\tau_0 < Z$, y por lo tanto, $\mathbb{E}(\tau_0) < \mathbb{E}(Z)$. Pero recordemos que $Z = \sum_{i=1}^G \Gamma_i$. Por la Proposición 5.1.4 del apéndice $\mathbb{E}(Z) < \infty$, y eso demuestra que $\mathbb{E}(\tau_0) < \infty$ también.

b) Como ya enunciamos en la Proposición 1.6.9, cuando una cadena de Markov irreducible tiene un estado recurrente positivo, la cadena tiene una única medida invariante. \square

Demostración Alternativa de a). Podemos pensar una demostración alternativa sin pasar por el esqueleto. Si consideramos el proceso que evoluciona a partir de las variables aleatorias exponenciales, podemos demostrar que en un tiempo finito, digamos 1, con la misma idea el proceso puede llegar al estado 0. Sea A_i , $1 \leq i \leq N$, el evento que indica que para \bar{S}_i , la marca que corresponde es $\omega^R(N)$. Es decir, ramificamos la partícula que se encuentra más a la derecha hasta que, gracias a la selección que hacemos salto a salto, se superponen todas. Entonces, nos interesa calcular

$$\mathbb{P}(A_1, \dots, A_N, J_N \leq 1), \quad (3.9)$$

y ver que la misma es positiva. Como los A_j son independientes de los saltos J_i , y además son independientes entre sí, tenemos que la ecuación (3.9) es igual a

$$\mathbb{P}(A_1) \dots \mathbb{P}(A_N) \mathbb{P}(J_N \leq 1). \quad (3.10)$$

Como explicamos en la primera demostración, $\mathbb{P}(A_i) = (\frac{1}{2N})^N$, mientras que para calcular los tiempos de saltos tenemos de nuevo dos formas: por un lado $J_N = \sum_{i=1}^N \bar{S}_i$, es decir, $J_N \sim \Gamma(N, \lambda)$. Por otro lado podemos pensarlo como un proceso de Poisson de parámetro $N\lambda$, donde nosotros queremos tener N éxitos en un tiempo $t = 1$. Utilizando esta segunda idea, si $w[0, t]$ cuenta las marcas de todo el proceso, $w[0, 1] \sim P(\lambda)$, y entonces

$$\mathbb{P}(J_N \leq 1) \geq \mathbb{P}(Z_1 = N) = e^{N-\lambda} \cdot \frac{(N\lambda)^N}{N!}.$$

Luego,

$$\begin{aligned} \mathbb{P}(\Delta_t = 0 \text{ para algún } t \leq 1) &\geq \mathbb{P}(A_1, \dots, A_N, J_N \leq 1) \\ &= \mathbb{P}(A_1) \dots \mathbb{P}(A_N) \cdot \mathbb{P}(J_N \leq 1) \geq \left(\frac{1}{2N}\right)^N \cdot e^{N-\lambda} \cdot \frac{(N\lambda)^N}{N!} > 0. \end{aligned}$$

Llamamos F_n al evento en que $\Delta_t = 0$ para algún $n < t \leq n + 1$. Como la cadena es homogénea y para el razonamiento anterior no usamos cuál era el estado inicial, los F_n son todos independientes y con la misma probabilidad p de que ocurran, que es mayor a $p' = \left(\frac{1}{2N}\right)^N \cdot e^{N-\lambda} \cdot \frac{(N\lambda)^N}{N!}$. Definimos $F = \sum_{n \in \mathbb{N}} \mathbf{1}_{\{F_n\}}$.

Con razonamientos equivalentes a los utilizados en la primera demostración, F es una geométrica de parámetro p y ahora $\tau_0 \leq F \Rightarrow \mathbb{E}(\tau_0) \leq 1/p' < \infty$. \square

3.5. Velocidad Asintótica

En esta sección vamos a demostrar que la primera partícula de la cadena alcanza una velocidad asintótica en el siguiente teorema:

Teorema 3.5.1. Existe un $a_N \in \mathbb{R}$ tal que

$$\frac{\eta_t(1)}{t} \xrightarrow[t \rightarrow \infty]{} a_N.$$

En [8] se demuestra un resultado análogo, pero con la utilización de resultados distintos a los que usamos acá. En dicho paper, se cita una versión del Teorema Ergódico Subaditivo, que Durrett ya había utilizado y demostrado en [7]. La demostración cierra el problema, pero utiliza argumentos que exceden al caso particular que estamos tratando. Con el objetivo de tomar un camino acorde a este trabajo y de proponer nuevos métodos de demostración, tomamos caminos que se basan directamente sobre el comportamiento de la cadena.

Nuestro objetivo es describir el comportamiento de $\eta_t(1)$ en función de nuestra cadena Δ_t , para así poder estudiar su velocidad y ver que en el infinito converge. Vamos a ir reduciendo el problema, siendo el primer paso estudiarlo en su esqueleto: recordemos que J_n es el tiempo correspondiente al n -ésimo salto de la cadena. Más adelante vamos a necesitar una segunda reducción.

La dinámica particular que tiene la cadena (donde renombramos siempre que se reordenan las partículas, y tenemos el hecho de eliminar en determinados pasos a la primera partícula) nos hace detenernos en sus detalles para poder hacer dicha descripción. Los saltos de la partícula se dan en tres casos distintos:

Posibles casos

1. cuando hay ramificación, dado que en realidad la primera partícula se elimina, y la «nueva» primera pasa a ser la que era la segunda. En el caso en que se ramifique la primera partícula no hay ningún salto.
2. cuando la primera partícula se desplaza a la izquierda. Si hay más de una en la primera posición, cualquiera de ellas que se mueva a la izquierda hace que «la primera» se desplace.
3. cuando la primera se mueve a la derecha. Esta va a importar cuando haya sólo una en la primera posición, porque sino otra pasa a ser la primera y no hay desplazamiento real de $\eta_t(1)$.

Descripción del desplazamiento en cada uno de ellos

1. Empecemos definiendo la sucesión $(\tau_k)_{k \in \mathbb{N}_0}$

$$\begin{aligned} \tau_0 &= 0 \\ \tau_{k+1} &= \text{mín} \{t \geq \tau_k / \text{la partícula } i \text{ se ramifica a tiempo } t, \text{ con } 2 \leq i \leq N-1\} \end{aligned} \quad (3.11)$$

Notemos que τ_1 tiene una distribución exponencial, con parámetro $\frac{(N-1)\lambda}{2}$. Sucesivamente, $\tau_{k+1} - \tau_k$ también, y son independientes entre sí. Por eso $\tau_n = \sum_{k=1}^n \tau_k - \tau_{k-1}$ es una suma de exponenciales, y por el Teorema 5.1.2

$$\tau_n \sim \Gamma\left(n, \frac{(N-1)\lambda}{2}\right).$$

Esta sucesión sirve tanto para probar la convergencia del primer término, como para definir de alguna manera el resto de los movimientos.

Con lo anterior, definimos $Y_k = \Delta_{\tau_k}$, y como dijimos cuando describimos el punto 1, sabemos que si hay ramificación, la primera partícula pasa a la posición de la segunda. Es decir, en un paso n tenemos que el efecto que se produce a partir de este primer caso se describe como

$$\sum_{k=1}^n \Delta_{\tau_k}(2) \quad (3.12)$$

2. Definamos $A_n = \#$ de desplazamientos de la primera partícula hacia la izquierda antes de τ_n . Como sucede que cada vez que tenemos un desplazamiento a la izquierda, la posición se disminuye sólo en 1, restando A_n describimos los efectos de este movimiento hasta el tiempo τ_n .

3. Análogamente, definimos $B_n = \#$ de desplazamientos de la primera partícula hacia la derecha antes de τ_n . A diferencia del caso anterior, acá tenemos que sumar B_n .

Tomando como referencia (3.12), y los puntos 2 y 3 nos queda descripto el movimiento de la primera partícula en estos tiempos de parada:

$$\eta_{\tau_n}(1) = \sum_{k=1}^n \Delta_{\tau_k}(2) - A_n + B_n \quad (3.13)$$

Y así es que hacemos la nueva reducción, concentrándonos en demostrar la siguiente proposición.

Proposición 3.5.2. La sucesión

$$\frac{\eta_{\tau_n}(1)}{n} = \frac{1}{n} \sum_{k=1}^n \Delta_{\tau_k}(2) - \frac{A_n}{n} + \frac{B_n}{n} \quad (3.14)$$

converge casi seguramente a un número determinístico, que llamamos \tilde{a}_N .

Demostración. **Probamos la convergencia término a término.**

1. *Primer término.* Como queremos utilizar el Teorema 1.5.2, el Teorema Ergódico que enunciamos en la Sección 1.5, utilizamos el siguiente resultado.

Proposición 3.5.3. El proceso estocástico $Y_k = \Delta_{\tau_k}$ cumple que

- a) es una cadena de Markov.
- b) es recurrente positiva.

Demostración. a) Sean $y_1, \dots, y_{k+1} \in \mathcal{S}$.

$$\begin{aligned}
 & \mathbb{P}(Y_{k+1} = y_{k+1} | Y_0 = y_0, \dots, Y_k = y_k) = \\
 & = \int_{t_0 < \dots < t_k} \mathbb{P}(Y_{k+1} = y_{k+1} | Y_0 = y_0, \dots, \tau_1 = t_1, \dots, \tau_k = t_k) \cdot d\mathbb{P}(\tau_1 = t_1, \dots) \\
 & = \int_{t_0 < \dots < t_k} \mathbb{P}(X_{t_k + \tilde{\tau}} = y_{k+1} | X_0 = y_0, \dots, \tau_1 = t_1, \dots, \tau_k = t_k) \cdot d\mathbb{P}(\tau_1 = t_1, \dots)
 \end{aligned} \tag{3.15}$$

Notemos que $\tilde{\tau} = \tau_{k+1} - \tau_k$ que en distribución es igual a τ_1 . Por otro lado, como el esqueleto es una cadena de Markov, y dado que todos los eventos por los que condicionamos son \mathcal{F}_k medibles por la propiedad fuerte de Markov nos queda:

$$\begin{aligned}
 & = \int_{t_0 < \dots < t_k} \mathbb{P}(X_{t_k + \tilde{\tau}} = y_{k+1} | X_{t_k} = y_k) \cdot d\mathbb{P}(\tau_1 = t_1, \dots, \tau_k = t_k) \\
 & = \int_{t_0 < \dots < t_k} \mathbb{P}(X_{\tau_1} = y_{k+1} | X_0 = y_k) \cdot d\mathbb{P}(\tau_1 = t_1, \dots, \tau_k = t_k) \\
 & = \mathbb{P}(X_{\tau_1} = y_{k+1} | X_0 = y_k) \cdot \int_{t_0 < \dots < t_k} d\mathbb{P}(\tau_1 = t_1, \dots, \tau_k = t_k) \\
 & = \mathbb{P}(X_{\tau_1} = y_{k+1} | X_0 = y_k) = \mathbb{P}(Y_{k+1} = y_{k+1} | Y_k = y_k).
 \end{aligned} \tag{3.16}$$

- b) Como hicimos en el Teorema 3.4.1, y de hecho con una demostración análoga, veamos que el tiempo de retorno de Y_k al $0 = (0, \dots, 0)$ tiene esperanza finita. Sea $\tau_0 := \inf \{k \geq 0 : \Delta_{\tau_k} = 0\}$. Si miramos N ramificaciones, y estas se dan sobre la última partícula, la cadena vuelve al 0. En este caso, como cada paso es una ramificación, la probabilidad de que en el paso N se hayan dado N ramificaciones de la última partícula es $(1/N)^N$.

Sea $G \sim Ge((1/N)^N)$. Partimos los números naturales en bloques de N pasos, y miramos en cada bloque si la geométrica da un éxito: de esta manera, $N \cdot G$ indica en qué paso se dio el éxito de la geométrica, y el tiempo de retorno al 0 es menor que esa cantidad de pasos. Como $\mathbb{E}(NG) = N \cdot N^N < \infty$, nos queda que $\mathbb{E}(\tau_0) < \infty$.

□

Corolario 3.5.4. La cadena Y_k tiene una medida invariante, que la llamamos $\hat{\mu}$.

Con el Corolario 3.5.4 tenemos las hipótesis que pide el Teorema Ergódico para la cadena. Por eso, sólo resta que la función elegida sea integrable con respecto a la invariante. Pero, ¿quién es f ?

Nosotros tenemos la suma de los $\Delta_{\tau_k}(2)$, y queremos que $\Delta_{\tau_k}(2) = Y_k(2) = f(Y_k)$. Entonces, f es simplemente la proyección en la segunda coordenada:

$$f : \mathbb{Z}^d \rightarrow \mathbb{R}$$

$$f(x) = x_2.$$

Lema 3.5.5. La función f así definida es integrable respecto a la medida invariante $\hat{\mu}$; es decir,

$$\sum_{i \in I} f(s_i) \hat{\mu}(s_i) < \infty.$$

Demostración. En [18, p.13] se da la siguiente caracterización para la medida invariante $\hat{\mu}$, que es una generalización de la que nosotros dimos en 1.4.6. Sea $z \in \mathcal{S}$ fijo y sea $T_z = \min \{n \geq 1 : Y_n = z\}$. Si $Y_0 = z$, a T_z se lo llama tiempo de retorno. Sea

$$\tilde{\mu}(s_i) = \mathbb{E}_z (\# \text{ visitas a } s_i \text{ antes de volver a } z) = \mathbb{E}_z \left(\sum_{n \geq 0} \mathbf{1}_{\{Y_n = s_i, T_z > n\}} \right).$$

Entonces $\hat{\mu}(s_i) = \frac{\tilde{\mu}(s_i)}{\mathbb{E}_z(T_z)}$. Si bien se cambia la notación que veníamos usando, los elementos son iguales a los definidos en la Sección 1.4: $\tilde{\mu}(s_i) = \gamma_i^z$ y $\mathbb{E}_z(T_z) = m_z$. De hecho, decimos que es una generalización de la caracterización que habíamos dado, porque en el caso particular en que s_i es el estado z , como $\gamma_z^z = 1$, nos queda $\hat{\mu}(s_z) = \frac{1}{m_z}$ como antes.

Elegimos $z = (0, \dots, 0)$, por ser un estado al que la cadena puede retornar en N pasos desde cualquier otro estado, como ya vimos en la Sección 3.4. Con esta caracterización tenemos que

$$\sum_{i \in I} f(s_i) \hat{\mu}(s_i) = \sum_{i \in I} \frac{f(s_i) \tilde{\mu}(s_i)}{\mathbb{E}_0(T_0)}. \quad (3.17)$$

La cadena es recurrente positiva, y por lo tanto su esqueleto también lo es. Eso nos permite afirmar, entonces, que $\mathbb{E}_0(T_0) < \infty$. Además $T_0 \geq 1$, por lo que $\mathbb{E}_0(T_0) > 0$. Continuamos sin el denominador, y al numerador lo desarrollamos:

$$\begin{aligned}
\sum_{i \in I} f(s_i) \tilde{\mu}(s_i) &= \sum_{i \in I} f(s_i) \mathbb{E}_0 \left(\sum_{n \geq 0} \mathbf{1}_{\{Y_n = s_i, T_0 > n\}} \right) = \\
&= \mathbb{E}_0 \left(\sum_{i \in I} f(s_i) \sum_{n \geq 0} \mathbf{1}_{\{Y_n = s_i, T_0 > n\}} \right).
\end{aligned} \tag{3.18}$$

En cada i -ésimo término de la sumatoria aparece $f(s_i)$ multiplicado por la cantidad de veces que la cadena pasa por ese estado s_i antes de volver al estado 0. Eso mismo se puede expresar como evaluar f en la cadena y sumarlo cada vez que se repite el estado s_i . Pero como además sumamos en todos los estados posibles, terminamos teniendo f sobre la cadena, hasta que retornamos al estado 0. Esto se puede ver reacomodando un poco los elementos de ambas sumas.

$$\begin{aligned}
&= \mathbb{E}_0 \left(\sum_{n \geq 0} \sum_{i \in I} f(Y_n) \mathbf{1}_{\{Y_n = s_i\}} \mathbf{1}_{\{T_0 > n\}} \right) = \mathbb{E}_0 \left(\sum_{n \geq 0} f(Y_n) \mathbf{1}_{\{T_0 > n\}} \sum_{i \in I} \mathbf{1}_{\{Y_n = s_i\}} \right) \\
&= \mathbb{E}_0 \left(\sum_{n \geq 0} f(Y_n) \mathbf{1}_{\{T_0 > n\}} \right) = \mathbb{E}_0 \left(\sum_{n=0}^{T_0} f(Y_n) \right).
\end{aligned} \tag{3.19}$$

Por un lado tenemos que $f(Y_n) = Y_n(2) = \Delta_{\tau_n}(2) \leq \Delta_{\tau_n}(N)$. Por otro, como la cadena empieza en el 0, sabemos que en el paso correspondiente al tiempo τ_n (llamémoslo $k^{(n)}$), los extremos de la cadena no pudieron separarse más que $k^{(n)}$ (puesto que la única manera de que aumente $\Delta(N)$ es cuando hay un desplazamiento del extremo izquierdo a la izquierda, o del derecho a la derecha). Además la sucesión de los $(k^{(n)})_{n \in \mathbb{N}}$ es estrictamente creciente, y $T_0 \leq k^{(T_0)}$. Entonces:

$$\leq \mathbb{E}_0 \left(\sum_{n=0}^{T_0} k^{(n)} \right) \leq \mathbb{E}_0 \left(T_0 k^{(T_0)} \right) \leq \mathbb{E}_0 \left(\left(k^{(T_0)} \right)^2 \right). \tag{3.20}$$

Y esta variable aleatoria tiene esperanza finita, como vemos en el siguiente lema.

Lema 3.5.6. Existe G una variable aleatoria con $G \sim Ge((1/2N)^N)$, que cumple que $k^{(T_0)} \leq N \cdot G$.

Demostración. $k^{(T_0)}$ cuenta cuántos pasos tarda en volver la cadena al 0, y recordemos que desde cualquier estado en N pasos la cadena tiene probabilidad mayor a $(1/2N)^N$ de volver al estado 0.

Retomando la demostración del Teorema 3.4.1, definimos G análogamente a como está definido en el paso II, ahora definiendo los \mathcal{G}_n para la cadena Δ_{τ_n} . Como explicamos al final del paso II, y como reproducimos en el lema 3.4.3, si $G = k$ entonces $k^{(T_0)} \leq Nk$ con probabilidad 1.

□

Con este lema nos queda, entonces, siguiendo desde (3.20), que

$$\mathbb{E}_0 \left(\sum_{n=0}^{T_0} k^{(n)} \right) \leq \mathbb{E}_0 (N^2 G^2) = N^2 \cdot \mathbb{E}_0 (G^2) < \infty. \quad (3.21)$$

□

Como la f es integrable respecto de la medida invariante, afirmamos, mediante el teorema ergódico, que el primer término converge:

$$\frac{1}{n} \sum_{k=1}^n \Delta_{\tau_k}(2) \xrightarrow[t \rightarrow \infty]{} \sum_{i \in I} f(s_i) \hat{\mu}(s_i) \quad c.s. \quad (3.22)$$

□

2. *Segundo término.* La descripción del movimiento de la cadena para esta parte se torna complicada. Cuando describimos el caso hicimos la siguiente salvedad: saber si en un paso de la cadena la primera partícula se desplaza a la izquierda, depende de hechos tales como cuántas partículas hay en la primera posición. En caso de que haya más de una, cualquiera que se desplace hace que la «primera» se desplace, y esto pasa porque nosotros renombramos a las partículas cuando hay alguna alteración en su orden. Entonces, no es cuestión de mirar sólo cuándo es que una determinada partícula se desplaza, sino que siempre dependemos de la configuración de la totalidad de ellas.

Sin embargo, tenemos una manera más práctica de describir el movimiento. Recordemos que

$$\begin{aligned} \Delta_t &= (\Delta_t(1), \dots, \Delta_t(N)) = (\eta_t(1) - \eta_t(1), \dots, \eta_t(N) - \eta_t(1)) \\ &= (0, \eta_t(2) - \eta_t(1), \dots, \eta_t(N) - \eta_t(1)). \end{aligned}$$

Si la primera partícula se desplaza a la izquierda en el paso n , tenemos que $\eta_{J_{n+1}}(1) = \eta_{J_n}(1) - 1$, por lo que se suma 1 en todas las coordenadas de $\Delta_{J_{n+1}}$ exceptuando la primera. De esta manera describimos siempre que se da este movimiento, incluso si hay más de una en la primera posición: en ese caso tenemos varios 0 en las primeras

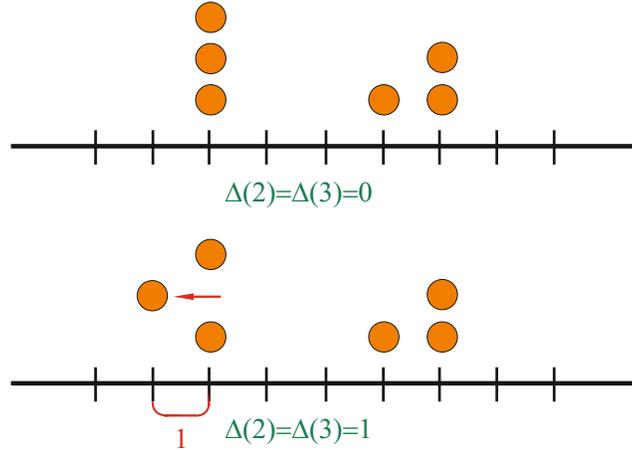


Figura 3.3: Desplazamiento de la primera partícula hacia la izquierda

coordenadas de Δ_t , pero cuando alguna de ellas se va a la izquierda, todas estas coordenadas pasan a ser un 1, como se ve en la Figura 2.

Por esto, entonces, sabemos que A_n suma de a uno cada vez que se ve esa transición antes de τ_n :

$$(0, x_2, \dots, x_N) \longrightarrow (0, x_2 + 1, \dots, x_N + 1)$$

donde $x_2, \dots, x_N \in \mathbb{Z}$, y $x_2 \leq \dots \leq x_N$. Nos definimos los siguientes elementos:

- a) Sea $x = (0, x_2, \dots, x_N) \in \mathcal{S}$. Dado que en la primera coordenada los elementos de \mathcal{S} tienen un 0, llamamos

$$x + 1 = (0, x_2 + 1, \dots, x_N + 1).$$

- b) Definimos la nueva cadena

$$Z_n = (\Delta_{J_n}, \Delta_{J_{n+1}}),$$

que nos sirve para contar cuántas veces se da esta transición

- c) Sea $A = \{(x, x + 1) \mid x \in \mathcal{S}\} \subset \mathcal{S} \times \mathcal{S}$.

Con estas definiciones, nos queda que $A_n = \sum_{j=0}^{k^{(n)}-1} \mathbf{1}_{\{Z_j \in A\}}$. Para seguir trabajando con estos elementos citamos la siguiente Proposición, que se puede encontrar demostrada en [14].

Proposición 3.5.7. El proceso estocástico $Z_n = (\Delta_{J_n}, \Delta_{J_{n+1}})$ cumple que

a) es una cadena de Markov.

b) tiene una medida invariante $\tilde{\mu}(x, y) = \mu(x) \cdot P(x, y)$, donde μ y P son respectivamente la medida invariante y la matriz de transición de la cadena Δ_{J_n} .

Si bien la convergencia de $\frac{A_n}{n}$ no la podemos ver directamente, la hacemos en varios pasos.

Observación 3.5.8. Notemos que la función $g : \mathcal{S} \times \mathcal{S} \rightarrow \mathbb{R}$ que hace $g(\omega) = \mathbf{1}_{\{\omega \in A\}}$ es una función acotada, y por lo tanto integrable con respecto a cualquier medida de probabilidad, en particular la medida $\tilde{\mu}$. Entonces podemos usar el Teorema Ergódico sobre $g(Z_j)$: por el Teorema 1.5.2 la sucesión $\left(\frac{1}{n} \sum_{j=1}^n \mathbf{1}_{\{Z_j \in A\}}\right)_{n \in \mathbb{N}}$ converge, y de hecho

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=1}^n \mathbf{1}_{\{Z_j \in A\}} = \sum_{i \in I} \mathbf{1}_{\{s_i \in A\}} \tilde{\mu}(s_i). \quad (3.23)$$

Ahora, para ver qué pasa en $k^{(n)}$, empecemos por caracterizarlos. Recordemos que $k^{(n)}$ es el paso correspondiente al tiempo τ_n , y por lo tanto $k^{(n)}$ cuenta cuántos pasos fueron necesarios que pasen hasta que se dé la n -ésima ramificación. Como en cada paso hay probabilidad $\frac{N-1}{2N}$ de que sea una ramificación, $k^{(n)}$ es una suma de n geométricas con dicho parámetro. Es decir,

$$k^{(n)} = \sum_{j=1}^n G_j \text{ con } (G_j)_{j \in \mathbb{N}} \text{ iid, y } G_j \sim Ge\left(\frac{N-1}{2N}\right).$$

Notemos que $k^{(n)} \geq n \forall n$ y, además, con esta descripción vemos que $k^{(n)}$ es creciente en n . Por lo tanto $k^{(n)} \xrightarrow{n \rightarrow \infty} \infty$.

Entonces la sucesión $\left(\frac{1}{k^{(n)}} \sum_{j=1}^{k^{(n)}} \mathbf{1}_{\{Z_j \in A\}}\right)_{n \in \mathbb{N}}$ es subsucesión de la que vimos que convergía en (3.23), y en consecuencia converge al mismo límite.

Por otro lado, por la Ley de los Grandes Números,

$$\frac{k^{(n)}}{n} = \frac{\sum_{j=1}^n G_j}{n} \xrightarrow{n \rightarrow \infty} \mathbb{E}(G_1) = \frac{2N}{N-1}. \quad (3.24)$$

Finalmente, basándonos en la convergencia de la subsucesión y (3.24), tenemos:

$$\begin{aligned} \frac{A_n}{n} &= \frac{\sum_{j=0}^{k^{(n)}-1} \mathbf{1}_{\{Z_j \in A\}}}{n} = \frac{k^{(n)}}{k^{(n)}} \cdot \frac{\sum_{j=0}^{k^{(n)}-1} \mathbf{1}_{\{Z_j \in A\}}}{n} \\ &= \frac{k^{(n)}}{n} \cdot \frac{\sum_{j=0}^{k^{(n)}-1} \mathbf{1}_{\{Z_j \in A\}}}{k^{(n)}} \xrightarrow{n \rightarrow \infty} \frac{2N}{N-1} \cdot \sum_{i \in I} \mathbf{1}_{\{s_i \in A\}} \tilde{\mu}(s_i). \end{aligned} \quad (3.25)$$

Llamemos

$$\alpha = \frac{2N}{N-1} \cdot \sum_{i \in I} \mathbf{1}_{\{s_i \in A\}} \tilde{\mu}(s_i). \quad (3.26)$$

Nos queda así demostrado que $\frac{A_n}{n}$ converge, y

$$\lim_{n \rightarrow \infty} \frac{A_n}{n} = \alpha.$$

□

3. *Tercer término.* Para estudiar B_n el razonamiento en líneas generales es similar al de A_n . Ahora tenemos que sumar 1 cada vez que la primera partícula se desplaza a la derecha, y notemos que este evento se puede dar sólo cuando en la primera posición hay una sola partícula. Si hubiese más de una en esa posición, aunque alguna de ellas se mueva a la derecha, la «primera» (que la definimos como la que está más a la izquierda) sigue estando en el mismo lugar. Con A_n no importaba si había más de una en el primer lugar, pero como sí ocurría ahí, es necesario mirar el comportamiento de toda la cadena para describir el movimiento.

Por eso necesitamos ver qué sucede con las transiciones de la cadena. El desplazamiento a la derecha en la primera partícula se da si $\eta_{J_{n+1}}(1) = \eta_{J_n}(1) + 1$, y por lo tanto, restamos 1 en cada una de las coordenadas de $\Delta_{J_{n+1}}$ exceptuando la primera.

Utilizando las mismas definiciones que en A_n , agregamos:

Definición 3.5.9. $B = \{(x, x-1) \mid x \in \mathcal{S}\}$.

De esta manera, y con razonamientos análogos, nos queda que

- $B_n = \sum_{j=0}^{k^{(n)}-1} \mathbf{1}_{\{Z_j \in B\}}$.
- la sucesión $\left(\frac{1}{n} \sum_{j=1}^n \mathbf{1}_{\{Z_j \in B\}}\right)_{n \in \mathbb{N}}$ converge, con

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=1}^n \mathbf{1}_{\{Z_j \in B\}} = \sum_{i \in I} \mathbf{1}_{\{s_i \in B\}} \tilde{\mu}(s_i).$$

- La sucesión $\left(\frac{1}{k^{(n)}} \sum_{j=1}^{k^{(n)}} \mathbf{1}_{\{Z_j \in B\}}\right)_{n \in \mathbb{N}}$ es subsucesión y converge al mismo límite.

▪

$$\frac{B_n}{n} = \frac{k^{(n)}}{n} \cdot \frac{\sum_{j=0}^{k^{(n)}-1} \mathbf{1}_{\{Z_j \in B\}}}{k^{(n)}} \xrightarrow{n \rightarrow \infty} \frac{2N}{N-1} \cdot \sum_{i \in I} \mathbf{1}_{\{s_i \in B\}} \tilde{\mu}(s_i) =: \beta$$

$$\implies \frac{B_n}{n} \text{ converge, y } \lim_{n \rightarrow \infty} \frac{B_n}{n} = \beta. \quad (3.27)$$

□

Probada la convergencia término a término en (3.22), (3.26) y (3.27), nos queda que

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{\eta_{\tau_n}(1)}{n} &= \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n \Delta_{\tau_k}(2) - \frac{A_n}{n} + \frac{B_n}{n} \\ &= \sum_{i \in I} f(s_i) \hat{\mu}(s_i) - \alpha + \beta = \tilde{a}_N. \end{aligned} \quad (3.28)$$

donde f era la proyección en la segunda coordenada.

□

Demostrada ya la convergencia en los tiempos de parada τ_n , veamos la convergencia en todo el esqueleto.

Proposición 3.5.10.

$$\lim_{k \rightarrow \infty} \frac{\eta_{J_k}(1)}{k} = \tilde{a}_N \frac{N-1}{2N} \quad c.s.$$

Demostración. Dado un k , llamamos $n(k)$ al paso en que se realizó la última ramificación antes de k . Es decir, $n(k) = \sup \{n \leq k : J_n = \tau_{n_0} \text{ para algún } n_0\}$. A su vez, llamamos $m(k)$ al primero después de k : $m(k) = \inf \{m > k : J_m = \tau_{n_0} \text{ para algún } n_0\}$, y así nos queda que $n(k) \leq k \leq m(k)$. Llamamos además $n_0(k)$ al número de ramificación en que se produce la última ramificación antes de k : al n_0 que cumple que $\tau_{n_0} = J_{n(k)}$. Notar que $n_0(k) \leq n(k)$. Tenemos

$$\begin{aligned} \left| \frac{\eta_{J_k}(1)}{k} - \tilde{a}_N \frac{N-1}{2N} \right| &\leq \left| \frac{\eta_{J_k}(1) - \eta_{J_{n(k)}}(1)}{k} \right| + \left| \frac{\eta_{J_{n(k)}}(1)}{k} - \tilde{a}_N \frac{N-1}{2N} \right| \\ &\leq \left| \frac{\eta_{J_k}(1) - \eta_{J_{n(k)}}(1)}{k} \right| + \left| \frac{\eta_{\tau_{n_0(k)}}(1)}{k} - \tilde{a}_N \frac{N-1}{2N} \right|. \end{aligned} \quad (3.29)$$

Veamos el comportamiento de cada término.

$$\blacksquare \quad |\eta_{J_k}(1) - \eta_{J_{n(k)}}(1)| \leq \sup_{n(k) \leq l \leq m(k)} \{|\eta_{J_l}(1) - \eta_{J_{n(k)}}(1)|\} =: X_k$$

X_k es la variable aleatoria que indica la mayor distancia que llega a tomar la cadena después de una ramificación $n_0(k)$. Entre un τ_j y τ_{j+1} la cadena sólo puede desplazarse o ramificar la primera partícula, por lo que en cada paso sólo puede, a lo sumo, sumar o restar 1. En consecuencia, X_k está acotada por la cantidad de

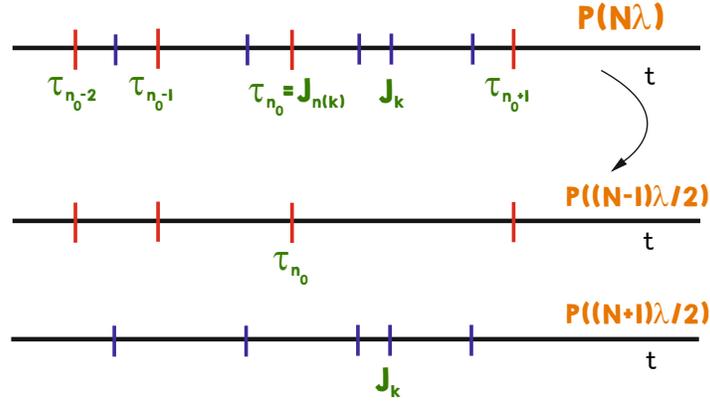


Figura 3.4: Marcas separadas

pasos que realiza la cadena entre el τ anterior y el siguiente a k . Gráficamente tendríamos los tiempos de los relojes divididos en dos conjuntos, como marcas pintadas de dos colores distintos en ω (el proceso de Poisson general), como se puede ver en la Figura 3.5. Un color corresponde a los τ_j que ya habíamos definido antes, y el otro a los pasos entre un τ_j y el siguiente, obteniendo así dos procesos de Poisson independientes.

Llamemos $L^{(k)}$ a la variable que cuenta la cantidad de marcas entre $\tau_{n_0} = J_{n(k)}$ y τ_{n_0+1} , y por lo ya explicado $X_k \leq L^{(k)}$. Trabajar con $L^{(k)}$ es más fácil, debido a que esta tiene una distribución conocida: supongamos que nos paramos en τ_{n_0} y desde ahí miramos el tiempo que tarda cada uno de los dos procesos de Poisson en tener una marca nueva. Si aparece primero en el de los τ lo consideramos un éxito, y sino un fracaso, y repetimos el experimento contando la cantidad de fracasos totales hasta el primer éxito. Basados en la independencia de los dos procesos, y en la pérdida de memoria de la variable exponencial que cuenta el tiempo correspondiente a que aparezca un nuevo τ , podemos afirmar que $L^{(k)}$ tiene distribución geométrica (con soporte desde 0). Formalicemos este razonamiento:

Si L_t es el proceso de Poisson que de todas las marcas, se queda con aquellas que no están incluidas entre los τ_j , su intensidad es

$$N\lambda - \frac{(N-1)\lambda}{2} = \frac{(N+1)\lambda}{2}.$$

Por lo tanto

$$L^{(k)}|_{\tau_{n_0+1}-\tau_{n_0}} = L_{\tau_{n_0+1}} - L_{\tau_{n_0}}|_{\tau_{n_0+1}-\tau_{n_0}} \sim P\left(\frac{(N+1)\lambda}{2}(\tau_{n_0+1} - \tau_{n_0})\right)$$

Además ya sabíamos que $\tau_{n_0+1} - \tau_{n_0} \sim \epsilon \left(\frac{(N-1)\lambda}{2}\right)$, y como demostramos en la Proposición 5.1.5 del Apéndice, estas dos últimas afirmaciones implican que

$$L^{(k)} \sim Ge\left(\frac{(N-1)\lambda}{2} / \left(\frac{(N-1)\lambda}{2} + \frac{(N+1)\lambda}{2}\right)\right).$$

o, simplificando,

$$L^{(k)} \sim Ge\left(\frac{N-1}{2N}\right)$$

Veamos que $\frac{L^{(k)}}{k}$ converge a 0 casi seguramente. Para eso planteamos lo siguiente, a partir de las desigualdades de Tchebychev y la varianza de la geométrica:

$$\begin{aligned} \mathbb{P}\left(\frac{L^{(k)}}{k} > \epsilon\right) &= \mathbb{P}\left(\left(\frac{L^{(k)}}{k}\right)^2 > (\epsilon)^2\right) \leq \frac{\mathbb{E}\left(\left(\frac{L^{(k)}}{k}\right)^2\right)}{(\epsilon)^2} \\ &= \frac{\mathbb{E}\left(\left(L^{(k)}\right)^2\right)}{(k\epsilon)^2} = \frac{2-p}{p^2(k\epsilon)^2} = C \frac{1}{k^2}. \end{aligned}$$

$$\text{Entonces } \sum_{k \in \mathbb{N}} \mathbb{P}\left(\frac{L^{(k)}}{k} > \epsilon\right) \leq \sum_{k \in \mathbb{N}} C \frac{1}{k^2} < \infty.$$

Como los $L^{(k)}$ son independientes, por el Teorema de Borel-Cantelli esto último nos dice que $\mathbb{P}\left(\frac{L^{(k)}}{k} > \epsilon \text{ para infinitos } k\text{'s}\right) = 0$ o lo que es lo mismo

$$\mathbb{P}\left(\frac{L^{(k)}}{k} \xrightarrow[k \rightarrow \infty]{} 0\right) = 1.$$

Y así queda demostrado que el primer término de (3.29) converge a 0.

- Para el segundo término vamos a ver directamente que

$$\lim_{k \rightarrow \infty} \frac{\eta_{\tau_{n_0(k)}}(1)}{k} = \frac{N-1}{2N} \tilde{a}_N.$$

En ese sentido reescribimos la sucesión:

$$\frac{\eta_{\tau_{n_0(k)}}(1)}{k} = \frac{n_0(k)}{n_0(k)} \frac{\eta_{\tau_{n_0(k)}}(1)}{k} = \frac{n_0(k)}{k} \frac{\eta_{\tau_{n_0(k)}}(1)}{n_0(k)}.$$

$n_0(k)$ es el número de la última ramificación antes del paso k , por lo que para calcularlo podemos contar la cantidad de veces que en un paso de la cadena se dio una ramificación. Así, nos queda expresado como $n_0(k) = \sum_{i=1}^k B_i$ con $B_i \sim Be((N-1)/2N)$ y la familia de $(B_i)_{i \in \mathbb{N}}$ *iid*.

Apelando nuevamente a la Ley de los Grandes Números,

$$\frac{n_0(k)}{k} = \frac{\sum_{i=1}^k B_i}{k} \xrightarrow[k \rightarrow \infty]{} \mathbb{E}(B_1) = \frac{N-1}{2N}.$$

Por otro lado, notemos que la sucesión $n_0(k)$ es creciente, y la afirmación siguiente:

Lema 3.5.11.

$$n_0(k) \xrightarrow[k \rightarrow \infty]{} \infty.$$

Es decir que, a medida que k crece, el número de la última ramificación anterior al paso k también crece tanto como queramos.

Demostración del Lema. La idea detrás de esta afirmación es que las ramificaciones siempre se siguen realizando, y por lo tanto aunque miremos al esqueleto en un paso muy grande, alguna ramificación cercana se va a dar. Hacemos esta aclaración porque con la notación se hace difícil entender qué estamos demostrando.

Sea $M \in \mathbb{N}$ y consideramos el paso en el esqueleto correspondiente a la M -ésima ramificación (que en el Teorema 3.5.5 lo llamamos $k^{(M)}$). Hagamos la aclaración (que no hicimos en ese momento) de que $\mathbb{P}(k^{(M)} < \infty) = 1$ para cualquier M . Esto sucede porque

$$k^{(M)} = \sum_{j=1}^M G_j,$$

donde G_j son geométricas *iid* con $G_j \sim Ge((N-1)/2N)$.

$$\text{Entonces } \mathbb{E}(k^{(M)}) = M\mathbb{E}(G_j) = M \frac{2N}{N-1} < \infty.$$

Por lo tanto, si $k_0 = k^{(M)}$, $\forall k \geq k_0$, $n_0(k) \geq n_0(k_0) = n_0(k^{(M)}) = M$

□

De esta manera tenemos que $\left(\frac{\eta_{\tau_{n_0(k)}}(1)}{n_0(k)}\right)_{k \in \mathbb{N}}$ es subsucesión de utilizada en la Proposición 3.5.2, y por eso, por lo demostrado en dicho Teorema,

$$\lim_{k \rightarrow \infty} \frac{\eta_{\tau_{n_0(k)}}(1)}{n_0(k)} = \lim_{n \rightarrow \infty} \frac{\eta_{\tau_n}(1)}{n} = \tilde{a}_N$$

Entonces

$$\lim_{k \rightarrow \infty} \frac{\eta_{\tau_{n_0(k)}}(1)}{k} = \lim_{k \rightarrow \infty} \frac{n_0(k)}{k} \frac{\eta_{\tau_{n_0(k)}}(1)}{n_0(k)} = \frac{N-1}{2N} \tilde{a}_N$$

A partir de la cota que dimos en (3.29), y con la convergencia de ambos ítems en el paso anterior, concluimos que $\lim_{k \rightarrow \infty} \frac{\eta_{J_k}(1)}{k} = \tilde{a}_N \frac{N-1}{2N}$ casi seguramente. \square

Corolario 3.5.12. La sucesión

$$\frac{\eta_{J_k}(1)}{J_k}$$

$$\text{converge casi seguramente a } a_N = \tilde{a}_N \frac{N-1}{2N} \frac{\lambda}{2} = \tilde{a}_N \frac{N-1\lambda}{N4}.$$

Demostración. Como $J_k = \sum_{j=1}^k S_j$ con $S_j \sim \epsilon(\lambda/2)$ e *iid*, por la Ley de los Grandes Números tenemos

$$\frac{J_k}{k} = \frac{1}{k} \sum_{j=1}^k S_j \xrightarrow{n \rightarrow \infty} \mathbb{E}(S_1) = \frac{2}{\lambda}.$$

Por lo tanto, y por la Proposición 3.5.10

$$\lim_{n \rightarrow \infty} \frac{\eta_{J_k}(1)}{J_k} = \lim_{n \rightarrow \infty} \frac{\eta_{J_k}(1)}{k} \frac{k}{J_k} = \tilde{a}_N \frac{N-1}{2N} \frac{\lambda}{2} = a_N. \quad (3.30)$$

\square

Como corolario tenemos la velocidad asintótica de la primera partícula en tiempo continuo, como enunciamos en el Teorema 3.5.1:

Corolario 3.5.13. Existe un $a_N \in \mathbb{R}$ tal que

$$\lim_{t \rightarrow \infty} \frac{\eta_t(1)}{t} = a_N.$$

Demostración. Utilizamos el siguiente lema, que vale en general para funciones con dominio real.

Lema 3.5.14. Sea $f : \mathbb{R} \rightarrow \mathbb{Z}$. Sea $(x_n)_{n \in \mathbb{N}}$ una sucesión de números reales positivos que cumple que

1. es estrictamente creciente
2. $\lim_{n \rightarrow \infty} x_n = \infty$
3. $f(t) = f(x_n) \quad \forall x_n \leq t < x_{n+1}$. Es decir, la función es «constante a trozos», dando los saltos sólo en la sucesión x_n .
- 4.

$$\lim_{n \rightarrow \infty} f(x_n) = a, \text{ con } a \in \mathbb{R} \quad (3.31)$$

Entonces $\lim_{t \rightarrow \infty} f(t) = a$

Demostración del lema. Sea $\epsilon > 0$. Como tenemos el límite en (3.31), sabemos que

$$\exists n_0 / \forall n \geq n_0 \quad |f(x_n) - a| < \epsilon \quad (3.32)$$

Tomamos $t_0 = x_{n_0}$, y veamos que este t_0 nos sirve para ver la convergencia del límite cuando $t \rightarrow \infty$.

Sea $t \geq t_0$. Por (2) sabemos que

$$\exists m / \forall n \geq m \quad x_n > t$$

Sin pérdida de generalidad, suponemos que m es el mínimo que cumple eso. De esa manera, $t \in [x_{m-1}, x_m)$, y por (3) tenemos que $f(t) = f(x_{m-1})$. Como $t \geq t_0 = x_{n_0}$, $m - 1 \geq n_0$. Por 3.32

$$|f(t) - a| = |f(x_{m-1}) - a| < \epsilon$$

□

El corolario es consecuencia directa del lema. Sea f la siguiente función:

$$\begin{aligned} f : \mathbb{R} &\rightarrow \mathbb{Z} \\ t &\mapsto \frac{\eta_t(1)}{t} \end{aligned}$$

Consideramos la sucesión $(J_n)_{n \in \mathbb{N}}$ de los tiempos de los saltos en la cadena original. La misma cumple las hipótesis del lema:

1. es estrictamente creciente.

2. $\lim_{n \rightarrow \infty} J_n = \infty$. Esto es lo que se llama que la cadena no tenga «explosión». Si la suma de los tiempos de espera de los relojes no converge, la cadena no tiene explosión, y en este caso pasa porque los tiempos de los relojes los definimos para que se mantengan constantes.
3. $f(t) = f(J_n) \quad \forall J_n \leq t < J_{n+1}$ por la definición de los (J_n)

Como tenemos en el primer corolario del teorema, en la ecuación (3.30):

$$\lim_{n \rightarrow \infty} f(J_n) = \lim_{n \rightarrow \infty} \frac{\eta_{J_n}(1)}{J_n} = a_N, \text{ con } a_N \in \mathbb{R}$$

Entonces, por el lema tenemos que

$$\lim_{t \rightarrow \infty} \frac{\eta_t(1)}{t} = a_N$$

□

Aquí concluimos con los resultados que obtuvimos durante el desarrollo de la tesis. Los mismos responden a las primeras propiedades que buscamos en este sistema de partículas. Si bien por cuestiones de tiempo y extensión no ahondamos en más propiedades, sí tenemos una orientación para continuar este trabajo, por ejemplo en el marco de un doctorado.

Primero, queremos mostrar que a medida que crece la cantidad de partículas, el comportamiento de cada una de ellas tiende a ser independiente del de las demás. Para ser más precisos, definimos la medida empírica dada por una configuración $\eta_t \in \mathbb{Z}^N$ como

$$m(\eta_t) := \frac{\sum_{x \in \mathbb{Z}} \sum_{i=1}^N \mathbf{1}_{\{\eta_t(i)=x\}} \cdot \delta_x}{N}$$

y notamos $m_x(\eta_t)$ a $m(\eta_t)(x)$. Luego, nuestro objetivo es demostrar la siguiente proposición:

Proposición 3.5.15. Dados $x, y \in \mathbb{Z}$, tenemos una función $K(N, t)$ que satisface

1. $\lim_{N \rightarrow \infty} K(N, t) = 0$, para cualquier $t > 0$ fijo.
2. Para cada N y $t > 0$,

$$\sup_{\eta_t \in \mathbb{Z}^N} \left| \mathbb{E}_\eta^N [m_x(\eta_t) m_y(\eta_t)] - \mathbb{E}_\eta^N [m_x(\eta_t)] \mathbb{E}_\eta^N [m_y(\eta_t)] \right| \leq K(N, t).$$

Como corolario de esta proposición se deduce que la correlación de la medida empírica en dos sitios cualesquiera converge a cero cuando N tiende a infinito. Podemos encontrar un resultado similar en [1, p.3]. A este tipo de resultados se los conoce como *Propagación del Caos*, del que tenemos referencia en [15, 26].

También nos interesa llegar a resultados análogos a los alcanzados en [8] por Durrett y Remenik para el sistema en \mathbb{R} . En su caso, también se trata de N partículas que se ramifican; cuando se da nacimiento a una nueva, su posición es elegida a partir de la ubicación de la original, usando una determinada distribución ρ sobre \mathbb{R} .

Los autores prueban que la medida empírica del sistema converge a una medida de probabilidad determinística ν_t para todo t . Dicha medida es absolutamente continua, con densidad $u(t, \cdot)$, solución del siguiente problema de frontera libre

$$\begin{aligned} \frac{\partial u}{\partial t}(t, x) &= \int_{-\infty}^{\infty} u(t, y) \rho(x - y) dy \quad \forall x > \gamma(t), \\ \int_{\gamma(t)}^{\infty} u(t, y) dy &= 1, \quad u(t, x) = 0, \quad \forall x \leq \gamma(t), \\ u(0, x) &= u_0(x). \end{aligned} \tag{3.33}$$

Asímismo, como comentamos al final del capítulo anterior, los autores demuestran que las velocidades v_N de la nube en el equilibrio convergen a la velocidad de la onda viajera minimal, otro punto a analizar para nuestro caso en \mathbb{Z} .

Capítulo 4

Sobre el contexto histórico

El artículo de Fisher ([12]) fue publicado en una revista llamada «Annals of Eugenics», que se editó sólo entre el 1925 y el 1952. El hecho de que la revista no se editase más nos hizo buscar información al respecto, llegando así a conocer el interesante y cuestionable contexto en que estas investigaciones se desarrollaron.

4.1. Estadística entre 1850 y 1950. Eugenesia y Darwinismo Social.

Como explica el sociólogo Donald Mackenzie ([21]), el crecimiento que la estadística vivió en Inglaterra desde los aportes de Sir Francis Galton (1822-1911) en adelante estuvo atravesado directamente por una ideología preponderante en los exponentes que caracterizaron esta época. Galton, primo de Darwin, estaba fuertemente interesado en el «progreso» del ser humano, las características de su reproducción y su futuro como especie. La herencia genética era un tema central para poder entender qué depararían las siguientes generaciones humanas, y un estudio estadístico de ella permitiría confirmar o refutar hipótesis y conjeturas planteadas en ese momento histórico.

En este contexto surgió la *eugenesia*: una corriente filosófica que pregonaba la mejora de los rasgos hereditarios del ser humano mediante la aplicación de políticas directas, la selección de genes «beneficiosos» y la intervención sobre la reproducción, basándose en conocimientos de distintas disciplinas como eran la biología y la estadística. Galton fue uno de los pioneros en este campo, ya desde 1865 defendiendo la necesidad concreta de hacer una selección artificial de seres humanos, controlando la reproducción de manera positiva (fomentando la reproducción de los más aptos) o negativa (limitando la de los menos aptos). Cabe destacar que la eugenesia surgió a principios de siglo XX como una disciplina científica, siguiendo el método científico aceptado para la época, y fue tratada como tal por lo menos hasta el fin de la Segunda Guerra Mundial.

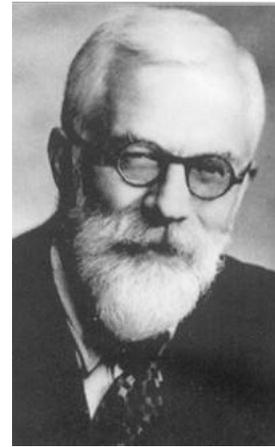
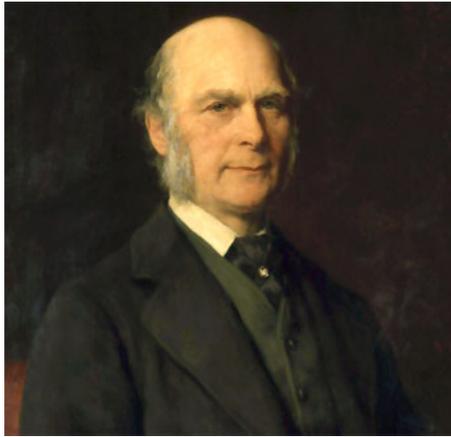


Figura 4.1: Sir Francis Galton y Sir Ronald Fisher. Estadísticos del siglo XIX y XX, y pensadores de la eugenesia.

En términos de los eugenesistas, los individuos más aptos son aquellos que mejor se adaptan a su sociedad, los que mejor posicionados se encuentran, los que más recursos tienen para subsistir. Los ricos, los inteligentes son los sujetos sociales que cuadran con esta descripción. Por el contrario, los no aptos son los pobres, los enfermos, los discapacitados, los delincuentes. En fin, se trata de un discurso que le adjudica una ventaja natural a las clases acomodadas y dominantes dentro de la especie humana.

Una disciplina emparentada con la eugenesia, que fue contemporánea y tuvo más repercusión, fue el llamado *darwinismo social*. Como su nombre lo indica, consistía en aplicar las leyes que Darwin mismo describió ([4]) a las sociedades humanas. Dentro de una especie aparecen nuevos rasgos genéticos de forma aleatoria, siendo algunos más propensos que otros a la adaptación de la especie al medio ambiente donde vive. Los individuos que poseen esos rasgos mejor adaptados son los que están más capacitados para reproducirse, reproduciendo así los rasgos nuevos, y haciendo que perduren en las siguientes generaciones frente a los rasgos viejos. A este fenómeno se lo conoce como *selección natural*. Spencer ([25]), que escribió en simultáneo a Darwin, fue el primero en utilizar el mismo razonamiento para la especie humana, concluyendo que por cuestiones hereditarias había individuos mejor preparados para la reproducción, siendo ellos los que mayor y mejor descendencia podían dejar. Hilando fino en sus postulados, el darwinismo social difería de la eugenesia en ver este desarrollo humano como una cuestión natural; por eso no compartía con la eugenesia la idea de tomar políticas de estado que estimulen su evolución, tales como el control de la reproducción.

No es necesario profundizar ampliamente en la historia del siglo XX para imaginarse

la relación entre estas filosofías, y las expresiones políticas con bases racistas, tales como el nazismo en la Alemania del Tercer Reich. En [17] se desarrolla dicho vínculo en detalle.

En [21] hay una explicación más detallada de los trabajos de los estadísticos de principios de siglo XX relacionados con la eugenesia. Pearson, socialista utópico y tecnócrata, dirigió el «Department of Applied Statistics» en la University College of London, y ganó el cargo «Galton Professor of Eugenics». Fisher se involucró en la fundación de la Sociedad de Eugenesia de la Universidad de Cambridge en 1911, como también se puede ver en [6]. En conferencias de la Sociedad Fisher hizo públicos sus estudios extracurriculares de estadística para aplicar a la herencia genética. Finalmente, Mackenzie también en [21] da un ejemplo comparativo de por qué la estadística desarrollada por Pearson era funcional a los postulados que busca justificar la eugenesia, frente a la estadística que en simultáneo desarrolló Yule.

4.2. Fisher y los genes ventajosos

En el Capítulo 2 comentamos el problema que Fisher abrió a partir de [12], donde planteaba una ecuación que modelaba la dispersión de *genes ventajosos* a lo largo del tiempo y espacio. Reconsiderando su artículo en este contexto concluimos que los genes a los que se refiere podrían ser genes humanos.

Quiero escribir ahora hablando a título personal, al menos en esta sección, para explicitar que no veo apropiado el direccionamiento de la ciencia hacia el estudio de determinados genes ventajosos y el potencial fomento de su reproducción.

Por un lado, sin haber profundizado en estudios de genética, no creo que los principios de la selección natural se puedan aplicar al género humano al igual que se aplican a otras especies. La razón es que el hombre en sociedad generó un medio ambiente propio, artificial, ampliamente modificado con respecto a la naturaleza previamente predominante. Esto hace que la capacidad de reproducción y supervivencia que tienen los individuos se vean muy afectadas por la posición social que les tocó vivir por razones históricas y sociales (la propiedad de los medios de producción, las políticas de estado, las luchas culturales, etc.). La selección natural afecta de manera mucho más lenta a las especies, quedando así diluido su efecto en los humanos.

Por otro lado, si bien me parece acertado que la ciencia se dirija en función de necesidades sociales, el sentido debe ser exactamente el opuesto al buscado por la eugenesia. El conocimiento producido, sobre todo por las universidades públicas, tiene que ser popular y emancipador, pensado en función de los intereses de las clases excluidas. La socialización de los medios de producción, la urbanización de las villas en las ciudades, la reforma agraria son tareas que tenemos pendientes como sociedad y donde la estadística, así como otras áreas de la matemática y otras disciplinas científicas, tienen mucho que aportar.

De todas maneras, no cuestiono que el problema que planteó Fisher y sus derivados se sigan estudiando al día de hoy, dado que el modelo abarca situaciones reales muy diversas y distintas de la correspondiente a los genes ventajosos. Como enumeran Brunet y Derrida [3], con la misma ecuación se puede modelar la dinámica de individuos enfermos en una población durante una epidemia, la proporción de gases quemados en una combustión [16], la concentración de especies en una reacción química, la distribución de polímeros dirigidos en un medio aleatorio [5] e incluso la dispersión de un rumor en un grupo social.

4.3. La ciencia como producción social y cultural

«Cuando se trata de grupos humanos foráneos, por otro aparte, se nos hace perceptible con bastante facilidad el condicionamiento práctico de los sistemas de conocimientos. (...) Por muy orgullosos que estemos de nuestra propia “ciencia”, por muy convencidos que estemos de su trascendencia, hagamos pues un esfuerzo para verla también a ella como una empresa con raíces terrestres. Y preguntémosnos: ¿Cuál puede ser el proyecto fundamental encarnado (entre otros) en la llamada “ciencia moderna”?»

Pierre Thuillier, La Trastienda del Sabio, Editorial Hachette, Barcelona 1983

A partir de las polémicas que abren estas investigaciones, y como conclusión, es pertinente prestar atención a determinados asuntos de corte más epistemológico. Como se trata de una producción de carácter social y cultural, la ciencia está atravesada por intereses y pujas que se ven expuestos de forma cotidiana en distintos ámbitos dentro de una sociedad. La eugenesia es una muestra muy gráfica al respecto, dado que su estudio y enseñanza se realizó formalmente en universidades muy prestigiosas dentro del ámbito científico (como es la Universidad de Cambridge), hasta que por determinados acontecimientos históricos se dejó de lado, al punto de ya no ser considerada una disciplina científica. Estos acontecimientos fueron y son conocidos en todo el mundo, y tuvieron efectos sobre la vida de millones de personas.

Citamos dos simples ejemplos que grafican esta relación estrecha: en 1939, año en que comienza la Segunda Guerra Mundial, se cierra la cátedra de eugenesia de la Universidad de Cambridge de la que Fisher era el titular. En 1952, unos años después de consumado el Holocausto, y con las investigaciones respectivas en curso, se deja de publicar la revista *Annals of Eugenics*.

Es interesante notar cómo aquello que se considera una verdad científica no está atado únicamente a la validez de preceptos lógicos, sino que está influenciado por la fuerza con la que se imponen ciertos discursos, y su origen en las entrañas de la sociedad y de la historia.

Para profundizar en estas discusiones y estudiar cómo se plasman en la política científica de la Argentina, recomendamos fuertemente la lectura del libro [27], que escribió

en 1969 Oscar Varsavsky, quien fue doctor en Química y profesor del Departamento de Matemática de la Facultad de Ciencias Exactas y Naturales de la UBA.

Por todo esto, y por muchos ejemplos más, queda claro que la ciencia no tiene una posición neutral en cada sociedad, sino que por el contrario es una trinchera más desde la cual se encuentran los intereses de clase, como ya mencionamos en la sección anterior. Hacemos énfasis en este aspecto, porque creemos que no es menor para pensarnos como científicos críticos en este lugar y en este momento histórico.

Capítulo 5

Apéndice

5.1. Exponenciales

En este apéndice vamos a comentar algunas propiedades de las variables aleatorias de distribución exponencial. Las mismas fueron usadas en distintos momentos del trabajo.

Teorema 5.1.1 (Propiedad de la Falta de Memoria). Sea $X \sim \epsilon(\lambda)$, entonces

$$\mathbb{P}(X > s + t | X > t) = \mathbb{P}(X > s) \quad \forall s, t \in \mathbb{R}_{\geq 0}.$$

Demostración.

$$\begin{aligned} \mathbb{P}(X > s + t | X > t) &= \frac{\mathbb{P}(X > s + t, X > t)}{\mathbb{P}(X > t)} = \frac{\mathbb{P}(X > s + t)}{\mathbb{P}(X > t)} = \\ &= \frac{1 - F(s + t)}{1 - F(t)} = \frac{e^{-(s+t)}}{e^{-t}} = e^{-s} = \mathbb{P}(X > s). \end{aligned}$$

□

En palabras, este teorema dice que la distribución de una exponencial es la misma si la observamos desde el principio, o si sabemos que es mayor que un tiempo t y la observamos desde ahí.

Es importante destacar que también vale el recíproco: si una variable aleatoria continua cumple esta propiedad, entonces es exponencial, como se puede ver en [22, p. 71].

Teorema 5.1.2. Sean X_1, \dots, X_n variables aleatorias *iid* con $X_i \sim \epsilon(\lambda) \forall 1 \leq i \leq n$, y sea $Z = \sum_{i=1}^n X_i$. Entonces $Z \sim \Gamma(n, \lambda)$.

Teorema 5.1.3. Sean S_1, \dots, S_N variables aleatorias *iid* con $S_i \sim \epsilon(\lambda_i)$. Entonces la variable aleatoria $S = \min_{1 \leq i \leq n} S_i$ tiene distribución exponencial de parámetro $\sum_{i=1}^N \lambda_i$, y $\mathbb{P}(S = S_i) = \lambda_i / \sum_{j=1}^N \lambda_j \quad \forall 1 \leq i \leq N$.

Demostración.

$$\begin{aligned}
\mathbb{P}(S \leq x) &= \mathbb{P}(S_i \leq x \text{ para algún } i) \\
&= 1 - \mathbb{P}(S_i \geq x \ \forall i) = 1 - \mathbb{P}(S_1 > x) \dots \mathbb{P}(S_N > x) \\
&= 1 - \left[1 - \left(1 - e^{-\lambda_1 x}\right)\right] \dots \left[1 - \left(1 - e^{-\lambda_N x}\right)\right] \\
&= 1 - e^{-\lambda_1 x} \dots e^{-\lambda_N x} = 1 - e^{-\sum_{i=1}^N \lambda_i x}.
\end{aligned}$$

$$\Rightarrow S \sim \epsilon(\sum_{i=1}^N \lambda_i). \quad \square$$

Proposición 5.1.4. Sea G una variable aleatoria con $G \sim Ge(p)$, $(X_i)_{i \geq 1}$ variables aleatorias *iid*, y sea $Z = \sum_{i=1}^G X_i$. Entonces $\mathbb{E}(Z) = \mathbb{E}(X_i)\mathbb{E}(G)$. En particular, si $\mathbb{E}(X_i) < \infty \Rightarrow \mathbb{E}(Z) < \infty$.

Enunciamos este resultado para $(X_i)_{i \geq 1}$ cualesquiera, dado que durante el trabajo lo usamos tanto para exponenciales como para gammas.

Demostración.

$$\begin{aligned}
\mathbb{E}(Z) &= \mathbb{E}\left(\sum_{i=1}^G X_i\right) = \sum_{n \in \mathbb{N}} \mathbb{E}\left(\sum_{i=1}^n X_i \mid G = n\right) \cdot \mathbb{P}(G = n) \\
&= \sum_{n \in \mathbb{N}} \mathbb{E}\left(\sum_{i=1}^n X_i\right) \cdot \mathbb{P}(G = n) = \sum_{n \in \mathbb{N}} n \cdot \mathbb{E}(X_i) \cdot \mathbb{P}(G = n) \\
&= \mathbb{E}(X_i) \cdot \sum_{n \in \mathbb{N}} n \cdot \mathbb{P}(G = n) = \mathbb{E}(X_i) \cdot \mathbb{E}(G).
\end{aligned}$$

\square

Proposición 5.1.5. Sean X y τ variables aleatorias tales que

$$\tau \sim \epsilon(\lambda') \quad \text{y} \quad X|_{\tau} \sim P(\lambda\tau).$$

Entonces $X \sim Ge\left(\frac{\lambda'}{\lambda + \lambda'}\right)$.

Demostración. Llamamos f a la función de densidad de τ

$$f(x) = \lambda' e^{-\lambda' x}$$

Entonces, teniendo en cuenta que $X|_{\tau} \sim P(\lambda\tau)$,

$$\begin{aligned}
\mathbb{P}(X = k) &= \int_0^{\infty} \mathbb{P}(X_k | \tau = t) f(t) dt = \int_0^{\infty} \frac{e^{-\lambda t} (\lambda t)^k}{k!} \lambda' e^{-\lambda' t} dt \\
&= \lambda' \lambda^k \int_0^{\infty} \frac{e^{-(\lambda + \lambda') t} t^k}{k!} dt
\end{aligned} \tag{5.1}$$

Para continuar, utilizamos el hecho de que si tenemos S una variable aleatoria con distribución $\Gamma(k, \lambda)$, su función de densidad cumple que:

1.

$$f_S(t) = \lambda e^{-\lambda t} \frac{(\lambda t)^{k-1}}{\Gamma(k)}$$

2.

$$\int_0^{\infty} f_S(t) dt = 1$$

□

Siguiendo desde (5.1):

$$\begin{aligned} \mathbb{P}(X = k) &= \lambda' \lambda^k \int_0^{\infty} \frac{e^{-(\lambda+\lambda')t} t^k}{k!} \frac{(\lambda + \lambda')^{k+1}}{(\lambda + \lambda')^{k+1}} dt \\ &= \frac{\lambda' \lambda^k}{(\lambda + \lambda')^{k+1}} \int_0^{\infty} (\lambda + \lambda') e^{-(\lambda+\lambda')t} \frac{[(\lambda + \lambda') t]^k}{\Gamma(k+1)} dt \\ &= \frac{\lambda' \lambda^k}{(\lambda + \lambda')^{k+1}} \\ &= \left(\frac{\lambda}{\lambda + \lambda'} \right)^k \frac{\lambda'}{\lambda + \lambda'} \\ &= \left(\frac{\lambda + \lambda' - \lambda'}{\lambda + \lambda'} \right)^k \frac{\lambda'}{\lambda + \lambda'} \\ &= \left(1 - \frac{\lambda'}{\lambda + \lambda'} \right)^k \frac{\lambda'}{\lambda + \lambda'} \\ &= (1 - p)^k p \end{aligned} \tag{5.2}$$

Proposición 5.1.6. Si X es una variable aleatoria exponencial de parámetro λ , y a es una variable aleatoria real, entonces

$$aX \sim \epsilon(\lambda/a).$$

Bibliografía

- [1] Amine Asselah, Pablo A. Ferrari, and Pablo Groisman. Quasistationary distributions and Fleming-Viot processes in finite spaces. *J. Appl. Probab.*, 48(2):322–332, 2011.
- [2] Eric Brunet and Bernard Derrida. Shift in the velocity of a front due to a cutoff. *Physical Review E*, 56(3):2597, 1997.
- [3] Eric Brunet and Bernard Derrida. Microscopic models of traveling wave equations. *Computer Physics Communications*, 121:376–381, 1999.
- [4] Charles Darwin and William F Bynum. *The origin of species by means of natural selection: or, the preservation of favored races in the struggle for life*. AL Burt. First Edition 1859, 2009.
- [5] B. Derrida and H. Spohn. Polymers on disordered trees, spin glasses, and traveling waves. *J. Statist. Phys.*, 51(5-6):817–840, 1988. New directions in statistical mechanics (Santa Barbara, CA, 1987).
- [6] Veronica di Mambro. The university of cambridge eugenics society from 1911-13 and 1930-33 and reasons for its ultimate demise. *Newsletter del Galton Institute de Inglaterra*, 2003.
- [7] Rick Durrett. *Probability: theory and examples*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, fourth edition, 2010.
- [8] Rick Durrett and Daniel Remenik. Brunet-Derrida particle systems, free boundary problems and Wiener-Hopf equations. *Ann. Probab.*, 39(6):2043–2078, 2011.
- [9] Analía Ferrari. Distribuciones cuasiestacionarias y procesos de ramificación multitempo.
- [10] Pablo A Ferrari and Antonio Galves. Construction of stochastic processes, coupling and regeneration. *XIII Escuela Venezolana de Matemática*, 2000.
- [11] 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.

- [12] Ronald Aylmer Fisher. The wave of advance of advantageous genes. *Annals of Eugenics*, 7(4):355–369, 1937.
- [13] Tertuliano Franco and Pablo Groisman. A particle system with explosions: law of large numbers for the density of particles and the blow-up time. *J. Stat. Phys.*, 149(4):629–642, 2012.
- [14] Hans-Otto Georgii. *Stochastics*. de Gruyter Textbook. Walter de Gruyter & Co., Berlin, extended edition, 2013. Introduction to probability and statistics, Translated from the German original [MR2397455] by Marcel Ortgiese, Ellen Baake and Georgii.
- [15] M. Kac. Foundations of kinetic theory. In *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, 1954–1955, vol. III*, pages 171–197. University of California Press, Berkeley and Los Angeles, 1956.
- [16] Alan R Kerstein. Computational study of propagating fronts in a lattice-gas model. *Journal of Statistical Physics*, 45(5-6):921–931, 1986.
- [17] Stefan Kühl. *Nazi Connection: Eugenics, American Racism, and German National Socialism*. Oxford University Press, 2002.
- [18] David A. Levin, Yuval Peres, and Elizabeth L. Wilmer. *Markov chains and mixing times*. American Mathematical Society, Providence, RI, 2009. With a chapter by James G. Propp and David B. Wilson.
- [19] Thomas M. Liggett. *Stochastic interacting systems: contact, voter and exclusion processes*, volume 324 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1999.
- [20] Thomas M. Liggett. *Continuous time Markov processes*, volume 113 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2010. An introduction.
- [21] Donald Mackenzie. Statistics and ideology: the british school of statistics 1865-1925. *Paper leído en el Encuentro de Verano de la Sociedad Británica de la Historia de la Ciencia*, pages 4–7, 1976.
- [22] J. R. Norris. *Markov chains*, volume 2 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge, 1998. Reprint of 1997 original.
- [23] Sidney Resnick. *Adventures in stochastic processes*. Birkhäuser Boston, Inc., Boston, MA, 1992.
- [24] Timo Seppäläinen. Translation invariant exclusion processes. *Lecture notes*, 2008.

- [25] Herbert Spencer. *First principles*. [sn], 1904.
- [26] Alain-Sol Sznitman. Topics in propagation of chaos. In *École d'Été de Probabilités de Saint-Flour XIX—1989*, volume 1464 of *Lecture Notes in Math.*, pages 165–251. Springer, Berlin, 1991.
- [27] Oscar Varsavsky. *Ciencia, política y cientificismo*. Centro Editor de América Latina Buenos Aires, 1969.