



**UNIVERSIDAD NACIONAL AUTÓNOMA  
DE MÉXICO**

---

---

**FACULTAD DE CIENCIAS**

**Modelos de Wright-Fisher para poblaciones de genes.**

**T E S I S**

**QUE PARA OBTENER EL TÍTULO DE:**

**M A T E M Á T I C A**

**P R E S E N T A:**

**Mariana Gleason Freidberg**



**DIRECTOR DE TESIS:  
Dra. Ana Meda Guardiola  
2012**



Universidad Nacional  
Autónoma de México



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.



# Modelos de Wright-Fisher para Poblaciones de Genes.

Mariana Gleason Freidberg.

29 de octubre de 2012

## Datos del Jurado

1. Datos del alumno  
Gleason  
Freidberg  
Mariana  
56 59 81 87  
Universidad Nacional Autónoma de México  
Facultad de Ciencias  
Matemáticas  
304771077
  
2. Datos del tutor  
Dra.  
Ana  
Meda  
Guardiola
  
3. Datos del sinodal 1  
Dra.  
Eliane Regina  
Rodrigues
  
4. Datos del sinodal 2  
Dr.  
Gerónimo Francisco  
Uribe  
Bravo
  
5. Datos del sinodal 3  
Dr.  
Pedro Eduardo  
Miramontes  
Vidal
  
6. Datos del sinodal 4  
Dra.

María Asunción Begoña  
Fernández  
Fernández

7. Datos del trabajo escrito  
Modelos de Wright-Fisher para poblaciones de genes  
95 p.  
2012

# Índice general

<b>Prefacio</b>	<b>VI</b>
<b>Introducción</b>	<b>x</b>
<b>1. Modelos de Wright-Fisher Discretos</b>	<b>2</b>
1.1. Haploide con 2 Alelos . . . . .	5
1.2. Haploide con r Alelos . . . . .	9
1.3. Haploide con 2 Alelos y Mutación . . . . .	11
1.4. Diploide con r Alelos . . . . .	13
1.5. Diploide con r Alelos y Selección . . . . .	18
1.6. Diploide r Alelos y Selección Multiplicativa . . . . .	23
1.7. Mutación y Selección General o Multiplicativa . . . . .	30
1.7.1. Mutación y Selección General. . . . .	31
1.7.2. Mutación y Selección Multiplicativa . . . . .	33
<b>2. Análisis Sencillo de los Modelos</b>	<b>43</b>
2.1. Haploide con 2 Alelos . . . . .	44
2.1.1. Clasificación de Estados . . . . .	44
2.1.2. Probabilidades de Absorción . . . . .	45
2.1.3. Simulación . . . . .	48
2.2. Haploide con r Alelos . . . . .	50
2.2.1. Clasificación de Estados . . . . .	50
2.2.2. Probabilidades de Absorción . . . . .	51
2.3. Haploide con 2 Alelos y Mutación . . . . .	54
2.3.1. Caso uno, $\mu_{1,2}$ y $\mu_{2,1}$ mayores a 0 . . . . .	54
2.3.2. Caso dos, $\mu_{1,2}$ mayor a 0 y $\mu_{2,1} = 0$ . . . . .	55
2.4. Diploide con r Alelos, Selección Multiplicativa y Mutación . . . . .	56

<b>3. Convergencia Débil a un Proceso de Difusión</b>	<b>59</b>
3.1. La sucesión $\mathcal{Y} = \{Y^N\}_{N \in \mathbb{N}}$ es Tensa . . . . .	61
3.2. El generador de la subsucesiones convergentes . . . . .	68
3.2.1. El generador y el Teorema de Taylor . . . . .	69
3.2.2. El Problema de Martingala y las esperanzas condicio- nales . . . . .	71
3.3. Unicidad de la solución al Problema de Martingala de $\mathcal{L}$ . . .	82
3.4. La Esperanza del Tiempo de Absorción . . . . .	84
<b>A. Momentos y otros cálculos de la variable <math>X_{n+1}</math> distribuida como una <math>Binomial(N, \frac{X_n}{N})</math></b>	<b>89</b>
A.1. Momentos. . . . .	89
A.2. La esperanza condicional del incremento en un paso al cuadrado.	90
A.3. La esperanza condicional del incremento en un paso al cubo .	92
<b>B. Notación.</b>	<b>93</b>
<b>Bibliografía</b>	<b>94</b>

# Prefacio

Tanto el capítulo 1 como el capítulo 2 están al alcance de cualquier alumno de la Facultad de Ciencias que haya aprobado el curso de Procesos Estocásticos. En cuanto al capítulo 3, considero que cualquier buen alumno que haya cursado Procesos Estocásticos 2 en la Facultad de Ciencias podría entender el material expuesto, aceptando dar por sentados los teoremas citados, la definición del Problema de Martingala, y la del generador de un proceso.

Me encuentro muy satisfecha con el capítulo 3 de mi tesis, no sólo porque a mi parecer expone una técnica muy elegante y práctica de aproximar a algunas cadenas de Markov por medio de procesos de difusión, sino también porque ha representado un verdadero reto para mí. Al inicio de esta tesis no lograba descifrar aquellas tres páginas de las notas de Donald Dawson en las que se encontraba la receta para construir el proceso de difusión de Wright-Fisher y he logrado demostrar y comprender (hasta cierto nivel) cada uno de los pasos, los cuales son expuestos con lujo de detalle en el capítulo 3.

También me siento muy orgullosa de muchos resultados del capítulo 1 y 2. Estos últimos no me costaron tanto trabajo como los del capítulo 3, pero también tienen su mérito, y los veo con cierto cariño.

La verdad es que la forma en que construí la mayor parte del capítulo 1 fue tratando de averiguar por qué definieron al modelo de Wright-Fisher de la forma en la que lo hicieron.

Hay muchas cosas en las que trabajé que no tienen lugar en esta tesis, pero que han dejado su huella haciendo que ésta sea como es. Incluso construí un nuevo modelo donde cambiaba completamente la forma en la que sucedía el proceso de selección, porque me parecía que se acercaba más a

la realidad. Pronto me dí cuenta de lo complicado que es crear un nuevo modelo que pueda aportar algo más allá de su definición.

Aún después de darme por vencida con mi nuevo modelo de poblaciones con selección, no me convencía en lo más mínimo la idea de una 'selección multiplicativa', reduce muchísimo el tipo de poblaciones a las que se puede modelar. Al final entendí la necesidad de este requisito, y considero que es lo mejor que se puede hacer con este modelo.

Después encontré un punto en el que considero que se puede mejorar el Modelo de Wright-Fisher. Como leerán más adelante, en el Modelo de Wright-Fisher completo, el proceso de selección sucede antes que el de mutación, lo cual entra en conflicto con lo que yo esperaba de cualquier modelo de una población biológica. Después de hacer muchas cosas al respecto (como tratar de demostrar que para la cadena de Markov daba lo mismo el orden, lo cual es falso) llegué a una excelente solución, definir un nuevo modelo con el orden adecuado y demostrar que obtenemos la misma cadena de Markov que en el modelo usual si cambiamos los parámetros del usual por cierta función de las probabilidades de mutación y supervivencia de la población (las cuales se pueden estimar experimentalmente).

En el capítulo 2 me topé con la misma pared contra la que chocaron las personas a las que se les ocurrió aproximar a estas cadenas por medio de un proceso de difusión (a menos que el modelo haya sido inventado con la aproximación en mente desde el principio).

Intenté calcular la distribución estacionaria del Modelo de Wright-Fisher con 2 alelos con mutación y sin selección, sin llegar a ninguna conclusión.

También intenté utilizar los programas Máxima y Octave para obtener un poco de intuición acerca de cómo sería la fórmula general de las probabilidades de absorción.

Para el Modelo de Wright-Fisher con 2 alelos sin mutación ni selección, utilicé Octave. Modelé de forma numérica una población con 100 individuos, y el resultado saltó a mis ojos. A partir de ese experimento fue mucho más fácil calcular analíticamente las probabilidades de absorción. La verdad es que en ese momento estaba yo tan orgullosa de mi demostración, que hasta me enojé al encontrar inmediatamente después una prueba de estas probabilidades en Internet.

Para el Modelo de Wright-Fisher con  $r$  alelos con selección multiplicativa y sin mutación, diseñé una hoja de código de Máxima que calculaba en términos de los coeficientes de selección, las probabilidades de absorción al

estado (0, 0, 6) de una modesta población con 6 individuos y 3 tipos de alelos. El programa tardó más de cuatro horas en correr en mi computadora, y su respuesta fue

<< *Expression too long to display!* >> .

Hay una escena de una película que me encanta que simplemente NO puedo evitar citarla. La película es "Hitchhiker's Guide to the Galaxy".

Many many millions of years ago a race of hyperintelligent pandimensional beings got so fed up with the constant bickering about the meaning of life that they commissioned two of their brightest and best (named Lunkwill and Fook) to design and build a stupendous super computer (named Deep Thought) to calculate the answer to life the universe and everything.

Once built, they went ahead and asked:

"Deep Thought, we want you to tell us . . ." he paused, ". . . the Answer!"

"The answer?" said Deep Thought. "The answer to what?"

"Life!" urged Fook.

"The Universe!" said Lunkwill.

"Everything!" they said in chorus.

Deep Thought paused for a moment's reflection.

"Tricky," he said finally.

"But can you do it?"

Again, a significant pause.

"Yes," said Deep Thought, "I can do it."

"There is an answer?" said Fook with breathless excitement.

"A simple answer?" added Lunkwill.

"Yes," said Deep Thought. "Life, the Universe, and Everything. There is an answer. But," he added, "I'll have to think about it."

"Return in exactly seven and a half million years."

Siete millones y quinientos mil años después regresan Fook y Lunkwill rodeados por miles de individuos y le preguntan a Deep Thought:

"Deep Thought ... Do you..."

"Have an answer for you? Yes. But you're not going to like it." interrupted Deep Thought majestically.

"Please tell us. We must know!" begged Fook.

"Okay. The answer to the ultimate question of life, the universe, and everything is..."

" ... "

" ... 42."

Este fragmento expresa perfectamente mis sentimientos en ese momento. Era la cuarta vez que dejaba correr al programa, y creo que pasé más de una semana intentando que sirviera de algo. Al pedirle que me enseñara la expresión, aunque fuera muy larga, la computadora se volvió a trabar una hora, pero abrió un archivo con un polinomio inmenso, en 3 variables, y de un grado gigantesco; ya no recuerdo cuál fue el término más grande que encontré, pero no lo podía creer.

Así fue como me dí cuenta que las aproximaciones por medio de procesos de difusión eran absolutamente necesarias.

Como he dicho ya, estoy muy satisfecha con mi capítulo 3, pero me gustaría confesar que en algún momento odié profundamente aquellas tres páginas de las notas de Donald Dawson donde se encuentra la estructura del tercer capítulo. Y quiero decir que mi capítulo no sigue al pie de la letra los pasos planteados por Donald Dawson, que hay pasos que no seguí, con los que aún no estoy de acuerdo. Por otro lado, una vez que logré entender un poco más, también encontré cosas maravillosas en esas tres páginas, que hacen que la prueba se vuelva elegante, admirable y práctica; cosas que hacen que las demostraciones sean mucho mas sencillas, que hicieron que redujera 20 páginas de cuentas a tan sólo 2. También quiero decir que sin el apoyo de tantas personas, no habría capítulo 3 en esta tesis.

La gran diferencia entre los capítulos 1 y 2, y el capítulo 3, es que en los primeros podía dejar volar a mi imaginación para después entender mejor por qué el modelo de Wright-Fisher es como es, en cambio, el capítulo 3 fue una gran lucha para comprender algo que en un inicio estaba fuera de mi alcance.

Aprendí muchisisisimo en los tres.

# Introducción

En esta tesis consideramos el Modelo de Wright-Fisher en su versión discreta, así como la versión continua del modelo sin selección, sin mutación y con dos alelos.

Supongamos que tenemos una población en la cual hay varios tipos de individuos y las características que determinan el tipo del individuo son hereditarias. Podríamos suponer también que sobre la población actúa un proceso de selección, en el que individuos de algún tipo tengan una mayor probabilidad de sobrevivir que el resto de la población. También podría pasar que los hijos no siempre fueran del mismo tipo que sus padres, que los hijos mutaran del tipo heredado de sus padres a algún otro tipo con cierta probabilidad.

El modelo de Wright-Fisher modela este tipo de poblaciones, tomando en cuenta únicamente la cantidad de individuos que hay de cada tipo en cada generación a lo largo del tiempo. Con este modelo, dado el estado actual de la población (en cuestión de cantidad de individuos de cada tipo), se puede saber qué esperar de la población en futuras generaciones.

Por ejemplo, si en una población la probabilidad de que ocurra una mutación es cero, gracias a cálculos hechos con el modelo de Wright-Fisher, sabemos que eventualmente (en una cantidad de tiempo finito con probabilidad uno) toda la población será exactamente del mismo tipo. Conocemos con qué probabilidad la población entera se volverá de cada uno de los tipos, y cual es la esperanza del tiempo que tardará la población en tener un solo tipo de individuos.

Para que el modelo tuviera una expresión matemática sencilla y se pudieran encontrar resultados con él, se hicieron muchas suposiciones adicionales sobre la población a la que se modela.

En el **capítulo 1** explicamos cada una de las componentes del modelo de Wright-Fisher discreto, dividiéndolo en casos particulares del modelo general, como el caso en el que la probabilidad de mutar es cero, o el caso en el que no hay selección.

También dejaremos de una forma muy clara qué suposiciones se hacen acerca de la población modelada, y cómo es que a partir de ella se construye el modelo.

Algo importante que encontramos al desarrollar esta tesis es la razón por la que se pide comúnmente que los coeficientes de selección sean multiplicativos. Esta condición sobre los coeficientes de selección es necesaria para que el tipo de los dos alelos de un individuo sean variables aleatorias independientes, y por lo tanto, se simplifique mucho la expresión del modelo.

Otra cosa importante que haremos en el capítulo 1 será encontrar los parámetros de mutación y selección  $(\overline{\mu_{i,j}}, \overline{v_s})$  que se deben usar para modelar a una población en la que el proceso de mutación sucede antes que el de selección con el modelo usual de Wright-Fisher con mutación y selección multiplicativa. Es necesario hacer estos cálculos debido a que la forma en la que se construye el modelo usual de Wright-Fisher en libros como el de Ethier y Kurtz [?] y las notas de Dawson [D], supone que el proceso de selección sucede antes que el de mutación, al revés que en los modelos biológicos.

Si la probabilidad de un alelo de un individuo de mutar de un tipo  $\{i\}$  a un tipo  $\{j\}$  está dada por  $\mu_{i,j}$ , y la probabilidad de un individuo de tipo  $\{s, a\}$  de sobrevivir hasta la edad adulta es  $v_s \cdot v_a$ , entonces los parámetros de mutación y selección que se deben usar en el modelo usual de Wright-Fisher son:

$$\overline{v_s} = \frac{\sum_a v_a \cdot \mu_{s,a}}{\sum_t v_t \cdot \mu_{1,t}} \cdot \overline{v_1} \qquad \overline{\mu_{s,i}} = \frac{v_i \cdot \mu_{s,i}}{\sum_t v_t \cdot \mu_{s,t}}.$$

Para utilizar el modelo usual, lo que tendría que hacerse es estimar los parámetros de selección sin considerar las posibles mutaciones de la población; y estimar los parámetros de mutación tomando en cuenta únicamente a individuos que hayan sobrevivido el proceso de selección.

En el **capítulo 2** calcularemos directamente dos de las cantidades importantes de las cadenas de Wright-Fisher. Encontraremos las probabilidades de absorción de cada uno de los estados absorbentes de las cadenas de Wright-Fisher sin mutación ni selección con  $r$  alelos y con *dos* alelos. Es

decir, para cada tipo, calcularemos la probabilidad de que eventualmente todos los individuos de la población sean de ese tipo.

No haremos el cálculo el resto de las cantidades importantes debido a la dificultad que involucran.

En general es muy complicado calcular cualquier cosa en los modelos de Wright-Fisher con mutación o selección, por lo que comúnmente se utilizan aproximaciones por procesos de difusión; se calcula lo que se desea calcular para el proceso de difusión, y luego se demuestra que este cálculo es una buena aproximación.

En el libro de Ethier y Kurtz [?] capítulo 10, se desarrolla la aproximación de los modelos de Wright-Fisher por medio de procesos de difusión.

En la sección 2 del capítulo 10 del libro de Ethier y Kurtz [?] se calculan la esperanza del tiempo de absorción, la distribución estacionaria y las probabilidades de absorción de los procesos de difusión que aproximan a los modelos de Wright-Fisher, y se demuestra que tanto la sucesión de tiempos de absorción como la sucesión de distribuciones estacionarias de las cadenas de Markov correspondientes, convergen débilmente al tiempo de absorción y la distribución estacionaria de los procesos de difusión que las aproximan.

En el **capítulo 3** veremos cómo se aproxima al modelo de Wright-Fisher por medio de un proceso de difusión en el caso particular en el que sólo se trabaja con *dos* alelos y no hay ni selección ni mutación. Para este proceso de difusión, calculamos la esperanza del tiempo de absorción.

Para hacer esta aproximación, seguimos los pasos planteados por Donald Dawson en el libro [D], aunque en varios puntos nos desviamos del camino propuesto. A grandes rasgos, el procedimiento es el siguiente:

- Se construye una sucesión de procesos cadlåg (continuos por la derecha), donde para cada  $N \in \mathbb{N}$  el proceso cadlåg preserva la información del modelo de Wright-Fisher para una población de tamaño  $N$ , para este nuevo proceso, en cada unidad de tiempo transcurren  $N$  generaciones.
- Después utilizamos el novedoso criterio para tensión de Aldous (el cual se puede encontrar en la última edición del libro de Billingsley [B]) para demostrar que la sucesión de procesos cadlåg construida es una sucesión tensa, y por lo tanto, relativamente compacta. Esto implica que para cada subsucesión existe una subsucesión que converge débilmente.
- Calculamos el generador de los procesos límite de las subsucesiones

convergentes para casi todos los tiempos. Y encontramos que los generadores de todos los procesos límite son iguales al generador de un proceso de difusión para casi todos los tiempos (para todos, excepto un conjunto de medida de Lebesgue 0).

- Que el generador del proceso límite de una subsucesión convergente sea igual al generador de un proceso de difusión para casi todos los tiempos, es suficiente para poder afirmar que el proceso límite es una solución al Problema de Martingala del generador de la difusión.
- Demostramos que este Problema de Martingala (del generador de la difusión) satisface las hipótesis de un teorema del libro de Ethier y Kurtz [?] con el cual podremos afirmar que ese Problema de Martingala tiene una única solución, y por lo tanto, habremos demostrado que todas las subsucesiones convergen débilmente al mismo proceso de difusión, y así concluimos que la sucesión entera converge a ese proceso.

Finalizaremos el capítulo 3 calculando la esperanza del tiempo de absorción del proceso de difusión, la cual, de acuerdo al libro de Ethier y Kurtz [?], es una buena aproximación a la esperanza del tiempo de absorción del modelo de Wright-Fisher (sin mutación ni selección con *dos* alelos) si el tamaño de la población es muy grande.

Para desarrollar esta tesis, fue necesario adentrarse profundamente en el libro [?] “Markov Processes, Characterization and Convergence” de los autores Stewart N. Ethier y Thomas Kurtz. (edición de 1986).

También utilizaremos el libro [B] “Convergence of Probability Measures” del autor Patrick Billingsley (edición de 1999);

Un capítulo fue inspirado en las notas [D] “Preliminary Version Notes for Lectures on Stochastic Population Systems” del autor Donald Dawson les llamaremos \*Donald Dawson\*.

Usaremos también el libro [HPS] “Introduction to Stochastic Processes” de los autores Paul G. Hoel, Sidney C. Port y Charles J. Stone.





# Capítulo 1

## Modelos de Wright-Fisher Discretos

En el libro de Ethier y Kurtz [?] en el capítulo 10 sección 1, se define el modelo discreto de Wright-Fisher, como la cadena de Markov  $\{P^n\}_{\mathbb{N}}$  con espacio de estados {Matrices de  $r \times r$  triangulares cuyas entradas suman  $N$ }, y con probabilidades de transición dadas por

$$\mathbb{P}\left(P^{n+1} = A \mid P^n\right) = \mathbb{P}\left[\text{Multinomial}\left(N, \left(\text{mut}G_{i,j}^n\right)_{i,j}\right) = A\right]$$

$$\text{con } \text{mut}G_{\{i,j\}}^n = \sum_{k \leq l} \frac{1}{1 + \delta_{i,j}} \cdot (\mu_{k,i} \cdot \mu_{l,j} + \mu_{k,j} \cdot \mu_{l,i}) \cdot G_{k,l}^n$$

$$\text{con } G_{i,j}^n = \frac{(2 - \delta_{i,j}) \cdot \psi_{i,j} \cdot p_i^n \cdot p_j^n}{\sum_{k \leq h} (2 - \delta_{k,h}) \cdot \psi_{k,h} \cdot p_k^n \cdot p_h^n}$$

$$\text{y con } p_i^n = 2 \cdot P_{\{i,i\}}^n + \sum_{j \neq i} P_{\{i,j\}}^n.$$

donde  $\mu_{i,j}$  y  $\psi_{i,j}$  son los parámetros de mutación y selección.

Se define y explora el caso particular en el que los parámetros de selección actúan de forma multiplicativa ( $\psi_{i,j} = v_i \cdot v_j$ ).

En las notas de Dawson [D] también se define de esta forma el modelo de

Wright-Fisher, pero se introduce primero el modelo en el caso neutral (sin mutación ni selección) con dos alelos.

En esta tesis, introduciremos el modelo de forma gradual, separándolo paso por paso. Describiendo en cada paso la forma en que el modelo se basa en la población, las suposiciones que se hacen de ésta y la forma como actúa cada uno de los elementos en la supuesta población a la que modela.

Empezaremos por el caso neutral (sin mutación ni selección) con dos alelos, y pasaremos por todas las combinaciones importantes hasta llegar al modelo completo.

En el transcurso definiremos varios modelos que se pueden obtener del modelo de Wright-Fisher completo haciendo los parámetros de mutación o selección iguales a cero, pero que son cruciales para entender el funcionamiento de cada una de sus partes.

Al dividir al modelo en pasos, me dí cuenta de un punto que podría causar problemas al tratar de modelar poblaciones reales con este modelo. En las definiciones del modelo de Wright-Fisher del libro de Ethier y Kurtz [?] y de las notas de Dawson [D], el modelo supone que en la formación de los individuos de la siguiente generación, el proceso de selección sucede antes que el proceso de mutación, en contra de los modelos biológicos actuales. Esto es natural si al calcular los parámetros de mutación y selección se estiman los parámetros de selección sin considerar las posibles mutaciones de la población y después se estiman los parámetros de mutación tomando en cuenta únicamente a individuos que hayan sobrevivido el proceso de selección. Sin embargo, si se estiman los parámetros de mutación antes de que la población sufra el proceso de selección, y se calculan los parámetros de selección tomando en cuenta que algunos individuos nacieron mutantes, no se puede utilizar directamente el modelo de Wright-Fisher, se necesita hacer un cambio en los parámetros que explicaremos en la sección **1.7.2**.

Con lo explicado en el párrafo anterior en mente, en la subsubsección **1.7.2** definimos un nuevo modelo, el cual corresponde al modelo de Wright-Fisher con mutación y selección multiplicativa, pero donde la mutación ocurre antes que la selección; y demostramos que este nuevo modelo se puede representar a partir del modelo usual de Wright-Fisher con mutación y selección multiplicativa, (donde la selección ocurre antes que la mutación), usando un cambio en los parámetros:

Si en una población la mutación ocurre antes que la selección multiplicativa, la probabilidad de un alelo de mutar de tipo  $\{i\}$  a tipo  $\{j\}$  está dada por el parámetro  $\mu_{i,j}$ , y la probabilidad de un individuo de tipo  $\{i, j\}$  de sobrevivir hasta la edad adulta está dada por el parámetro  $\psi_{i,j} = v_i \cdot v_j$  (donde la probabilidad de sobrevivir se mide sobre individuos que no volverán a mutar), entonces los parámetros que se deberían utilizar en el modelo usual de Wright-Fisher con mutación y selección multiplicativa, (donde la selección ocurre antes que la mutación) son  $\{\overline{\mu_{i,j}}\}_{i,j}$  y  $\{\overline{v_s}\}_s$  dados por

$$\overline{v_s} = \frac{\sum_a v_a \cdot \mu_{s,a}}{\sum_t v_t \cdot \mu_{1,t}} \cdot \overline{v_1} \qquad \overline{\mu_{s,i}} = \frac{v_i \cdot \mu_{s,i}}{\sum_t v_t \cdot \mu_{s,t}}.$$

El modelo usual desarrollado en el libro de Ethier y Kurtz [?], donde la selección multiplicativa ocurre antes que la mutación, es descrito en esta tesis en la sección **1.7.2**.

Otro aspecto importante que se desarrolla en esta tesis, son las razones por las que se pide que la selección sea multiplicativa.

Empezaremos en la sección **1.1** presentando el modelo más sencillo, el modelo de Wright-Fisher haploide con 2 tipos de alelos, sin mutación y sin selección.

En la sección **1.2** mantendremos la neutralidad en la mutación y selección, pero aumentaremos el número de tipos de alelos posibles. Presentaremos el modelo de Wright-Fisher haploide con  $r$  tipos de alelos, sin mutación y sin selección.

En la sección **1.3** regresaremos a considerar únicamente 2 tipos de alelos, pero introducimos el factor de la mutación de un tipo de alelo a otro.

En la sección **1.4** cambiaremos la naturaleza haploide de la población estudiada. Obtendremos el modelo de Wright-Fisher que representa a poblaciones diploides que se reproducen de manera sexual, manteniendo la neutralidad en la mutación y selección. Lo que encontraremos es que este modelo de Wright-Fisher diploide con  $r$  alelos, sin mutación y sin selección, es equivalente al modelo de la sección **1.2**, pero del doble del tamaño.

En la sección **1.5** agregaremos a la población diploide el factor de selección. Presentaremos el Modelo Diploide con  $r$  Alelos y Selección.

En la sección 1.6 consideramos poblaciones diploides con cierto tipo de selección al que llamaremos selección multiplicativa, la cual tiene la ventaja de hacer que la cadena que representa a la cantidad de alelos de cada tipo en la generación tenga probabilidades de transición muy sencillas. Obtenemos el Modelo Diploide con r Alelos y Selección Multiplicativa.

En la sección 1.7 unimos todos los elementos, modelamos una población diploide con mutación y selección. En la primera subsección modelamos a la población con la mutación general, en la segunda subsección modelamos a la población con la mutación multiplicativa, en ésta subsección consideraremos primero el caso en el que la selección sucede antes que la mutación (el caso que se desarrolla comúnmente en la literatura), después consideramos el caso en el que la mutación sucede antes que la selección, y al final encontramos la forma de representar al modelo en el que la mutación sucede antes que la selección a través del primer modelo (con la selección antes que la mutación) con los parámetros adecuados.

Para desarrollar este capítulo, nos basamos en el libro [?] “Markov Processes, Characterization and Convergence” de los autores Stewart N. Ethier y Thomas Kurtz. (edición de 1986).

## 1.1. Modelo Haploide con 2 Alelos

**Definición 1.1.1.** *El modelo discreto de Wright-Fisher haploide, con dos tipos de alelos, sin mutación ni selección, es una cadena de Markov discreta  $\{Y_n\}_{n \in \mathbb{N}}$  con espacio de estados  $\{0, \dots, N\}$ , y probabilidades de transición dadas por*

$$\mathbb{P}\left(Y_{n+1} = j \mid Y_n = i\right) = \binom{N}{j} \left(\frac{i}{N}\right)^j \left(1 - \frac{i}{N}\right)^{N-j}, \quad j = 0, \dots, N.$$

Esta cadena de Markov modela una población que se encuentra dividida en dos tipos por alguna característica hereditaria, donde los hijos siempre son del mismo tipo que los padres y donde lo que más nos interesa es cómo fluctúa de generación en generación la proporción de la población que es de uno de los dos tipos. El modelo se inspira en poblaciones de genes, por lo que hablaremos de tipos de alelos. Un requisito de este modelo es que la población mantenga un tamaño constante  $N$  a lo largo del tiempo. Suponer

esto de la población es algo natural si lo que nos interesa es la proporción en la que se distribuyen los tipos.

Construyamos la cadena de Wright-Fisher de este modelo a partir de las poblaciones a las que modela.

La población de la cual vamos a extraer el modelo de Wright-Fisher de esta sección, es una población que empieza con  $N$  individuos divididos en dos tipos, tipo  $\alpha$  y tipo  $\beta$ , que se reproducen de manera mitótica produciendo hijos del mismo tipo al que pertenecen y teniendo todos la misma habilidad para reproducirse (como una célula que duplica su material genético y simplemente se divide en dos, generando otra célula idéntica a ella, del mismo tipo).

Los  $N$  individuos con los que empezamos conforman la generación 1. De la generación 1 nacen exactamente  $N$  hijos, los cuales conforman la generación 2. De la generación 2 nacen exactamente  $N$  hijos que conforman la generación 3, y así se construyen el resto de las generaciones. Además, los individuos de una generación no interactúan con individuos de otras generaciones, digamos que tienen hijos que conforman la siguiente generación y se mueren. Para tener una notación clara, a los individuos de la generación  $n$  les llamamos "n-individuos".

Sea  $Y_n =$  el número de n-individuos de tipo  $\alpha$ .

Tenemos una población que evoluciona de generación en generación, donde empezamos con  $Y_1$  cantidad de 1-individuos de tipo  $\alpha$  que siempre tienen hijos de tipo  $\alpha$ , tenemos  $N - Y_1$  individuos de tipo  $\beta$ , que siempre tienen hijos de tipo  $\beta$ , y sabemos que la suma del número de hijos de todos los individuos es exactamente  $N$ .  $\{Y_n\}_{n \in \mathbb{N}}$  conforma una cadena de la cual queremos calcular las probabilidades de transición.

En algunas poblaciones, el hecho de que un individuo sea de tipo  $\alpha$  provoca que tenga una mayor o menor habilidad para reproducirse. Pero en este modelo, el tipo del individuo no afecta su habilidad para reproducirse.

La cantidad de hijos que puede tener un individuo, y la probabilidad con la que los tiene, varía ampliamente de una población a otra. La forma clásica y mundialmente aceptada para modelar la cantidad de hijos que tiene un individuo es a través de la distribución  $Poisson(\lambda)$  para la  $\lambda$  adecuada. Modelaremos a nuestra población con ésta distribución, sin olvidar que nuestra población tiene además la restricción de que la suma de los hijos por generación es  $N$ .

Supongamos que para alguna  $\lambda$  fija, la cantidad de hijos de cada n-individuo se distribuye  $Poisson(\lambda)$  condicionado a que la suma de todas las variables  $Poisson(\lambda)$  de los n-individuos es  $N$ .

Para explicar con mayor detalle el enunciado anterior, tomemos una numeración de los n-individuos. Sean  $I_j$  variables aleatorias  $Poisson(\lambda)$  independientes entre sí, para todo  $j$  entre 1 y  $N$ ,  $I_j \sim Poisson(\lambda)$  y sea  $W$  la suma de estas variables,  $W = \sum_{j=1}^N I_j$ .

El número de hijos del n-individuo  $j$  es la variable aleatoria  $I_j$  condicionado a que  $W = \sum_{j=1}^N I_j = N$ ,

$$Z_j = I_j \Big|_{W=N}.$$

Notemos que los  $Z_j$  no son independientes entre sí, aunque los  $I_j$  si lo sean. La distribución de  $Z_j$  (el número de hijos del n-individuo  $j$ ) está dada por

$$\mathbb{P}(Z_j = k) = \mathbb{P}\left(I_j = k \mid \sum_{i=1}^N I_i = N\right) = \frac{\mathbb{P}\left(I_j = k \cap \sum_{i \neq j} I_i = N - k\right)}{\mathbb{P}\left(\sum_{i=1}^N I_i = N\right)}$$

$$0 \leq k \leq N$$

Recordemos que  $\sum_{i=1}^N I_i$  es una suma de  $N$  variables aleatorias  $Poisson(\lambda)$ , y por lo tanto se distribuye de manera  $Poisson(N\lambda)$ . Recordemos también que los  $I_i$  son independientes entre sí.

$$\begin{aligned} \mathbb{P}(Z_j = k) &= \frac{\mathbb{P}(I_j = k) \cdot \mathbb{P}\left(\sum_{i \neq j} I_i = N - k\right)}{\mathbb{P}\left(Poisson(N \cdot \lambda) = N\right)} \\ &= \frac{e^{-\lambda} \cdot (\lambda)^k}{k!} \cdot \frac{e^{-(N-1)\lambda} \cdot \left((N-1) \cdot \lambda\right)^{N-k}}{(N-k)!} \cdot \frac{N!}{(N \cdot \lambda)^N e^{-N \cdot \lambda}} \\ &= \binom{N}{k} \left(\frac{1}{N}\right)^k \left(\frac{N-1}{N}\right)^{N-k} \end{aligned}$$

Lo cual significa que  $Z_j \sim Binomial(N, \frac{1}{N})$  para cada n-individuo  $j$ .

Ahora, como cada  $(n+1)$ -individuo de tipo  $\alpha$  es hijo de un  $n$ -individuo de tipo  $\alpha$  :

$$Y_{n+1} = \sum_{\substack{j \text{ tales que} \\ \text{el individuo } j \\ \text{es de tipo } \alpha}} Z_j$$

Hay que recordar que  $Z_i$  no es independiente de  $Z_a$ . Pero podemos calcular la distribución de  $Y_{n+1}$  de una manera análoga a la forma en que calculamos la distribución de  $Z_j$ :

$$\mathbb{P}(Y_{n+1} = k \mid Y_n = m) = \mathbb{P} \left( \sum_{\substack{j \text{ tales que} \\ \text{el individuo } j \\ \text{es de tipo } \alpha}} Z_j = k \mid \begin{array}{l} \text{hay exactamente } m \\ \text{n-individuos de tipo } \alpha \end{array} \right)$$

Para facilitar la notación, supongamos que los  $m$   $n$ -individuos de tipo  $\alpha$ , son los primeros  $m$  en la enumeración de los  $n$ -individuos.

$$\begin{aligned} \mathbb{P} \left( \sum_{j=1}^m Z_j = k \right) &= \mathbb{P} \left( \sum_{j=1}^m I_j = k \mid \sum_{i=1}^N I_i = N \right) \\ &= \mathbb{P} \left( \sum_{j=1}^m I_j = k \cap \sum_{i=m+1}^N I_i = N - k \right) \cdot \frac{1}{\mathbb{P} \left( \sum_{i=1}^N I_i = N \right)} \\ &= \mathbb{P} \left( \sum_{j=1}^m I_j = k \right) \cdot \mathbb{P} \left( \sum_{i=m+1}^N I_i = N - k \right) \cdot \frac{N!}{(N \cdot \lambda)^N e^{-N \cdot \lambda}} \\ &= \frac{e^{-m\lambda} \cdot (m\lambda)^k}{k!} \cdot \frac{e^{-(N-m)\lambda} \cdot ((N-m) \cdot \lambda)^{N-k}}{(N-k)!} \cdot \frac{N!}{(N \cdot \lambda)^N e^{-N \cdot \lambda}} \\ &= \binom{N}{k} \left( \frac{m}{N} \right)^k \left( \frac{N-m}{N} \right)^{N-k} \\ &= \binom{N}{k} \left( \frac{m}{N} \right)^k \left( 1 - \frac{m}{N} \right)^{N-k} \end{aligned}$$

Lo cual implica que :

$$\mathbb{P}(Y_{n+1} = k \mid Y_n) = \mathbb{P}\left(\sum_{j=1}^{Y_n} Z_j = k \mid Y_n\right) = \binom{N}{j} \left(\frac{Y_n}{N}\right)^k \left(1 - \frac{Y_n}{N}\right)^{N-k}$$

( si  $0 \leq k \leq N, k \in \mathbb{N}$  ).

Hemos obtenido las probabilidades de transición de la cadena  $\{Y_n\}_{n \in \mathbb{N}}$ .

Como podemos observar, estas probabilidades de transición son las de una cadena de Markov homogénea. Su espacio de estados es  $\{0, \dots, N\}$  y es idéntica a la cadena definida en la **Definición 1.1.1**.

También se pueden interpretar estas probabilidades de transición como las de una cadena que modela a una población donde cada  $(n+1)$ -individuo es hijo de cualquiera de los  $n$ -individuos con la misma probabilidad, en lugar de modelar la cantidad de hijos que tiene cada  $n$ -individuo con una variable  $Poisson(\lambda)$ .

La probabilidad de cualquier  $(n+1)$ -individuo de ser de tipo  $\alpha$  en este modelo es  $\frac{Y_n}{N}$ , y la de ser de tipo  $\beta$  es  $(1 - \frac{Y_n}{N})$ , con lo cual tenemos las mismas probabilidades de transición.

En general, podríamos haber construido una población como la que acabamos de construir, donde el  $n$ -individuo 's' tienen una cantidad aleatoria de hijos  $I_s$  distribuida casi como queramos, condicionada a que  $\sum_{s=1}^N I_s = N$ , y construir de esta forma todo tipo de modelos, muy probablemente distintos del Wright-Fisher, que no estudiaremos en esta tesis.

Un ejemplo:  $I_s = \text{Uniforme } \{0, N\}$ , en este modelo trivial, cada  $n$ -individuo es padre de todos los  $(n+1)$ -individuos o de ninguno, y a partir de la generación 2, todos los individuos son claramente del mismo tipo.

## 1.2. Modelo Haploide con r Alelos

El nuevo modelo que presentaremos será muy parecido al modelo anterior, con la diferencia de que ahora consideraremos  $r$  tipos de individuos. Recurriremos a  $\mathbb{N}^r$ , cada  $r$ -ada representará la composición de la población (en cuestión de los  $r$  tipos), cada entrada del vector representa el número de individuos de cada tipo en la generación a la que representa la  $r$ -ada.

La población que modelaremos en esta sección se reproduce mitóticamente, y cada  $(n+1)$ -individuo es hijo de cualquiera de los  $n$ -individuos con la misma probabilidad. Podríamos tomar una población que tiene una cantidad de hijos  $Poisson(\lambda)$  condicionada a que la cantidad total de hijos es  $N$  y llegaríamos al mismo modelo, como sucedió en el modelo de la sección anterior.

**Definición 1.2.1.** *El modelo discreto de Wright-Fisher haploide, con  $r$  tipos de alelos, sin mutación ni selección, es una cadena de Markov discreta  $\{X_n\}_{n \in \mathbb{N}}$  con espacio de estados  $\{(a_1, \dots, a_r) \in \mathbb{N}^r \mid \sum_{i=1}^r a_i = N\}$ , y probabilidades de transición dadas por:*

$$\mathbb{P}\left(X_{n+1} = (a_1, \dots, a_r) \mid X_n = (b_1, \dots, b_r)\right) = \binom{N}{a_1, \dots, a_r} \left(\frac{b_1}{N}\right)^{a_1} \cdots \left(\frac{b_r}{N}\right)^{a_r}$$

Construyamos este modelo a partir de la población que queremos modelar:

Tomemos una población que tenga  $N$  individuos en cada generación, donde:

- cada individuo sea de uno sólo de los  $r$  tipos que puede haber en la población,
- los hijos sean del mismo tipo que los padres,
- cada  $(n+1)$ -individuo sea hijo de cualquiera de los  $n$ -individuos con la misma probabilidad y de forma independiente al tipo del resto de los  $n$ -individuos.

Sea  $X_n = (x_1^n, x_2^n, \dots, x_r^n) \in \mathbb{N}^r$  donde  $x_i^n$  es el número de  $n$ -individuos de tipo  $i$ . Por como es la población, tenemos que  $\sum_{i=1}^r x_i^n = N$ .

Tenemos de esta forma una cadena con espacio de estados

$$\left\{ (a_1, \dots, a_r) \in \mathbb{N}^r \mid \sum_{i=1}^r a_i = N \right\}.$$

Calculemos  $\mathbb{P}\left(X_{n+1} = (a_1, \dots, a_r) \mid X_n = (b_1, \dots, b_r)\right)$ :

Si  $X_n = (b_1, \dots, b_r)$ , entonces hay  $b_i$   $n$ -individuos de tipo  $i$ , por lo que hay una probabilidad de  $\frac{b_i}{N}$  de que cierto  $(n+1)$ -individuo sea de tipo  $i$ . Para que  $X_{n+1} = (a_1, \dots, a_r)$  tiene que suceder que  $a_i$   $(n+1)$ -individuos sean de tipo  $i$  ( $\forall i$ ). Hay  $\binom{N}{a_1, \dots, a_r}$  formas de dividir al conjunto de  $(n+1)$ -individuos en  $r$  subconjuntos, un subconjunto con  $a_i$   $(n+1)$ -individuos para cada  $a_i$ . Por cada una de estas formas de tomar los subconjuntos, hay una probabilidad de  $\left(\frac{b_1}{N}\right)^{a_1} \cdots \left(\frac{b_r}{N}\right)^{a_r}$  de que los individuos del subconjunto con  $a_i$  elementos sean del tipo  $i$  ( $\forall i$ ). Por lo que tenemos que:

$$\mathbb{P}\left(X_{n+1} = (a_1, \dots, a_r) \mid X_n = (b_1, \dots, b_r)\right) = \binom{N}{a_1, \dots, a_r} \left(\frac{b_1}{N}\right)^{a_1} \cdots \left(\frac{b_r}{N}\right)^{a_r}$$

Siempre y cuando  $\sum_{i=1}^r a_i = N$ .

Podemos ver que la variable aleatoria  $X_{n+1}$  dado  $X_n$ , es decir  $X_{n+1}|X_n$ , se distribuye como una variable aleatoria  $Multinomial(x_1^n, x_2^n, \dots, x_r^n)$ .

Esta cadena es claramente una cadena de Markov homogénea, con la misma distribución que la cadena de la **Definición 1.2.1**.

### 1.3. Modelo Haploide con 2 Alelos y Mutación

Podemos también modelar poblaciones donde no todos los hijos son del mismo tipo que los padres.

Empezaremos a hablar más de alelos. En esta sección, modelaremos una población donde los alelos que determinan el tipo del individuo mutan a otro alelo con cierta probabilidad.

Tomemos una población

- dividida en dos tipos, tipo *uno* ‘1’ y tipo *dos* ‘2’
- que se reproduzca mitóticamente,
- donde cada  $(n+1)$ -individuo sea hijo de cualquiera de los  $n$ -individuos con la misma probabilidad,
- y haya  $N$  individuos en cada generación.
- En lugar de suponer que los hijos son siempre del mismo tipo que sus padres, suponemos que los hijos pueden nacer de un tipo distinto al de

sus padres, y que esto sucede con probabilidades dadas por constantes  $\mu_{1,2}$  y  $\mu_{2,1}$ . La constante  $\mu_{1,2}$  corresponde a la probabilidad de que un  $(n+1)$ -individuo hijo de un  $n$ -individuo de tipo *uno* nazca de tipo *dos*. De forma análoga, la constante  $\mu_{2,1}$  corresponde a la probabilidad de que un  $(n+1)$ -individuo hijo de un  $n$ -individuo de tipo *dos* nazca de tipo *uno*. Estas constantes son generalmente muy pequeñas. Para facilitar la notación, diremos que un individuo **muta** de tipo  $i$  a tipo  $j$  cuando sea de tipo  $j$  y sea hijo de un individuo de tipo  $i$ .

Ahora la probabilidad de que un individuo sea del mismo tipo que su padre es  $(1 - \mu_{i,j})$ ,  $i, j$  dependiendo del tipo del padre.

Sea  $M_n =$  número de  $n$ -individuos de tipo *uno*.

La probabilidad de que un  $(n+1)$ -individuo sea de tipo *uno* dado que hay  $M_n$   $n$ -individuos de tipo *uno* es:

$$p_n^* = (1 - \mu_{1,2}) \cdot \frac{M_n}{N} + \mu_{2,1} \cdot \left(1 - \frac{M_n}{N}\right). \quad (1.1)$$

El primer término representa la probabilidad de que el padre del individuo sea de tipo *uno* y que el individuo no mute al tipo *dos*; el segundo término representa la probabilidad de que el padre sea de tipo *dos* y el individuo mute del tipo *dos* al tipo *uno*.

La probabilidad de que exactamente  $k$  de los  $(n+1)$ -individuos sean de tipo *uno*, conociendo cuantos  $n$ -individuos son de tipo *uno*, es una variable aleatoria  $Binomial(N, p_n^*)$ .

$$\mathbb{P}\left(M_{n+1} = k \mid M_n\right) = \binom{N}{k} (p_n^*)^k \cdot (1 - p_n^*)^{N-k}. \quad (1.2)$$

Sustituyendo  $p_n^*$  tenemos que:

$$\begin{aligned} \mathbb{P}\left(M_{n+1} = k \mid M_n\right) &= \binom{N}{k} \left( (1 - \mu_{1,2}) \frac{M_n}{N} + \mu_{2,1} \left(1 - \frac{M_n}{N}\right) \right)^k \\ &\quad \cdot \left[ 1 - \left( (1 - \mu_{1,2}) \frac{M_n}{N} + \mu_{2,1} \left(1 - \frac{M_n}{N}\right) \right) \right]^{N-k} \\ &= \binom{N}{k} \left( (1 - \mu_{1,2}) \cdot \frac{M_n}{N} + \mu_{2,1} \cdot \left(1 - \frac{M_n}{N}\right) \right)^k \\ &\quad \cdot \left[ (1 - \mu_{2,1}) \cdot \left(1 - \frac{M_n}{N}\right) + \mu_{1,2} \cdot \frac{M_n}{N} \right]^{N-k}. \end{aligned}$$

La igualdad entre los dos términos entre paréntesis [ ] se puede calcular directamente, o deducir del hecho de que  $\mathbb{P}(\text{un individuo sea de tipo uno}) = 1 - \mathbb{P}(\text{un individuo sea de tipo dos})$ .

Como podemos ver,  $\mathbb{P}\left(M_{n+1} = k \mid M_n\right)$  depende únicamente de  $M_n$  una vez que los coeficientes de mutación ( $\mu_{1,2}, \mu_{2,1}$ ) están fijos, por lo que se cumple la propiedad de Markov, y entonces la cadena  $\{M_n\}$  es una cadena de Markov. El espacio de estados de  $\{M_n\}$  es  $\{0, 1, 2, \dots, N\}$  y las probabilidades de transición están determinadas por (1.2).

**Definición 1.3.1.** *El modelo discreto de Wright-Fisher haploide, con dos tipos de alelos, con mutación y sin selección, es una cadena de Markov discreta  $\{M_n\}_{n \in \mathbb{N}}$  con espacio de estados  $\{0, \dots, N\}$ , y probabilidades de transición dadas por*

$$\mathbb{P}\left(M_{n+1} = k \mid M_n\right) = \binom{N}{k} (p_n^*)^k \cdot (1 - p_n^*)^{N-k}$$

$$\text{con } p_n^* = (1 - \mu_{1,2}) \cdot \frac{M_n}{N} + \mu_{2,1} \cdot \left(1 - \frac{M_n}{N}\right).$$

## 1.4. Modelo Diploide con r Alelos

Ahora vamos a exponer la construcción de un modelo para poblaciones diploides que se reproducen de manera sexual. Poblaciones con una cantidad fija de individuos, sin mutación ni selección y con  $r$  posibles alelos en el locus que nos interesa.

El proceso de reproducción sexual de organismos diploides comienza en la meiosis de las células reproductivas primarias. Estas son células diploide que tiene  $k$  pares de cromosomas homólogos. Homólogos se refiere a que ocupan el mismo lugar en el código genético. En el caso de los humanos  $k = 23$ . De cada par de cromosomas homólogos, uno proviene del padre del individuo y otro de la madre.

El proceso de meiosis consiste en varios pasos:

- Dentro de la célula reproductiva primaria, se agrupan los cromosomas homólogos.
- Cada cromosoma se duplica, y por lo tanto, quedan 4 cromosomas homólogos de cada tipo.
- Se produce el entrecruzamiento, donde los cromosomas se mezclan con sus cromosomas homólogos produciendo, de cada tipo, 4 cromosomas distintos a los originales y distintos entre sí.
- Se forman 4 células llamadas gametos, cada una con un cromosoma de cada tipo. Notemos que los gametos son células haploides (tienen un sólo cromosoma de cada tipo) en lugar de ser diploides (dos cromosomas de cada tipo). Como las células reproductivas primarias de los humanos.

En el siguiente paso, un gameto de un individuo fecunda a un gameto de otro individuo, formando una célula diploide llamada cigoto, a partir de la cual se forma otro individuo.

En esta sección explicaremos la construcción de un modelo para poblaciones diploides que se reproducen de manera sexual, donde la característica que nos interesa, (que divide a la población en varios tipos), es una característica provocada por el tipo de alelo que tiene el individuo en cierto “locus” de su genoma. Supondremos que hay únicamente  $r$  tipos de alelos en este “locus” y que en esta población no hay mutación ni selección.

Tomemos una población con  $N$  individuos. Cada individuo tiene un par no ordenado de alelos  $\{i, j\}$  en el locus que nos interesa. Uno aportado por la madre y otro por el padre.

Recordemos que hay  $r$  posibles alelos en el locus.

Cada  $n$ -individuo con alelos  $\{i, j\}$  puede producir, en cualquier momento, a través de la meiosis, un gameto con un alelo de tipo  $\{i\}$  y otro gameto con un alelo de tipo  $\{j\}$ .

Los  $(n+1)$ -individuos son formados a partir de dos gametos generados uniformemente por los  $n$ -individuos de la siguiente forma:

- Se escoge uniformemente y con reemplazo un  $n$ -individuo para que produzca dos gametos, un gameto por cada uno de sus dos alelos.
- Se escoge uniformemente uno de estos dos gametos y se desecha el otro.
- De forma independiente, se escoge otra vez uniformemente y con reemplazo un  $n$ -individuo para que produzca sus dos gametos y también se escoge uniformemente uno de sus dos gametos y se desecha el otro. Debido a que los escogemos con reemplazo, este segundo  $n$ -individuo puede ser el mismo que el anterior.
- Con ambos gametos escogidos uniformemente, se forma el nuevo  $(n+1)$ -individuo.

Se crean  $N$   $(n+1)$ -individuos de esta forma, y de manera independiente.

Este proceso con el que formamos a los  $(n+1)$ -individuos también se puede representar por medio de urnas. Para crear a un  $(n+1)$ -individuo:

- Se forman dos urnas. En cada urna, cada  $n$ -individuo de tipo  $\{i, j\}$  deposita dos gametos, un gameto tipo  $\{i\}$  y otro gameto tipo  $\{j\}$ .
- Se escoge uniformemente un gameto de cada urna.
- Se unen los dos gametos escogidos formando al cigoto del cual se desarrolla el  $(n+1)$ -individuo.

Usualmente los dos gametos provienen de dos  $n$ -individuos de sexos distintos, pero necesitamos que el modelo sea sencillo para poder obtener información de él, por lo que no distinguiremos ni el sexo, ni cuál es el individuo del cuál proviene el gameto. Únicamente tomaremos en cuenta el tipo de alelo. Incluso podrían ocurrir clonaciones. Se puede hacer esta simplificación ya que si solo tomamos en cuenta una generación si y otra no, saltándonos las generaciones de en medio, si puede ocurrir que un  $(n+1)$ -individuo haya heredado ambos alelos de un mismo  $(n-1)$ -individuo.

Entonces tenemos  $2N$  urnas con la misma composición, 2 urnas por cada uno de los  $(n+1)$ -individuos que se van a formar, y cada urna con los dos gametos de los  $n$ -individuos. Vamos a representar a la composición de estas urnas por el vector  $(q_1^n, q_2^n, \dots, q_r^n)$ , donde  $q_i^n$  = número de gametos en la urna de tipo  $\{i\}$ . A este vector le llamaremos la “ $n$ -urna”. También sabemos que hay  $\sum_{i=1}^r q_i^n = 2N$  gametos en la  $n$ -urna, porque cada  $n$ -individuo aportó 2 gametos a ésta. Este vector (la  $n$ -urna) está determinada por la composición de la generación  $n$  de la siguiente forma:

Sea  $P_{\{i,j\}}^n$  = número de  $n$ -individuos que tienen alelos  $\{i, j\}$ .

Recordando que  $\{i, j\} = \{j, i\}$ , y por lo tanto  $P_{\{i,j\}}^n = P_{\{j,i\}}^n$ , tenemos que el número de gametos de tipo  $\{i\}$  en la  $n$ -urna es:

$$q_i^n = 2 \cdot P_{\{i,i\}}^n + \sum_{j \neq i} P_{\{i,j\}}^n, \quad (1.3)$$

ya que cada individuo de tipo  $\{i, i\}$  aporta 2 gametos de tipo  $\{i\}$  a la urna y cada individuo de tipo  $\{i, j\}$  aporta exactamente un gameto de tipo  $\{i\}$  para cualquier  $j$ .

Entonces la probabilidad de tomar un gameto tipo  $\{h\}$  de la  $n$ -urna es  $\frac{q_h^n}{2N}$ . La probabilidad de formar un individuo de tipo  $\{h, h\}$  es  $\frac{q_h^n}{2N} \cdot \frac{q_h^n}{2N}$ . Y debido a que el tipo  $\{h, k\}$  es el mismo tipo que  $\{k, h\}$ , es decir, como no importa qué gameto se escoja primero, la probabilidad de formar un  $(n+1)$ -individuo de tipo  $\{h, k\}$  con  $h \neq k$  a partir de las  $n$ -urnas es:

$$\frac{q_h^n}{2N} \cdot \frac{q_k^n}{2N} + \frac{q_k^n}{2N} \cdot \frac{q_h^n}{2N} = 2 \cdot \frac{q_h^n}{2N} \cdot \frac{q_k^n}{2N} \quad (1.4)$$

La composición de la población en la generación  $n$  puede ser representada por la matriz triangular, de dimensión  $r \times r$ , con entradas en  $\mathbb{N}$  que suman  $2N$  :

$$\begin{pmatrix} P_{\{1,1\}}^n & P_{\{1,2\}}^n & P_{\{1,3\}}^n & P_{\{1,4\}}^n & \cdots & P_{\{1,r\}}^n \\ & P_{\{2,2\}}^n & P_{\{2,3\}}^n & P_{\{2,4\}}^n & \cdots & P_{\{2,r\}}^n \\ & & P_{\{3,3\}}^n & P_{\{3,4\}}^n & \cdots & P_{\{3,r\}}^n \\ & & & P_{\{4,4\}}^n & \cdots & P_{\{4,r\}}^n \\ & & & & \ddots & \vdots \\ & & & & & P_{\{r,r\}}^n \end{pmatrix} \quad (1.5)$$

A esta matriz le corresponde, a través de la ecuación (1.3), la n-urna  $(q_1^n, q_2^n, \dots, q_r^n)$ , que representa a las  $2N$  urnas de gametos generadas por los n-individuos de la matriz.

A partir de las urnas representadas por la n-urna  $(q_1^n, q_2^n, \dots, q_r^n)$ , se construye aleatoriamente la generación (n+1), utilizando las probabilidades descritas en ( 1.4 ) y su párrafo anterior. Esta generación puede ser representada por la matriz triangular:

$$\begin{pmatrix} P_{\{1,1\}}^{n+1} & P_{\{1,2\}}^{n+1} & \cdots & P_{\{1,r\}}^{n+1} \\ & P_{\{2,2\}}^{n+1} & \cdots & P_{\{2,r\}}^{n+1} \\ & & \cdots & P_{\{3,r\}}^{n+1} \\ & & & \vdots \\ & & & P_{\{r,r\}}^{n+1} \end{pmatrix} \quad (1.6)$$

A la cual le corresponde la (n+1)-urna de gametos  $(q_1^{n+1}, q_2^{n+1}, \dots, q_r^{n+1})$ . Y así sucesivamente.

En esta sección nos fijaremos únicamente en las n-urnas, y consideraremos a las matrices triangulares ( i.e. la configuración particular de cada una de las generaciones) como parte del proceso a partir del cual evolucionan las urnas. Podemos hacer esto porque las poblaciones de esta sección no sufren ningún proceso de selección ni mutación, y por lo tanto, calcular la distribución de las (n+1)-urnas a partir de las n-urnas resulta ser muy sencillo.

Observemos cómo la configuración de los individuos de las matrices (1.5) y (1.6) no juega ningún papel importante:

Al formar un (n+1)-individuo, lo que hacemos es tomar aleatoriamente de la n-urna dos gametos. Al formar la (n+1)-urna, el individuo lo que hace es depositar dos gametos idénticos a los que tomó de la n-urna. A final de cuentas, la única elección aleatoria que se hizo fue la elección de qué gametos conforman al (n+1)-individuo, y una vez escogidos, unos idénticos a él van a parar a la (n+1)-urna. Entonces, **en este modelo** podemos escoger uniformemente los gametos que conforman la (n+1)-urna directamente de la n-urna, y llegar al mismo resultado.

Denotemos por  $Q_n$  a la n-urna  $(q_1^n, q_2^n, \dots, q_r^n)$ .

Por lo visto en el párrafo anterior, cada gameto de la (n+1)-urna se escoge uniformemente de la n-urna, y por lo tanto, su probabilidad de ser de

tipo  $\{i\}$  es  $\left(\frac{q_i^n}{2N}\right)$ . Así es que las probabilidades de transición de la  $n$ -urna a la  $(n+1)$ -urna están dadas por:

$$\mathbb{P}\left(Q_{n+1} = (e_1, \dots, e_r) \mid Q_n = (q_1^n, \dots, q_r^n)\right) = \binom{2N}{e_1, \dots, e_r} \left(\frac{q_1^n}{2N}\right)^{e_1} \cdots \left(\frac{q_r^n}{2N}\right)^{e_r}$$

Siempre que  $\sum_i e_i = 2N$ .

Estas probabilidades de transición son las de una cadena de Markov homogénea. Y su espacio de estados es:

$$\left\{ (e_1, \dots, e_r) \in \mathbb{N}^r \mid \sum_{i=1}^r e_i = 2N \right\}$$

Como podemos observar, esta cadena es igual a la de la **sección 1.2**, con una población de tamaño  $2N$ .

En la siguiente sección construiremos un modelo como este, pero agregando un proceso de selección natural. Una de las principales consecuencias será que el paso de una urna a la siguiente dependerá de la forma en que los gametos se agrupan (en pares) formando individuos y no será un paso equivalente a escoger los gametos de la  $(n+1)$ -urna directamente de la  $n$ -urna.

## 1.5. Modelo Diploide con $r$ Alelos y Selección

Este modelo es muy parecido al de la sección 1.4, la diferencia es que agregamos el factor de la selección.

Será un modelo para poblaciones diploides que sufren un proceso de selección aleatorio que depende únicamente de los tipos de los individuos y toma lugar antes de la etapa de reproducción del individuo. Una vez que el individuo ha sobrevivido hasta su etapa de reproducción, tiene la misma habilidad que todos los demás para reproducirse sin importar su tipo.

A lo que nos referimos diciendo que el proceso de selección depende únicamente de los tipos de los individuos, es a que el proceso de selección se mantiene constante a lo largo del tiempo, depende de la composición genética del individuo, y no depende de la composición del resto de la población. Esto implica que se puede calcular la probabilidad de que un individuo de tipo  $\{i, j\}$  de cualquier generación sobreviva al proceso de selección (para

cualquier  $\{i, j\}$ ). Cada población tendrá sus probabilidades correspondientes, y las tomaremos como parámetros.

A la probabilidad de sobrevivir al proceso de selección de cualquier individuo de tipo  $\{i, j\}$  le llamamos  $\psi_{i,j}$ .

Podemos agrupar estas probabilidades en una matriz triangular de  $r \times r$  a la que llamaremos  $\Psi = (\psi_{i,j})_{i,j}$  y nos referiremos a ella como el coeficiente de selección.

Tomemos una población diploide que se reproduzca de forma sexual, donde la característica que nos interesa está determinada por el genoma de los individuos en cierto “locus” que admite  $r$  posibles alelos. Decimos que un individuo es de tipo  $\{i, j\}$  si tiene un alelo de tipo  $i$  y uno de tipo  $j$ . La población sufre un proceso de selección, donde cada individuo que nace (de tipo  $\{u, v\}$ ) tiene una probabilidad de sobrevivir hasta la edad adulta (etapa de reproducción) dada por la constante  $\psi_{u,v}$ , cada individuo que nace es hijo de cualquiera de los individuos de la generación anterior con la misma probabilidad y hereda un alelo de cada uno de sus padres. Además, se mantiene una cantidad constante de  $N$  individuos adultos en cada generación, los individuos que nacen, pero no sobreviven el proceso de selección no son contados.

Podemos modelar la creación de los  $(n+1)$ -individuos a partir de los  $n$ -individuos de la siguiente forma:

- Se forman dos urnas. En cada urna, cada  $n$ -individuo de tipo  $\{i, j\}$  deposita un gameto de tipo  $\{i\}$  y otro de tipo  $\{j\}$ . Les llamamos  $n$ -urnas.
- Se escoge independiente y uniformemente un alelo de cada urna.
- Uniendo los dos gametos escogidos nace un nuevo individuo.
- El nuevo individuo crece, pero debido al proceso de selección, sólo sobrevive hasta la edad adulta con una probabilidad dada por  $\psi_{u,v}$  (si el individuo es de tipo  $u, v$ ).
- Si el individuo sobrevivió, entonces forma parte de los  $(n+1)$ -individuos. Si el individuo no sobrevivió, no hemos obtenido ningún  $(n+1)$ -individuos en este proceso.

Este proceso se repite hasta obtener  $N$  individuos adultos que sobrevivieron el proceso de selección, y estos conforman a la generación  $(n + 1)$ .

Al igual que en la **sección 1.4**, tenemos que:

- $P_{\{i,i\}}^n$  es el número de  $n$ -individuos de tipo  $\{i, j\}$ ,
- $P^n$  es la matriz con entradas  $\left(P_{\{r,s\}}^n\right)_{r,s}$  que representa la composición de la población en cuestión de cantidad de individuos de cada tipo.
- $p_i^n$  es la cantidad de gametos de las  $n$ -urnas que son de tipo  $\{i\}$ ,

$$p_i^n = 2 \cdot P_{\{i,i\}}^n + \sum_{j \neq i} P_{\{i,j\}}^n.$$

- Cada hijo que nace tiene una probabilidad de ser de tipo  $\{h, k\}$  igual a

$$(2 - \delta_{h,k}) \cdot \frac{p_h^n}{2N} \cdot \frac{p_k^n}{2N}$$

donde  $(2 - \delta_{h,k})$  es igual a 2 si  $h$  y  $k$  son distintos, y es igual a 1 si  $h = k$ .

Calculemos la probabilidad de un  $(n+1)$ -individuo de ser de tipo  $\{i, j\}$ .

Si tomamos un  $(n+1)$ -individuo, sabemos de antemano que ha sobrevivido al proceso de selección, por lo tanto, la probabilidad de que éste sea de tipo  $\{i, j\}$ , a la que llamaremos  $G_{i,j}^n$ , está dada por la siguiente probabilidad

condicional:

$$\begin{aligned}
G_{i,j}^n &= \mathbb{P} \left( \begin{array}{c} \text{nazca y} \\ \text{sobreviva un} \\ \text{hijo tipo } \{i,j\} \end{array} \middle| \begin{array}{c} \text{nace y} \\ \text{sobrevive el hijo} \end{array} \right) \\
&= \frac{\mathbb{P} \left( \begin{array}{c} \text{nazca y} \\ \text{sobreviva un} \\ \text{hijo tipo } \{i,j\} \end{array} \cap \begin{array}{c} \text{nace y} \\ \text{sobrevive el hijo} \end{array} \right)}{\mathbb{P} \left( \begin{array}{c} \text{nace y} \\ \text{sobrevive el hijo} \end{array} \right)} = \frac{\mathbb{P} \left( \begin{array}{c} \text{nazca y} \\ \text{sobreviva un} \\ \text{hijo tipo } \{i,j\} \end{array} \right)}{\mathbb{P} \left( \begin{array}{c} \text{nace y} \\ \text{sobrevive el hijo} \end{array} \right)} \\
&= \frac{\mathbb{P} \left( \begin{array}{c} \text{nazca} \\ \text{un hijo tipo } \{i,j\} \end{array} \right) \cdot \mathbb{P} \left( \begin{array}{c} \text{sobreviva un} \\ \text{hijo tipo } \{i,j\} \end{array} \right)}{\mathbb{P} \left( \begin{array}{c} \text{nace y} \\ \text{sobrevive el hijo} \\ \text{y sea de tipo } \{s,u\} \end{array} \right)} = \frac{\mathbb{P} \left( \begin{array}{c} \text{nazca} \\ \text{un hijo tipo } \{i,j\} \end{array} \right) \cdot \psi_{i,j}}{\sum_{\{s,u\}} \mathbb{P} \left( \begin{array}{c} \text{nazca y} \\ \text{sobreviva el hijo} \\ \text{y sea tipo } \{s,u\} \end{array} \right)} \\
&= \frac{(2 - \delta_{i,j}) \cdot \frac{p_i^n}{2N} \cdot \frac{p_j^n}{2N} \cdot \psi_{i,j}}{\sum_{\{s,u\}} (2 - \delta_{s,u}) \cdot \frac{p_s^n}{2N} \cdot \frac{p_u^n}{2N} \cdot \psi_{s,u}} = \frac{(2 - \delta_{i,j}) \cdot p_i^n \cdot p_j^n \cdot \psi_{i,j}}{\sum_{s \leq u} (2 - \delta_{s,u}) \cdot p_s^n \cdot p_u^n \cdot \psi_{s,u}}.
\end{aligned}$$

Con las  $G_{i,j}^n$ , podemos calcular las probabilidades de transición de la cadena  $\{P^n\}_{n \in \mathbb{N}}$ .

Sabemos que nacen  $N$   $(n+1)$ -individuos, sabemos que el tipo de cada  $(n+1)$ -individuo es una variable aleatoria independiente de las otras variables que determinan el tipo de los demás  $(n+1)$ -individuos y como acabamos de ver, su distribución es una especie de Bernoulli con probabilidades  $\{G_{s,u}^n\}_{s,u}$ . Además, estas probabilidades dependen únicamente de los valores de  $P^n$ . Entonces, dados los valores de  $P^n$ , tenemos que la matriz triangular  $P^{n+1}$  (la matriz que tiene en cada entrada el número de  $(n+1)$ -individuos de cada tipo) es una variable aleatoria *Multinomial*  $\left( N, \left( G_{i,j}^n \right)_{i,j} \right)$ .

La cantidad de  $(n+1)$ -individuos que son de tipo:

$$\begin{pmatrix} \{1, 1\} & \{1, 2\} & \dots & \{1, r\} \\ & \{2, 2\} & \dots & \{2, r\} \\ & & \ddots & \vdots \\ & & & \{r, r\} \end{pmatrix} \text{ se distribuyen multinomial} \\ \text{N y con probabilidades} \begin{pmatrix} G_{1,1}^n & G_{1,2}^n & \dots & G_{1,r}^n \\ & G_{2,2}^n & \dots & G_{2,r}^n \\ & & \ddots & \vdots \\ & & & G_{r,r}^n \end{pmatrix}.$$

Es decir:

$$\begin{aligned}
\mathbf{P}^{n+1} \Big|_{\mathbf{P}^n} \text{ se distribuye } & \text{Multinomial} \left( \mathbf{N}, \begin{pmatrix} G_{1,1}^n & \cdots & G_{1,r}^n \\ & \ddots & \vdots \\ & & G_{r,r}^n \end{pmatrix} \right) \\
\text{con} & G_{i,j}^n = \frac{(2 - \delta_{i,j}) \cdot \psi_{i,j} \cdot p_i^n \cdot p_j^n}{\sum_{k \leq h} (2 - \delta_{k,h}) \cdot \psi_{k,h} \cdot p_k^n \cdot p_h^n} \\
\text{y con} & p_i^n = 2 \cdot P_{\{i,i\}}^n + \sum_{j \neq i} P_{\{i,j\}}^n
\end{aligned}$$

Como  $G_{s,u}^n$  depende únicamente de los valores de  $P^n$ , la cadena  $\{P^n\}_{n \in \mathbb{N}}$  es una cadena de Markov homogénea, con probabilidades de transición dadas por la variable aleatoria multinomial descrita en la ecuación anterior, y cuyo espacio de estados es el conjunto de todas las matrices triangulares cuyas entradas suman  $N$ .

**Definición 1.5.1.** *El modelo discreto de Wright-Fisher diploide, con  $r$  tipos de alelos, con selección y sin mutación, es una cadena de Markov discreta  $\{P_n\}_{n \in \mathbb{N}}$  con espacio de estados*

$$\{ \text{Matrices de } r \times r \text{ triangulares cuyas entradas suman } N \},$$

y probabilidades de transición dadas por

$$\begin{aligned}
\mathbb{P} \left( P^{n+1} = A \mid P^n \right) &= \mathbb{P} \left[ \text{Multinomial} \left( N, \left( G_{i,j}^n \right)_{i,j} \right) = A \right] \\
\text{con} & G_{i,j}^n = \frac{(2 - \delta_{i,j}) \cdot \psi_{i,j} \cdot p_i^n \cdot p_j^n}{\sum_{k \leq h} (2 - \delta_{k,h}) \cdot \psi_{k,h} \cdot p_k^n \cdot p_h^n} \\
\text{y con} & p_i^n = 2 \cdot P_{\{i,i\}}^n + \sum_{j \neq i} P_{\{i,j\}}^n.
\end{aligned}$$

Si tratáramos de representar a estas poblaciones por medio de las  $n$ -urnas, es decir por la cadena  $\{n\text{-urna}\}_{n \in \mathbb{N}}$ , nos quedarían probabilidades de transición muy complicadas. Sin embargo, como veremos en la siguiente sección, agregando una hipótesis acerca de la forma en que ocurre la selección, que hace que no influya la forma en que se agrupan los alelos en pares,

podemos modelar a esta población por medio de la cadena  $\{n\text{-urna}\}_{n \in \mathbb{N}}$ , y sus probabilidades de transición se simplifican tanto que queda una cadena más sencilla que la definida en esta sección.

## 1.6. Modelo Diploide con r Alelos y Selección Multiplicativa

Nos gustaría tener una cadena de Markov que modele a una población con selección y que sea más sencilla que la cadena de la sección anterior.

En esta sección vamos a representar a un subconjunto de las poblaciones de la sección anterior a través de la cadena de Markov  $\{n\text{-urna}\}_{n \in \mathbb{N}}$  que generan. Daremos condiciones sobre la población bajo las cuales esta cadena tiene probabilidades de transición más sencillas que las de la cadena de la sección anterior.

En la sección anterior el primer paso para la creación de la siguiente generación, era crear una urna en donde cada individuo  $\{i, j\}$  deposita un gameto tipo  $\{i\}$  y otro tipo  $\{j\}$ ; lo cual es equivalente a meter a toda la población en una licuadora y olvidarse de cualquier estructura aparte de las cantidades de alelos. Desde este paso lo único que nos queda de la generación anterior es la cantidad de alelos de cada tipo. Gracias a esto, es razonable crear un modelo de esa población donde el espacio de estados nos proporciona información únicamente acerca de las cantidades de alelos de cada tipo.

Sea  $p^n$  la  $n$ -urna  $(p_1^n, \dots, p_i^n, \dots, p_r^n)$ .

Dada la  $n$ -urna  $p^n$ , cada alelo de la  $(n+1)$ -urna tiene una probabilidad de ser de tipo  $\{i\}$  a la que llamaremos  $g_i^n$  igual a

$$\begin{aligned}
& \mathbb{P} \left( \begin{array}{l} \text{ser un alelo} \\ \text{de tipo } \{i\} \end{array} \right) \\
&= \sum_j \mathbb{P} \left( \begin{array}{l} \text{ser el alelo } \{i\} \text{ de} \\ \text{un } (n+1)\text{-individuo} \\ \text{de tipo } \{i, j\} \end{array} \middle| \begin{array}{l} \text{el } (n+1)\text{-individuo} \\ \text{del que proviene el} \\ \text{alelo es de tipo } \{i, j\} \end{array} \right) \cdot \mathbb{P} \left( \begin{array}{l} \text{el } (n+1)\text{-individuo} \\ \text{del que proviene el} \\ \text{alelo sea de tipo } \{i, j\} \end{array} \right) \\
&= \sum_{j \neq i} \frac{1}{2} \cdot G_{i,j}^n + G_{i,i}^n = g_i^n
\end{aligned}$$

$$\text{con } G_{i,j}^n = \frac{(2 - \delta_{i,j}) \cdot \psi_{i,j} \cdot p_i^n \cdot p_j^n}{\sum_{k \leq h} (2 - \delta_{k,h}) \cdot \psi_{k,h} \cdot p_k^n \cdot p_h^n}$$

definida como en la sección anterior, donde  $G_{i,j}^n$  es la probabilidad de cualquier  $(n+1)$ -individuo de ser de tipo  $\{i, j\}$ .

Sin embargo, esto no implica que podamos calcular las probabilidades de transición fácilmente.

**Si las variables aleatorias que determinan el tipo de los alelos de la  $(n+1)$ -urna fueran variables aleatorias independientes**, como son variables aleatorias idénticamente distribuidas, discretas y con la forma de una Bernoulli múltiple, tendríamos que su suma conforma a una variable aleatoria multinomial. Las probabilidades de transición de  $\{p^n\}_{n \in \mathbb{N}} = \{(p_1^n, \dots, p_r^n)\}_{n \in \mathbb{N}}$  serían las probabilidades de una multinomial con parámetros dados por una función del estado anterior, y sería una cadena de Markov homogénea.

$$\begin{aligned}
& \mathbb{P} \left( p^{n+1} = (a_1, \dots, a_r) \mid p^n \right) \\
&= \binom{2N}{a_1, \dots, a_r} \left( \sum_{j \neq 1} \frac{1}{2} \cdot G_{1,j}^n + G_{1,1}^n \right)^{a_1} \cdots \left( \sum_{j \neq r} \frac{1}{2} \cdot G_{r,j}^n + G_{r,r}^n \right)^{a_r} \\
&= \binom{2N}{a_1, \dots, a_r} (g_1^n)^{a_1} \cdots (g_r^n)^{a_r}
\end{aligned}$$

$$\text{con } g_i^n = \sum_{j \neq i} \frac{1}{2} \cdot G_{i,j}^n + G_{i,i}^n = \frac{\sum_j \psi_{i,j} \cdot p_i^n \cdot p_j^n}{\sum_{k \leq h} (2 - \delta_{k,h}) \cdot \psi_{k,h} \cdot p_k^n \cdot p_h^n}$$

$$\begin{aligned}
&= \binom{2N}{a_1, \dots, a_r} \left( \frac{\sum_j \psi_{1,j} \cdot p_1^n \cdot p_j^n}{\sum_{k \leq h} (2 - \delta_{k,h}) \cdot \psi_{k,h} \cdot p_k^n \cdot p_h^n} \right)^{a_1} \cdots \left( \frac{\sum_j \psi_{r,j} \cdot p_r^n \cdot p_j^n}{\sum_{k \leq h} (2 - \delta_{k,h}) \cdot \psi_{k,h} \cdot p_k^n \cdot p_h^n} \right)^{a_r} \\
&= \binom{2N}{a_1, \dots, a_r} \left( \sum_j \psi_{1,j} \cdot p_1^n \cdot p_j^n \right)^{a_1} \cdots \left( \sum_j \psi_{r,j} \cdot p_r^n \cdot p_j^n \right)^{a_r} \left( \frac{1}{\sum_k \sum_h \psi_{k,h} \cdot p_k^n \cdot p_h^n} \right)^{2N}
\end{aligned}$$

Como podemos observar,  $g_i^n = \frac{\sum_j \psi_{i,j} \cdot p_i^n \cdot p_j^n}{\sum_{k \leq h} (2 - \delta_{k,h}) \cdot \psi_{k,h} \cdot p_k^n \cdot p_h^n}$  es una función determinista de  $p^n$  para todo  $i$ , y por lo tanto, una vez fijo  $p^n = (p_1^n, \dots, p_i^n, \dots, p_r^n)$ , las probabilidades de transición a  $p^{n+1}$  están totalmente determinadas, son constantes y tampoco dependen del tiempo  $n$ . Por lo tanto, son probabilidades de transición de una cadena de Markov homogénea.

Tendríamos de esta forma una cadena de Markov con espacio de estados

$$\left\{ (a_1, \dots, a_r) \in \mathbb{N}^r \mid \sum_{i=1}^r a_i = 2N \right\} \quad (1.7)$$

y probabilidades de transición:

$$\mathbb{P} \left( p^{n+1} = (a_1, \dots, a_r) \mid p^n \right) = \binom{2N}{a_1, \dots, a_r} \cdot \frac{\left( \sum_j \psi_{1,j} \cdot p_1^n \cdot p_j^n \right)^{a_1} \cdots \left( \sum_j \psi_{r,j} \cdot p_r^n \cdot p_j^n \right)^{a_r}}{\left( \sum_k \sum_h \psi_{k,h} \cdot p_k^n \cdot p_h^n \right)^{2N}}$$

Sin embargo, esto no siempre sucede, **los tipos de los alelos de la (n+1)-urna no siempre son variables independientes**. Tenemos el siguiente contraejemplo:

**Ejemplo:** Tomemos una población diploide con 2 tipos de alelos (tipo {1}, tipo {2}), un total de 20000 individuos en cada generación, y una selección como en la sección 1.5, con probabilidades de supervivencia tales que sólo sobreviven los individuos heterocigotos, es decir, sólo los individuos de tipo {1, 2} tienen probabilidad de sobrevivir mayor a 0.

Consideremos las probabilidades de transición a partir del estado

$$P^n = \begin{pmatrix} 10000 & 0 \\ & 10000 \end{pmatrix}$$

$$N = 20000 \quad r = 2 \quad \Psi = \begin{pmatrix} 0 & 1 \\ & 0 \end{pmatrix} \quad p^n = (20000, 20000)$$

El proceso es el siguiente:

- Tenemos la n-urna  $p^n = (20000, 20000)$ .
- Se escogen 20000 individuos con probabilidades dadas por  $(G_{i,j}^n)_{i,j}$ .

Donde  $\psi_{1,1} = 0$ ,  $\psi_{1,2} = 1$ ,  $\psi_{2,2} = 0$  y  $G_{i,j}^n = \frac{(2-\delta_{i,j}) \cdot \psi_{i,j} \cdot p_i^n \cdot p_j^n}{\sum_{k \leq h} (2-\delta_{k,h}) \cdot \psi_{k,h} \cdot p_k^n \cdot p_h^n}$ .

$$G_{1,1}^n = \frac{0 \cdot p_1^n \cdot p_1^n}{0 + 0 + 2 \cdot \psi_{1,2} \cdot p_1^n \cdot p_2^n} = 0$$

$$G_{1,2}^n = \frac{2 \cdot 1 \cdot p_1^n \cdot p_2^n}{0 + 0 + 2 \cdot \psi_{1,2} \cdot p_1^n \cdot p_2^n} = 1$$

$$G_{2,1}^n = \frac{0 \cdot p_2^n \cdot p_2^n}{0 + 0 + 2 \cdot \psi_{1,2} \cdot p_1^n \cdot p_2^n} = 0.$$

- A partir de los individuos escogidos, se forman las (n+1)-urnas.

En el segundo punto se escogen aleatoriamente 20000 individuos, cada individuo tiene una probabilidad 0 de ser de tipo  $\{1, 1\}$ , una probabilidad 0 de ser de tipo  $\{2, 2\}$ , y una probabilidad 1 de ser de tipo  $\{1, 2\}$ , por lo que podemos afirmar que la (n+1)-generación y la (n+1)-urna serán

$$P^{n+1} = \begin{pmatrix} 0 & 20000 \\ & 0 \end{pmatrix}, \quad p^{n+1} = (20000, 20000).$$

Queda claro que en esta población, en la cadena  $\{n - \text{urna}\}_{n \in \mathbb{N}}$ , las variables que denotan el tipo de los alelos de la (n+1)-urna dada la n-urna, no son variables independientes.

Una de las muchas maneras de demostrarlo es considerar el último de los alelos elegidos. Al sobrevivir únicamente individuos heterocigotos, en la (n+1)-urna debe haber la misma cantidad de alelos de ambos tipos, entonces el total de los alelos de cada tipo debe ser 20000, y justo antes de escoger el último alelo debe haber 20000 alelos de un tipo (digamos tipo A) y 19999 alelos del otro tipo (tipo B). El último alelo está destinado a ser de tipo B. Los tipos del resto de los alelos han determinado completamente el tipo del último alelo, no son independientes.  $\square$

Hemos demostrado con este ejemplo que no en todas las poblaciones los tipos de los alelos son variables independientes. Algo que se puede hacer en estos casos es encontrar las características de la población que provocan que las variables "tipo de alelo" sí sean independientes.

Recordemos la sección 1.4, en el modelo de esa sección, se calculan de una forma muy sencilla las probabilidades de transición de la n-urna a la (n+1)-urna. Todo el argumento recae en que al formar a los (n+1)-individuos se escogen aleatoria e independientemente dos alelos de la n-urna, y para formar la (n+1)-urna se depositan los dos alelos escogidos. Estos pasos se reducen a escoger directamente de la n-urna los alelos que conforman la (n+1)-urna, y esto se hace de forma independiente.

La única diferencia entre el modelo de la sección 1.5 y el de la sección 1.4 es la selección. Por lo tanto, las características buscadas deben estar relacionadas principalmente con el coeficiente de selección.

Buscamos una característica que provoque que escoger un (n+1)-individuo sea equivalente a escoger de forma independiente dos alelos.

Escoger un alelo tipo  $\{i\}$  tiene una probabilidad de:

$$g_i^n = \sum_{j \neq i} \frac{1}{2} \cdot G_{i,j}^n + G_{i,i}^n = \frac{\sum_j \psi_{i,j} \cdot p_i^n \cdot p_j^n}{\sum_{k \leq h} (2 - \delta_{k,h}) \cdot \psi_{k,h} \cdot p_k^n \cdot p_h^n} = \frac{\sum_j \psi_{i,j} \cdot p_i^n \cdot p_j^n}{\sum_k \sum_h \psi_{k,h} \cdot p_k^n \cdot p_h^n}.$$

Escoger a un individuo tipo  $\{i, j\}$  tiene una probabilidad de:

$$G_{i,j}^n = \frac{(2 - \delta_{i,j}) \cdot \psi_{i,j} \cdot p_i^n \cdot p_j^n}{\sum_{k \leq h} (2 - \delta_{k,h}) \cdot \psi_{k,h} \cdot p_k^n \cdot p_h^n}.$$

Buscamos que suceda:

$$G_{i,j}^n = (2 - \delta_{i,j}) \cdot g_i^n \cdot g_j^n$$

$$\begin{aligned} \frac{(2 - \delta_{i,j}) \cdot \psi_{i,j} \cdot p_i^n \cdot p_j^n}{\sum_{k \leq h} (2 - \delta_{k,h}) \cdot \psi_{k,h} \cdot p_k^n \cdot p_h^n} &= (2 - \delta_{i,j}) \cdot \frac{\sum_k \psi_{i,k} \cdot p_i^n \cdot p_k^n}{\sum_{k \leq h} (2 - \delta_{k,h}) \cdot \psi_{k,h} \cdot p_k^n \cdot p_h^n} \\ &\cdot \frac{\sum_h \psi_{h,j} \cdot p_h^n \cdot p_j^n}{\sum_{k \leq h} (2 - \delta_{k,h}) \cdot \psi_{k,h} \cdot p_k^n \cdot p_h^n}. \end{aligned}$$

Simplifiquemos esta ecuación:

$$\begin{aligned} (2 - \delta_{i,j}) \cdot \psi_{i,j} \cdot p_i^n \cdot p_j^n &= (2 - \delta_{i,j}) \cdot \frac{(\sum_k \psi_{i,k} \cdot p_i^n \cdot p_k^n) \cdot (\sum_h \psi_{h,j} \cdot p_h^n \cdot p_j^n)}{\sum_{k \leq h} (2 - \delta_{k,h}) \cdot \psi_{k,h} \cdot p_k^n \cdot p_h^n} \\ \psi_{i,j} \cdot p_i^n \cdot p_j^n &= \frac{p_i^n \cdot p_j^n \cdot (\sum_k \psi_{i,k} \cdot p_k^n) \cdot (\sum_h \psi_{h,j} \cdot p_h^n)}{\sum_k \sum_h \psi_{k,h} \cdot p_k^n \cdot p_h^n} \\ \psi_{i,j} &= \frac{\sum_k \sum_h \psi_{i,k} \cdot p_k^n \cdot \psi_{h,j} \cdot p_h^n}{\sum_k \sum_h \psi_{k,h} \cdot p_k^n \cdot p_h^n} \\ \psi_{i,j} \cdot \left( \sum_k \sum_h \psi_{k,h} \cdot p_k^n \cdot p_h^n \right) &= \sum_k \sum_h \psi_{i,k} \cdot \psi_{h,j} \cdot p_k^n \cdot p_h^n \\ 0 &= \sum_k \sum_h (\psi_{i,k} \cdot \psi_{h,j} - \psi_{k,h} \cdot \psi_{i,j}) \cdot p_k^n \cdot p_h^n \end{aligned}$$

$p_k^n$  y  $p_h^n$  deberían poder tomar cualquier valor entre 0 y 1. Tenemos así que cualquier combinación lineal de los elementos  $(\psi_{i,k} \cdot \psi_{h,j} - \psi_{k,h} \cdot \psi_{i,j})$  es cero, y

por lo tanto:

$$\psi_{i,k} \cdot \psi_{h,j} - \psi_{k,h} \cdot \psi_{i,j} = 0 \quad \forall i, j, k, h$$

Es decir:

$$\psi_{i,k} \cdot \psi_{h,j} = \psi_{k,h} \cdot \psi_{i,j}. \quad (1.8)$$

Reuniendo lo dicho hasta el momento, el modelo se puede expresar fácilmente en términos de las n-urnas (el vector de la cantidad de alelos de cada tipo en la generación n) con probabilidades de transición distribuidas como una

*Multinomial*( $g_1^n, \dots, g_r^n$ ) si las variables que determinan el tipo de cada alelo son independientes entre sí. Lo cual sucede si y sólo si escoger un (n+1)-individuo es equivalente a escoger de forma independiente dos alelos. Lo cual sucede si y sólo si  $G_{i,j}^n = (2 - \delta_{i,j})g_i^n \cdot g_j^n$  y como acabamos de probar, esto pasa si y sólo si  $\psi_{i,k} \cdot \psi_{h,j} = \psi_{k,h} \cdot \psi_{i,j}$ .

La característica  $\psi_{i,k} \cdot \psi_{h,j} = \psi_{k,h} \cdot \psi_{i,j}$  nos da mucha más información de lo que aparenta. Esta característica es equivalente a decir que los coeficientes de selección  $\Psi = \{\psi_{i,j}\}_{i,j}$  son “multiplicativos”, es decir, que existen constantes  $\{v_i\}_i$  tales que  $\psi_{h,k} = v_h \cdot v_k \quad \forall h, k$ . La demostración de esta equivalencia la damos a continuación:

*Demostración.* Recordemos que  $\psi_{r,s} = \psi_{s,r}$  para todas las  $s$  y  $r$ , y supongamos que  $\psi_{i,k} \cdot \psi_{h,j} = \psi_{k,h} \cdot \psi_{i,j}$  para todo  $i, k, h, j$ , entonces

$$\begin{aligned} (\psi_{i,j})^2 &= \psi_{i,j} \cdot \psi_{i,j} = \psi_{i,i} \cdot \psi_{j,j} \\ \psi_{i,j} &= \sqrt{\psi_{i,i} \cdot \psi_{j,j}} = \sqrt{\psi_{i,i}} \cdot \sqrt{\psi_{j,j}}. \end{aligned}$$

Es decir, si  $\psi_{i,k} \cdot \psi_{h,j} = \psi_{k,h} \cdot \psi_{i,j}$  para toda  $i, j, k, h$ , entonces existen constantes  $v_k, v_h$  tales que para toda  $h, k$  sucede que  $\psi_{h,k} = v_h \cdot v_k$ .

$$\left\{ v_h = \sqrt{\psi_{h,h}} \right\}_h$$

Ahora supongamos que los coeficientes de selección son multiplicativos ( para toda  $h, k$  existen  $v_h, v_k$  tales que  $\psi_{h,k} = v_h \cdot v_k$  ) entonces se cumple que  $\psi_{i,k} \cdot \psi_{h,j} = \psi_{k,h} \cdot \psi_{i,j}$ :

$$\psi_{i,k} \cdot \psi_{h,j} = (v_i \cdot v_k) \cdot (v_h \cdot v_j) = (v_i \cdot v_j) \cdot (v_h \cdot v_k) = \psi_{i,j} \cdot \psi_{h,k}$$

□

Como conclusión tenemos que en nuestra población  $P^n$  sucede que  $G_{i,j}^n = (2 - \delta_{i,j})g_i^n \cdot g_j^n$  si y sólo si los coeficientes de selección son multiplicativos ( $\psi_{i,j} = v_i \cdot v_j$ ).

En este caso, el tipo de alelo de cada gameto escogido es independiente al de los demás, y entonces podemos utilizar el modelo desarrollado en la página 24 al inicio de esta sección, al que denotamos por  $\{p^n\}_{n \in \mathbb{N}} = \{(p_1^n, \dots, p_r^n)\}_{n \in \mathbb{N}}$ .

En la ecuación (1.7) en la página 25 vimos que esta cadena era una cadena de Markov, con espacio de estados

$$\left\{ (a_1, \dots, a_r) \in \mathbb{N}^r \mid \sum_{i=1}^r a_i = 2N \right\}$$

y probabilidades de transición dadas por:

$$\begin{aligned} \mathbb{P}(p^{n+1} = (a_1, \dots, a_r) \mid p^n) &= \binom{2N}{a_1, \dots, a_r} \cdot (g_1^n)^{a_1} \cdots (g_r^n)^{a_r} \\ &= \binom{2N}{a_1, \dots, a_r} \cdot \left( \frac{\sum_j \psi_{1,j} \cdot p_1^n \cdot p_j^n}{\sum_k \sum_h \psi_{k,h} \cdot p_k^n \cdot p_h^n} \right)^{a_1} \cdots \left( \frac{\sum_j \psi_{r,j} \cdot p_r^n \cdot p_j^n}{\sum_k \sum_h \psi_{k,h} \cdot p_k^n \cdot p_h^n} \right)^{a_r} \\ &= \binom{2N}{a_1, \dots, a_r} \cdot \frac{\left( \sum_j \psi_{1,j} \cdot p_1^n \cdot p_j^n \right)^{a_1} \cdots \left( \sum_j \psi_{r,j} \cdot p_r^n \cdot p_j^n \right)^{a_r}}{\left( \sum_k \sum_h \psi_{k,h} \cdot p_k^n \cdot p_h^n \right)^{2N}}. \end{aligned}$$

Sin embargo, utilizando otra vez que los coeficientes de selección son multiplicativos, la expresión de  $g_i^n$  se simplifica todavía más.

$$g_i^n = \frac{\sum_j \psi_{i,j} \cdot p_i^n \cdot p_j^n}{\sum_k \sum_h \psi_{k,h} \cdot p_k^n \cdot p_h^n} = \frac{v_i \cdot p_i^n \cdot \left( \sum_j v_j \cdot p_j^n \right)}{\left( \sum_k v_k \cdot p_k^n \right) \cdot \left( \sum_h v_h \cdot p_h^n \right)} = \frac{v_i \cdot p_i^n}{\sum_k v_k \cdot p_k^n}$$

y con esto, las probabilidades de transición se ven como:

$$\mathbb{P}(p^{n+1} = (a_1, \dots, a_r) \mid p^n) = \binom{2N}{a_1, \dots, a_r} \cdot \left( \frac{v_1 \cdot p_1^n}{\sum_k v_k \cdot p_k^n} \right)^{a_1} \cdots \left( \frac{v_r \cdot p_r^n}{\sum_k v_k \cdot p_k^n} \right)^{a_r}$$

$$p^n \sim \text{Multinomial} \left[ N, \left( \frac{v_1 \cdot p_1^n}{\sum_k v_k \cdot p_k^n}, \dots, \frac{v_r \cdot p_r^n}{\sum_k v_k \cdot p_k^n} \right) \right].$$

En la literatura, cuando en una población sucede que  $G_{i,j}^n = (2 - \delta_{i,j})g_i^n \cdot g_j^n$ , (lo cual sucede si y sólo si los coeficientes de selección son multiplicativos), se dice que la población preserva **proporciones de Hardy-Weinberg**.

**Definición 1.6.1.** *El modelo discreto de Wright-Fisher diploide, con  $r$  tipos de alelos, con selección **multiplicativa** y sin mutación, es una cadena de Markov discreta  $\{p_n^n\}_{n \in \mathbb{N}} = \{(p_1^n \dots, p_r^n)_n\}_{n \in \mathbb{N}}$  con espacio de estados*

$$\left\{ (a_1, \dots, a_r) \in \mathbb{N}^r \mid \sum_{i=1}^r a_i = 2N \right\}$$

y probabilidades de transición dadas por

$$\mathbb{P}\left( p^{n+1} = (a_1, \dots, a_r) \mid p^n \right) = \binom{2N}{a_1, \dots, a_r} \cdot \left( \frac{v_1 \cdot p_1^n}{\sum_k v_k \cdot p_k^n} \right)^{a_1} \cdots \left( \frac{v_r \cdot p_r^n}{\sum_k v_k \cdot p_k^n} \right)^{a_r}$$

con  $v_i = \sqrt{\psi_{i,i}}$

## 1.7. Modelo Diploide con r Alelos, Mutación y Selección General o Multiplicativa

En esta sección generalizamos los modelos de las dos secciones anteriores agregando la noción de mutación de la sección 1.3. Obtenemos de esta forma dos modelos que generalizan a todos los modelos de este capítulo. Cualquier modelo de Wright-Fisher presentado se puede expresar en términos de uno de los siguientes dos modelos con los parámetros adecuados.

El modelo se basa en una población diploide con  $r$  tipos alelos en el locus que nos interesa, que sufre un proceso de selección (ya sea general o multiplicativa) y cuyos alelos mutan con cierta probabilidad.

Tenemos  $r \cdot (r - 1)$  parámetros  $\mu_{i,j}$  con  $i \neq j$  que representan la probabilidad de un alelo de tipo  $i$  de mutar a un alelo de tipo  $j$ . Para facilitar la notación, le llamaremos  $\mu_{i,i}$  a la probabilidad de un alelo de tipo  $i$  de no mutar. Como la probabilidad de no mutar debería ser  $(1 - \text{la probabilidad de mutar})$ ,  $\mu_{i,i} = 1 - \sum_{j \neq i} \mu_{i,j}$ .

Para seguir los modelo expuestos en el libro de Ethier y Kurtz [?], aunque vaya en contra de la intuición, la mutación sucederá después de la selección.

En la subsección 1.7.2, donde desarrollamos el modelo con mutación de una población Diploide con r Alelos y Selección Multiplicativa, analizamos cuál es la relación entre agregar la mutación antes y después de la selección (para la selección multiplicativa). Encontramos que ambas opciones generan el mismo modelo, con un cambio en los parámetros, y daremos la relación entre éstos. Debido a la complejidad del modelo con Selección General (no necesariamente multiplicativa) de la subsección 1.7.1, no analizaremos la relación entre el modelo donde la mutación sucede antes que la selección, y el modelo donde la mutación sucede después.

### 1.7.1. Mutación y Selección General.

Modelaremos una población idéntica a la de la sección 1.5 (Modelo Diploide con  $r$  Alelos y Selección General) pero cuyos alelos pueden mutar de un tipo a otro.

Tomemos una población diploide que se reproduzca de forma sexual, donde la característica que nos interesa está determinada por el genoma de los individuos en cierto “locus” que admite  $r$  posibles alelos. Decimos que un individuo es de tipo  $\{i, j\}$  si tiene un alelo de tipo  $i$  y uno de tipo  $j$ . La población sufre un proceso de selección, donde cada individuo que nace (de tipo  $\{u, v\}$ ) tiene una probabilidad de sobrevivir hasta la edad adulta (etapa de reproducción) dada por la constante  $\psi_{u,v}$ , cada individuo que nace es hijo de cualesquiera dos individuos de la generación anterior con la misma probabilidad y hereda un alelo de cada uno de sus dos padres. Además, se mantiene una cantidad constante de  $N$  individuos adultos en cada generación, los individuos que nacen, pero no sobreviven el proceso de selección no son contados.

El proceso de una generación a otra es el siguiente:

- Tenemos  $N$  n-individuos distribuidos en tipos como indica la matriz  $P^n$ .
- Se forman las n-urnas, en cada n-urna, cada n-individuo de tipo  $\{i, j\}$  deposita un gameto de tipo  $\{i\}$  y otro de tipo  $\{j\}$ . Las n-urnas son  $p^n = (p_1^n, \dots, p_r^n)$  con  $p_i^n$  cantidad de gametos de tipo  $\{i\}$ .

$$p_i^n = 2 \cdot P_{\{i,i\}}^n + \sum_{j \neq i} P_{\{i,j\}}^n.$$

- Se escoge uniformemente un alelo de dos de las n-urna.
- Uniendo los dos gametos escogidos nace un nuevo individuo.
- El nuevo individuo crece, pero debido al proceso de selección, sólo sobrevive hasta la edad adulta con una probabilidad dada por  $\psi_{u,v}$  (si el individuo es de tipo  $\{u, v\}$ ).
- Si el nuevo individuo sobrevivió, formará parte de los próximos  $(n+1)$ -individuos. Si el nuevo individuo no sobrevivió, no se ha obtenido nada.
- Se vuelven a crear individuos a partir de las n-urnas hasta obtener  $N$  individuos adultos que sobrevivieron el proceso de selección.

Hasta este punto, cualquiera de los  $N$  individuos sobrevivientes tiene una probabilidad  $G_{i,j}^n$  de ser de tipo  $\{i, j\}$ .

$$G_{i,j}^n = \frac{(2 - \delta_{i,j}) \cdot \psi_{i,j} \cdot p_i^n \cdot p_j^n}{\sum_{k \leq h} (2 - \delta_{k,h}) \cdot \psi_{k,h} \cdot p_k^n \cdot p_h^n}.$$

■ **Mutación.**

En este paso, los gametos de los  $N$  individuos sobrevivientes mutan. Cada gameto muta de tipo  $\{i\}$  a tipo  $\{j\}$  con una probabilidad  $\mu_{i,j}$ . Y estos individuos mutados son a los que llamamos (n+1)-individuos.

Como consecuencia, la probabilidad de un (n+1)-individuo de ser de tipo  $\{i, j\}$ , a la que llamaremos  $^{mut}G_{\{i,j\}}^n$  es

$$\begin{aligned} ^{mut}G_{\{i,j\}}^n &= \mathbb{P}\left(\begin{array}{c} \text{ser de tipo} \\ \{i,j\} \end{array}\right) = \sum_{k \leq l} \frac{1}{1 + \delta_{i,j}} \cdot (\mu_{k,i} \cdot \mu_{l,j} + \mu_{k,j} \cdot \mu_{l,i}) \cdot G_{k,l}^n \\ &= \frac{1}{2} \cdot \sum_k \sum_l (\mu_{k,i} \cdot \mu_{l,j} + \mu_{k,j} \cdot \mu_{l,i}) \cdot G_{k,l}^n \\ &= \sum_{\{u,v\}} \mathbb{P}\left(\begin{array}{c} \text{los gametos muten} \\ \text{de tipo } \{u\} \text{ y } \{v\} \\ \text{a gametos de tipo} \\ \{i\} \text{ y } \{j\} \end{array}\right) \cdot \mathbb{P}\left(\begin{array}{c} \text{el individuo} \\ \text{nació y sobrevivió} \\ \text{siendo de tipo} \\ \{u, v\} \end{array}\right) \cdot \mathbb{P}\left(\begin{array}{c} \text{el individuo} \\ \text{nació y sobrevivió} \\ \text{siendo de tipo} \\ \{u, v\} \end{array}\right). \end{aligned}$$

En otras palabras, la probabilidad de un (n+1)-individuo de ser de tipo  $\{i, j\}$  es igual a la probabilidad del individuo de haber nacido y sobrevivido siendo de cualquier tipo (digamos  $\{u, v\}$ ) por la probabilidad de que sus gametos muten de esos tipos ( $\{u\}, \{v\}$ ) a gametos de tipo  $\{i\}$  y  $\{j\}$ .

Entonces la cantidad de (n+1)-individuos que hay de cada tipo, dada la cantidad de n-individuos que hay de cada tipo, se distribuye como una variable aleatoria *Multinomial*  $\left(N, \left(^{mut}G_{\{i,j\}}^n\right)_{i,j}\right)$ .

Lo cual define a una cadena de Markov homogénea con probabilidades de transición dadas por:

$$\mathbb{P} \left( P^{n+1} = \begin{pmatrix} a_{1,1} & \cdots & a_{1,r} \\ & \ddots & \vdots \\ & & a_{r,r} \end{pmatrix} \mid P^n = \begin{pmatrix} P_{\{1,1\}}^n & \cdots & P_{\{1,r\}}^n \\ & \ddots & \vdots \\ & & P_{\{r,r\}}^n \end{pmatrix} \right) =$$

$$\binom{N}{a_{1,1} \dots a_{i,j} \dots a_{1,r}} \cdot \prod_{i,j} \left( \sum_{k \leq l} \frac{1}{1 + \delta_{i,j}} \cdot (\mu_{k,i} \cdot \mu_{l,j} + \mu_{k,j} \cdot \mu_{l,i}) \cdot G_{k,l}^n \right)^{a_{i,j}}$$

*con*

$$G_{i,j}^n = \frac{(2 - \delta_{i,j}) \cdot \psi_{i,j} \cdot p_i^n \cdot p_j^n}{\sum_{k \leq h} (2 - \delta_{k,h}) \cdot \psi_{k,h} \cdot p_k^n \cdot p_h^n}$$

*y*

$$p_i^n = 2 \cdot P_{\{i,i\}}^n + \sum_{j \neq i} P_{\{i,j\}}^n$$

**Definición 1.7.1.** *El modelo discreto de Wright-Fisher diploide, con  $r$  tipos de alelos, con selección y con mutación general (no necesariamente multiplicativa), es una cadena de Markov discreta  $\{P_n\}_{n \in \mathbb{N}}$  con espacio de estados*

$$\{ \text{Matrices de } r \times r \text{ triangulares cuyas entradas suman } N \},$$

*y probabilidades de transición dadas por la ecuación anterior. También las podemos describir de la siguiente forma:*

$$\mathbb{P} \left( P^{n+1} = A \mid P^n \right) = \mathbb{P} \left[ \text{Multinomial} \left( N, \left( {}^{mut}G_{i,j}^n \right)_{i,j} \right) = A \right]$$

$$\text{con } {}^{mut}G_{\{i,j\}}^n = \sum_{k \leq l} \frac{1}{1 + \delta_{i,j}} \cdot (\mu_{k,i} \cdot \mu_{l,j} + \mu_{k,j} \cdot \mu_{l,i}) \cdot G_{k,l}^n$$

$$\text{con } G_{i,j}^n = \frac{(2 - \delta_{i,j}) \cdot \psi_{i,j} \cdot p_i^n \cdot p_j^n}{\sum_{k \leq h} (2 - \delta_{k,h}) \cdot \psi_{k,h} \cdot p_k^n \cdot p_h^n}$$

$$\text{y con } p_i^n = 2 \cdot P_{\{i,i\}}^n + \sum_{j \neq i} P_{\{i,j\}}^n.$$

### 1.7.2. Mutación y Selección Multiplicativa

Ahora modelaremos una población idéntica a la de la sección 1.6 (Modelo Diploide con  $r$  Alelos y Selección Multiplicativa), pero cuyos alelos pueden mutar de un tipo a otro.

Tomemos una población idéntica a la de la subsección anterior pero con coeficientes de selección multiplicativos. Tenemos así las constantes  $(v_k)_{(1 \leq k \leq r)}$  que al multiplicarse nos dan los coeficientes de selección  $\psi_{f,g} = v_f \cdot v_g$  (la probabilidad de cualquier individuo de tipo  $\{f, g\}$  de sobrevivir hasta la edad adulta). Y tenemos que la probabilidad de que un alelo mute de tipo  $\{u\}$  a tipo  $\{s\}$  es el parámetro constante  $\mu_{u,s}$  para todo  $1 \leq u, s \leq r$ .

Recordemos que si la selección es multiplicativa,  $G_{i,j}^n = (2 - \delta_{i,j}) \cdot g_i^n \cdot g_j^n$ , es decir, escoger a un alelo de tipo  $\{i, j\}$  es equivalente a escoger un alelo de tipo  $\{i\}$  y otro de tipo  $\{j\}$  de forma independiente. Como todos los individuos selectos se escogen de manera independiente, esto significa que el tipo de los alelos de la  $(n+1)$ -urna son variables aleatorias independientes, y la probabilidad de que un alelo de la  $(n+1)$ -urna sea de tipo  $\{i\}$  es  $g_i^n$ .

Empezaremos siguiendo al modelo presentado en el libro de Ethier y Kurtz [?], donde la mutación de los alelos ocurre después del proceso de selección, cuando los individuos ya son adultos. Más adelante analizaremos un modelo en el que la mutación se da antes que la selección. Ambos modelos resultan generar la misma cadena de Markov, pero con distintos parámetros. Daremos en la subsección 1.7.2 la correspondencia entre parámetros de ambos modelos.

### Primero Selección Multiplicativa y después Mutación.

Este modelo se basa en poblaciones en las que primero sucede un proceso de selección multiplicativa y después ocurren mutaciones en los alelos de los individuos sobrevivientes. El modelo se puede consultar en el capítulo 10 del libro de Ethier y Kurtz [?], en las páginas 411, 412, 413.

El proceso de una generación a otra es el siguiente:

- Tenemos las  $n$ -urnas,  $p^n = (p_1^n, \dots, p_r^n)$  con  $p_i^n$  cantidad de gametos de tipo  $\{i\}$ .
- Se escoge uniformemente un alelo de dos de las  $n$ -urna.
- Uniendo los dos gametos escogidos nace un nuevo individuo.
- El nuevo individuo crece, pero debido al proceso de selección, sólo sobrevive hasta la edad adulta con una probabilidad dada por  $\psi_{u,s} = v_u \cdot v_s$  (si el individuo es de tipo  $\{u, s\}$ ).
- Si el nuevo individuo sobrevivió, formará parte de los próximos  $(n+1)$ -individuos. Si el nuevo individuo no sobrevivió, no se ha obtenido nada.
- Se vuelven a crear individuos a partir de las  $n$ -urnas hasta obtener  $N$  individuos adultos que sobrevivieron el proceso de selección.

Hasta este punto, cualquiera de los  $N$  individuos sobrevivientes tiene una probabilidad  $G_{i,j}^n$  de ser de tipo  $\{i, j\}$ .

$$G_{i,j}^n = \frac{(2 - \delta_{i,j}) \cdot \psi_{i,j} \cdot p_i^n \cdot p_j^n}{\sum_{k \leq h} (2 - \delta_{k,h}) \cdot \psi_{k,h} \cdot p_k^n \cdot p_h^n}.$$

Y como consecuencia, utilizando el hecho de que la selección es multiplicativa en la última igualdad, cualquiera de los alelos de los  $N$  individuos sobrevivientes tiene una probabilidad  $g_i^n$  de ser un alelo de tipo  $\{i\}$ :

$$g_i^n = \sum_{j \neq i} \frac{1}{2} G_{i,j}^n + G_{i,i}^n = \frac{\sum_j \psi_{i,j} \cdot p_i^n \cdot p_j^n}{\sum_k \sum_h \psi_{k,h} \cdot p_k^n \cdot p_h^n} = \frac{v_i \cdot p_i^n}{\sum_k v_k \cdot p_k^n}.$$

- Los individuos sobrevivientes forman urnas con los gametos que los conforman (lo que correspondería a formar las  $(n+1)$ -urnas). Para formar estas urnas, cada individuo de tipo  $\{r, s\}$  deposita en cada urna un gameto de tipo  $\{r\}$  y otro de tipo  $\{s\}$ .

Como la selección de esta subsección es multiplicativa, las variables que representan el tipo de los alelos son independientes entre sí, y entonces las urnas creadas por los individuos sobrevivientes tienen una distribución *Multinomial* $(N, (g_1^n, \dots, g_r^n))$ .

- **Mutación.**

En este paso, los gametos de la urna mutan. Cada gameto muta de tipo  $\{i\}$  a tipo  $\{j\}$  con una probabilidad  $\mu_{i,j}$ . La urna resultante es a la que llamamos  $(n+1)$ -urna.

La probabilidad de cada alelo de la  $(n+1)$ -urna de ser de tipo  $\{i\}$ , a la que llamaremos  ${}^{sel,mut}g_i^n$ , es:

$${}^{sel,mut}g_i^n = \sum_k \mu_{k,i} \cdot g_k^n.$$

Entonces, dada la  $n$ -urna, la  $(n+1)$ -urna se distribuye como una variable aleatoria *Multinomial*  $[N, ({}^{sel,mut}g_1^n, \dots, {}^{sel,mut}g_r^n)]$ .

Estas probabilidades de transición definen a una cadena de Markov homogénea.

**Definición 1.7.2.** *El modelo discreto de Wright-Fisher diploide, con  $r$  tipos de alelos, con selección multiplicativa y con mutación, donde **la selección ocurre antes que la mutación**, es una cadena de Markov discreta  $\{p_n\}_{n \in \mathbb{N}} = \{(p_1^n \dots p_r^n)_n\}_{n \in \mathbb{N}}$  con espacio de estados*

$$\left\{ (a_1, \dots, a_r) \in \mathbb{N}^r \mid \sum_{i=1}^r a_i = 2N \right\}$$

y probabilidades de transición dadas por

$$\mathbb{P}\left(p^{n+1} = (a_1, \dots, a_r) \mid p^n\right) = \binom{2N}{a_1, \dots, a_r} \cdot ({}^{sel,mut}g_1^n)^{a_1} \dots ({}^{sel,mut}g_r^n)^{a_r}$$

$$\text{con } {}^{sel,mut}g_s^n = \sum_k \mu_{k,s} \cdot g_k^n = \frac{\sum_k \mu_{k,s} \cdot v_k \cdot p_k^n}{\sum_a v_a \cdot p_a^n}$$

$$\text{con } g_k^n = \frac{v_k \cdot p_k^n}{\sum_a v_a \cdot p_a^n}$$

$$\text{y con } v_s = \sqrt{\psi_{s,s}} \quad .$$

Se puede ver en el libro de Ethier y Kurtz [?] en la página 412 que:

$${}^{sel,mut}g_i^n = \sum_j \frac{1}{2} \cdot (1 + \delta_{i,j}) \cdot {}^{mut}G_{i,j}^n.$$

### Primero Mutación y luego Selección Multiplicativa.

En los modelos biológicos, la mutación ocurre cuando se está gestando el individuo, y la selección ocurre cuando éste ya ha sido gestado. Además, la mutación influye en las características fisiológicas y por lo tanto en su habilidad para sobrevivir, es decir, afecta su coeficiente de selección.

En el modelo que vamos a construir en esta subsección, el proceso de mutación sucede antes que el de selección, y una vez que el individuo ha mutado, su coeficiente de selección corresponde al del tipo de individuo al cual mutó.

Tomamos el mismo tipo de población que en la subsección anterior, pero con la mutación antes que la selección. El proceso de una generación a la siguiente es casi el mismo, sólo se intercambia el orden.

El proceso de una generación a otra es el siguiente:

- Tenemos las n-urnas,  $p^n = (p_1^n, \dots, p_r^n)$  con  $p_i^n$  cantidad de gametos de tipo  $\{i\}$ .

- Mutación.** Los gametos de las  $n$ -urnas mutan. Cada gameto muta de tipo  $\{i\}$  a tipo  $\{j\}$  con una probabilidad  $\mu_{i,j}$ .  
 Una vez mutada la  $n$ -urna, la probabilidad de tomar al azar un gameto de tipo  $\{i\}$ , a la que llamaremos  $^{mut}p_i^n$  es

$$^{mut}p_i^n = \sum_{s=1}^r p_s^n \cdot \mu_{s,i}.$$

A los vectores  $^{mut}p^n = ( ^{mut}p_1^n, \dots, ^{mut}p_r^n )$  les llamaremos “ $n$ -urnas mutadas”.

- A partir de las  $n$ -urnas mutadas se construye la generación  $(n+1)$ . La forma en que se construye es la misma que en el modelo diploide con selección multiplicativa y sin selección (sección 1.6), únicamente intercambiamos las  $n$ -urnas por  $n$ -urnas mutadas.

El proceso de selección multiplicativa (idéntico al de la sección 1.6) es el siguiente:

- Se escoge uniformemente un alelo de dos de las  $n$ -urnas mutadas.
- Uniando los dos gametos escogidos nace un nuevo individuo.
- El nuevo individuo crece, pero debido al proceso de selección, sólo sobrevive hasta la edad adulta con una probabilidad dada por  $\psi_{u,s} = v_u \cdot v_s$  (si el individuo es de tipo  $\{u, s\}$ ).
- Si el nuevo individuo sobrevivió, formará parte de los próximos  $(n+1)$ -individuos. Si el nuevo individuo no sobrevivió, no se ha obtenido nada.
- Se vuelven a crear individuos a partir de las  $n$ -urnas hasta obtener  $N$  individuos adultos que sobrevivieron el proceso de selección.

Hasta este punto, cualquiera de los  $N$  individuos sobrevivientes tiene una probabilidad de ser de tipo  $\{i, j\}$  a la que llamaremos  $F_{i,j}^n$ , que sería lo equivalente a la probabilidad  $G_{i,j}^n$  si no hubiéramos agregado la mutación.

$$F_{i,j}^n = \frac{(2 - \delta_{i,j}) \cdot \psi_{i,j} \cdot ^{mut}p_i^n \cdot ^{mut}p_j^n}{\sum_{k \leq h} (2 - \delta_{k,h}) \cdot \psi_{k,h} \cdot ^{mut}p_k^n \cdot ^{mut}p_h^n}.$$

Y como consecuencia, utilizando el hecho de que la selección es multiplicativa en la última igualdad, cualquiera de los alelos de los  $N$  individuos sobrevivientes tiene una probabilidad  $^{mut,sel}g_i^n$  de ser un alelo de tipo  $\{i\}$ :

$$\begin{aligned} ^{mut,sel}g_i^n &= \sum_{j \neq i} \frac{1}{2} F_{i,j}^n + F_{i,i}^n = \frac{\sum_j \psi_{i,j} \cdot ^{mut}p_i^n \cdot ^{mut}p_j^n}{\sum_k \sum_h \psi_{k,h} \cdot ^{mut}p_k^n \cdot ^{mut}p_h^n} \\ &= \frac{v_i \cdot ^{mut}p_i^n}{\sum_k v_k \cdot ^{mut}p_k^n}. \end{aligned}$$

6. Los individuos sobrevivientes forman urnas con los gametos que los conforman (lo que correspondería a formar las  $(n+1)$ -urnas). Para formar estas urnas, cada individuo de tipo  $\{r, s\}$  deposita en cada urna un gameto de tipo  $\{r\}$  y otro de tipo  $\{s\}$ .

Como la selección es multiplicativa, las variables que representan el tipo de los alelos son independientes entre sí, y entonces las urnas creadas por los individuos sobrevivientes tienen una distribución

$$Multinomial\left(N, \left({}^{mut,sel}g_1^n, \dots, {}^{mut,sel}g_r^n\right)\right).$$

Entonces, dada la  $n$ -urna, la  $(n+1)$ -urna se distribuye como una variable aleatoria  $Multinomial\left(N, \left({}^{sel,mut}g_1^n, \dots, {}^{sel,mut}g_r^n\right)\right)$ .

Estas probabilidades de transición definen a una cadena de Markov homogénea.

**Definición 1.7.3.** *El modelo discreto de Wright-Fisher diploide, con  $r$  tipos de alelos, con selección multiplicativa y con mutación, donde **la mutación ocurre antes que la selección**, es una cadena de Markov discreta  $\{p_n\}_{n \in \mathbb{N}} = \{(p_1^n \dots p_r^n)_n\}_{n \in \mathbb{N}}$  con espacio de estados*

$$\left\{ (a_1, \dots, a_r) \in \mathbb{N}^r \mid \sum_{i=1}^r a_i = 2N \right\}$$

y probabilidades de transición dadas por

$$\mathbb{P}\left(p^{n+1} = (a_1, \dots, a_r) \mid p^n\right) = \binom{2N}{a_1, \dots, a_r} \cdot \left({}^{mut,sel}g_1^n\right)^{a_1} \dots \left({}^{mut,sel}g_r^n\right)^{a_r}$$

$$\text{con} \quad {}^{mut,sel}g_s^n = \frac{v_i \cdot {}^{mut}p_i^n}{\sum_k v_k \cdot {}^{mut}p_k^n} = \frac{v_i \cdot (\sum_a \mu_{a,i} \cdot p_a^n)}{\sum_k v_k \cdot (\sum_s \mu_{s,k} \cdot p_s^n)}$$

$$\text{con} \quad {}^{mut}p_i^n = \sum_{s=1}^r p_s^n \cdot \mu_{s,i}$$

$$\text{y con} \quad v_s = \sqrt{\psi_{s,s}} \quad .$$

**Relación entre ambos modelos.**

Recordemos la probabilidad de obtener un alelo de tipo  $\{i\}$  de la  $(n+1)$ -urna a partir de la  $n$ -urna del modelo con la selección antes que la mutación.

$${}^{sel,mut}g_i^n = \sum_k g_k^n \cdot \mu_{k,i} = \sum_k \mu_{k,i} \cdot \frac{v_k \cdot p_k^n}{\sum_s v_s \cdot p_s^n} = \frac{\sum_k \mu_{k,i} \cdot v_k \cdot p_k^n}{\sum_s v_s \cdot p_s^n}$$

A primera vista, las probabilidades de los dos modelos son muy diferentes:

$${}^{sel,mut}g_i^n = \frac{\sum_k \mu_{k,i} \cdot v_k \cdot p_k^n}{\sum_s v_s \cdot p_s^n} \quad {}^{mut,sel}g_i^n = \frac{v_i \cdot (\sum_k \mu_{k,i} \cdot p_k^n)}{\sum_k v_k \cdot (\sum_z \mu_{z,k} \cdot p_z^n)}$$

Sin embargo, ambos modelos pueden representar al otro, haciendo un cambio en los parámetros.

Nosotros sólo demostraremos que el modelo de Wright-Fisher con la mutación antes de la selección se puede representar por el modelo de Wright-Fisher con la selección antes de la mutación, debido a que el converso es más complicado y mucho menos útil. Es más complicado porque involucra resolver dos sistemas de ecuaciones lineales, y es menos útil porque el modelo que se utiliza comúnmente, y del que hay muchos resultados ya probados es el modelo de Wright-Fisher con la selección antes de la mutación.

Para demostrar que el modelo de Wright-Fisher con la mutación antes de la selección se puede representar por el modelo de Wright-Fisher con la selección antes de la mutación, demostraremos que para cada conjunto de parámetros de población  $\{\{\mu_{i,j}\}_{i,j}, \{v_s\}_s\}$  podemos encontrar un nuevo conjunto de parámetros  $\{\{\bar{\mu}_{i,j}\}_{i,j}, \{\bar{v}_s\}_s\}$  de tal forma que el modelo de Wright-Fisher con la mutación antes de la selección y el primer conjunto de parámetros, tiene la misma distribución que el modelo de Wright-Fisher con selección primero y mutación después y el segundo conjunto de parámetros.

Lo que podemos interpretar de esta equivalencia es que en realidad no importa si la selección sucede antes o después de la mutación, siempre que se calculen de la manera apropiada los parámetros, y a la hora de calcularlos si es importante tomar en cuenta el orden.

Encontremos los parámetros correspondientes:

$$sel,mut g_i^n = mut,sel g_i^n$$

si y sólo si

$$\frac{\sum_s \overline{\mu_{s,i}} \cdot \overline{v_s} \cdot p_s^n}{\sum_j \overline{v_j} \cdot p_j^n} = \frac{v_i \cdot (\sum_k \mu_{k,i} \cdot p_k^n)}{\sum_t v_t \cdot (\sum_z \mu_{z,t} \cdot p_z^n)}$$

$$\left( \sum_s \overline{\mu_{s,i}} \cdot \overline{v_s} \cdot p_s^n \right) \cdot \left( \sum_t \sum_z v_t \cdot \mu_{z,t} \cdot p_z^n \right) = \left( \sum_j \overline{v_j} \cdot p_j^n \right) \cdot \left( v_i \cdot \sum_k \mu_{k,i} \cdot p_k^n \right)$$

$$\sum_s \sum_z \left( \overline{\mu_{s,i}} \cdot \overline{v_s} \cdot p_s^n \cdot p_z^n \cdot \left( \sum_t v_t \cdot \mu_{z,t} \right) \right) = v_i \cdot \sum_j \sum_k \overline{v_j} \cdot p_j^n \cdot \mu_{k,i} \cdot p_k^n$$

tomando en el lado derecho de la ecuación

$$k = s \quad y \quad j = z$$

obtenemos

$$v_i \cdot \sum_z \sum_s \overline{v_z} \cdot p_z^n \cdot \mu_{s,i} \cdot p_s^n$$

y despejando obtenemos :

$$\sum_s \sum_z p_s^n \cdot p_z^n \cdot \left( \overline{\mu_{s,i}} \cdot \overline{v_s} \cdot \left( \sum_t v_t \cdot \mu_{z,t} \right) - v_i \cdot \mu_{s,i} \cdot \overline{v_z} \right) = 0$$

como esta igualdad se debe dar para cualquier vector  $(p_1^n, \dots, p_r^n)$ , tenemos esto sucede si y sólo si

$$\overline{\mu_{s,i}} \cdot \overline{v_s} \cdot \sum_t v_t \cdot \mu_{z,t} - v_i \cdot \mu_{s,i} \cdot \overline{v_z} = 0$$

y si y sólo si

$$\frac{\overline{\mu_{s,i}} \cdot \overline{v_s}}{\overline{v_z}} = \frac{v_i \cdot \mu_{s,i}}{\sum_t v_t \cdot \mu_{z,t}}. \quad (1.9)$$

Con esta ecuación podemos encontrar los valores de  $\overline{\mu_{s,i}}$  a partir de los valores  $\{\{\mu_{i,j}\}_{i,j}, \{v_s\}_s\}$ , con sólo tomar  $z = s$ .

$$\overline{\mu_{s,i}} = \frac{\overline{\mu_{s,i}} \cdot \overline{v_s}}{\overline{v_s}} = \frac{v_i \cdot \mu_{s,i}}{\sum_t v_t \cdot \mu_{s,t}}.$$

Como podemos observar, ambas probabilidades  $sel,mut g_i^n$  y  $mut,sel g_i^n$  permanecen inalteradas si sustituimos los valores de  $\{v_s\}_{s=1}^r$  por valores que mantengan las mismas proporciones, como  $\{constante \cdot v_s\}_{s=1}^r$ , y por lo tanto, los modelos de Wright-Fisher son los mismos si multiplicamos todos los coeficientes de selección por la misma constante.

Tomemos a  $\bar{v}_1$  como un valor fijo, entonces podemos calcular cuanto valen los demás  $\bar{v}_i$  con respecto a  $\bar{v}_1$  y a los valores  $\{\{\mu_{i,j}\}_{i,j}, \{v_s\}_s\}$  (con los cuales además quedan determinados los valores  $\overline{\mu_{i,j}}$ ). Tomando en la ecuación (1.9)  $z = 1$  obtenemos que

$$\begin{aligned} \bar{v}_s &= \frac{v_i \cdot \mu_{s,i}}{\sum_t v_t \cdot \mu_{1,t}} \cdot \frac{\bar{v}_1}{(\overline{\mu_{s,i}})} = \frac{v_i \cdot \mu_{s,i}}{\sum_t v_t \cdot \mu_{1,t}} \cdot \frac{\sum_a v_a \cdot \mu_{s,a}}{v_i \cdot \mu_{s,i}} \cdot \bar{v}_1 \\ &= \frac{\sum_a v_a \cdot \mu_{s,a}}{\sum_t v_t \cdot \mu_{1,t}} \cdot \bar{v}_1. \end{aligned}$$

Entonces tenemos que el modelo de Wright-Fisher con mutación antes que selección, y parámetros dados por  $\{\{\mu_{i,j}\}_{i,j}, \{v_s\}_s\}$  se puede obtener a partir del modelo de Wright-Fisher con selección antes que mutación, y parámetros dados por

$$\bar{v}_s = \frac{\sum_a v_a \cdot \mu_{s,a}}{\sum_t v_t \cdot \mu_{1,t}} \cdot \bar{v}_1 \quad \overline{\mu_{s,i}} = \frac{v_i \cdot \mu_{s,i}}{\sum_t v_t \cdot \mu_{s,t}}.$$

Verifiquemos que estos parámetros si satisfacen la ecuación (1.9),

$$\begin{aligned} \frac{\overline{\mu_{s,i}} \cdot \bar{v}_s}{\bar{v}_z} &= \frac{\overbrace{v_i \cdot \mu_{s,i}}^{\overline{\mu_{s,i}}}}{\sum_t v_t \cdot \mu_{s,t}} \cdot \frac{\overbrace{\sum_a v_a \cdot \mu_{s,a}}^{\bar{v}_s}}{\sum_t v_t \cdot \mu_{1,t}} \cdot \bar{v}_1 \cdot \frac{\overbrace{\sum_t v_t \cdot \mu_{1,t}}^{\frac{1}{\bar{v}_z}}}{\sum_a v_a \cdot \mu_{z,a}} \cdot \frac{1}{\bar{v}_1} \\ &= \frac{v_i \cdot \mu_{s,i}}{\sum_t v_t \cdot \mu_{z,t}}. \end{aligned}$$



## Capítulo 2

# Análisis Sencillo de los Modelos

En este capítulo obtendremos las probabilidades de absorción de los modelos Wright-Fisher con dos alelos y con  $r$  alelos, sin mutación y sin selección, y mencionaremos cuales estados son absorbentes, recurrentes o transitorios, y cuales son los conjuntos cerrados e irreducibles.

En general es muy complicado calcular cualquier cosa en los modelos de Wright-Fisher con mutación o selección, por lo que comúnmente se utilizan aproximaciones por procesos de difusión, se calcula lo que se desea calcular para el proceso de difusión, y luego se demuestra que este cálculo es una buena aproximación.

En el libro de Ethier y Kurtz [?] capítulo 10, se desarrolla la aproximación de los modelos de Wright-Fisher por medio de procesos de difusión, y se aproximan la esperanza del tiempo de absorción, las distribuciones estacionarias y las probabilidades de absorción.

Para definir al proceso de difusión que aproxima a cada modelo, se construye una sucesión de procesos cadlåg, donde para cada  $N \in \mathbb{N}$  el proceso cadlåg preserva la información del modelo de Wright-Fisher del tipo correspondiente para una población de tamaño  $N$ , donde además, en cada unidad de tiempo transcurren  $N$  generaciones y los parámetros de mutación y selección varían con respecto a  $N$ . La sucesión se construye de tal forma que converge débilmente a un proceso de difusión.

En la sección 2 del capítulo 10 del libro de Ethier y Kurtz [?] se calculan la esperanza del tiempo de absorción, la distribución estacionaria y las probabilidades de absorción de los procesos de difusión que aproximan a los modelos, y se demuestra que tanto la sucesión de tiempos de absorción como la sucesión de distribuciones estacionarias de las cadenas de Markov correspondientes, convergen débilmente al tiempo de absorción y la distribución estacionaria de los procesos de difusión que

las aproximan.

En el capítulo 3 mostraremos este mecanismo en el caso más sencillo, para el modelo de Wright-Fisher haploide con dos alelos sin mutación y sin selección. Encontraremos al proceso de difusión que lo aproxima y calcularemos su esperanza de tiempo de absorción.

No enunciaremos los resultados importantes de los modelos de Wright-Fisher (a excepción de las probabilidades de absorción de los modelos sin selección y sin mutación, con  $r$  y con dos alelos, y la esperanza del tiempo de absorción del de dos alelos), pero como hemos dicho ya, estos resultados se pueden consultar en la sección 2 del capítulo 10 del libro de Ethier y Kurtz [?].

En esta sección calcularemos únicamente las probabilidades de absorción de los modelos sin selección y sin mutación, con  $r$  alelos y con dos alelos.

Utilizaremos en este capítulo la notación y los teoremas del libro [HPS] “Introduction to Stochastic Processes” de los autores Paul G. Hoel, Sidney C. Port y Charles J. Stone.

Los cálculos de este capítulo fueron desarrollados en su totalidad por el autor de esta tesis.

## 2.1. Modelo Haploide con 2 Alelos

$Y_n$  = número de  $n$ -individuos de tipo *uno* ( $\alpha$ ).

$$Y_{n+1} \Big|_{Y_n} \sim \text{Binomial}(N, Y_n)$$

Es decir, la variable  $Y_{n+1}$  dada la variable  $Y_n$  se distribuye como una variable aleatoria  $\text{Binomial}(N, Y_n)$ .

Espacio de estados de  $Y_n = \{0, 1, 2, \dots, N\}$

$$\mathbb{P}(Y_{n+1} = j | Y_n) = \binom{N}{j} \left(\frac{Y_n}{N}\right)^j \left(1 - \frac{Y_n}{N}\right)^{N-j}. \quad (2.1)$$

### 2.1.1. Clasificación de Estados

**Teorema 2.1.1.** *Del espacio de estados de  $Y_n$ ,  $\{0, 1, \dots, N\}$ , los estados 0 y  $N$  son estados absorbentes y por lo tanto recurrentes positivos. Desde los estados*

$\{1, \dots, N-1\}$  se puede acceder a 0 y a  $N$ , y por lo tanto son estados transitorios.

- 0 es un estado absorbente.  $\mathbb{P}(Y_{n+1} = 0 \mid Y_n = 0) = 1$
- $N$  es un estado absorbente.  $\mathbb{P}(Y_{n+1} = N \mid Y_n = N) = 1$
- Sea  $i \in \{1, \dots, N-1\}$ , se puede acceder a 0 desde  $i$  :

$$\mathbb{P}(Y_{n+1} = 0 \mid Y_n = i) = \binom{N}{0} \left(\frac{i}{N}\right)^0 \left(1 - \frac{i}{N}\right)^{N-0} = \left(1 - \frac{i}{N}\right)^N > 0 \quad (2.2)$$

- Sea  $i \in \{1, \dots, N-1\}$ , se puede acceder a  $N$  desde  $i$  :

$$\mathbb{P}(Y_{n+1} = N \mid Y_n = i) = \binom{N}{N} \left(\frac{i}{N}\right)^N \left(1 - \frac{i}{N}\right)^{N-N} = \left(\frac{i}{N}\right)^N > 0 \quad (2.3)$$

A partir del **Teorema 2** de la sección **1.6** del libro [HPS] de Hoel, Port y Stone, se concluye que si  $y$  es un estado absorbente entonces es recurrente, y también se deduce que si se puede acceder a  $y$  desde  $x$ , entonces  $x$  es un estado transitorio.  $\square$

### 2.1.2. Probabilidades de Absorción

Utilizaremos el **Teorema 5** de la sección **1.6.1** del libro [HPS].

**Teorema 2.1.2** (Teorema 5 de la sección **1.6.1** del libro [HPS]).

*Supongamos que el conjunto de estados transitorios ( $S_T$ ) es finito, sea  $C$  un conjunto cerrado e irreducible de estados recurrentes. La probabilidad de absorción de la cadena al conjunto cerrado  $C$  empezando desde un estado  $x$  se puede escribir como:*

$$\rho_C(x) = \sum_{w \in C} \mathbb{P}(x, w) + \sum_{y \in S_T} \mathbb{P}(x, y) \cdot \rho_C(y). \quad \text{“ ecuación } (x) \text{ ”}$$

*Tenemos de esta forma un sistema de  $|S_T|$  ecuaciones ( una “ecuación  $(x)$ ” para cada  $x \in S_T$  ) con  $|S_T|$  incógnitas (  $\rho_C(x)$  para cada  $x \in S_T$  ).*

*El teorema dice que este sistema lineal de ecuaciones*

$$f(x) = \sum_{w \in C} \mathbb{P}(x, w) + \sum_{y \in S_T} \mathbb{P}(x, y) \cdot f(y) \quad \forall x \in S_T$$

*tiene una única solución.*

Tenemos que  $\{N\}$  es un conjunto cerrado, irreducible y recurrente, por lo que la probabilidad de que la cadena  $Y_n$  se absorba en  $\{N\}$  dado que la cadena empieza en  $x$  ( $\rho_C(x)$ ) la podemos obtener resolviendo el sistema de ecuaciones correspondiente:

“Ecuación (x)”

$$\begin{aligned} f(x) &= \mathbb{P}(x, N) + \sum_{y \in \{1, \dots, N-1\}} \mathbb{P}(x, y) \cdot f(y) \quad \forall x \in \{1, \dots, N-1\} \\ f(x) &= \left(\frac{x}{N}\right)^N + \sum_{y=1}^{N-1} \binom{N}{y} \left(\frac{x}{N}\right)^y \left(1 - \frac{x}{N}\right)^{N-y} \cdot f(y) \end{aligned} \quad (2.4)$$

Ordenando todas las variables ( $f(z)$ ) de un mismo lado y los términos constantes del otro lado,

$$-\left(\frac{x}{N}\right)^N = (-1) \cdot f(x) + \sum_{y=1}^{N-1} \binom{N}{y} \left(\frac{x}{N}\right)^y \left(1 - \frac{x}{N}\right)^{N-y} \cdot f(y).$$

Demos explícitamente los coeficientes de este sistema de ecuaciones. El primer subíndice del coeficiente corresponde al número de ecuación, y el segundo, al número de variable al que el coeficiente está multiplicando. La numeración de las ecuaciones y de las variables se da de manera natural debido al tipo de conjunto  $S_T = \{1, 2, \dots, N-1\}$  que tenemos.

$$\begin{aligned} A_{x,y} &= \begin{cases} \binom{N}{y} \left(\frac{x}{N}\right)^y \left(1 - \frac{x}{N}\right)^{N-y} & \text{si } x \neq y \\ \binom{N}{x} \left(\frac{x}{N}\right)^x \left(1 - \frac{x}{N}\right)^{N-x} - 1 & \text{si } x = y \end{cases} \\ b_x &= -\left(\frac{x}{N}\right)^N \end{aligned}$$

La solución de este sistema de ecuaciones se puede calcular con una computadora para valores de  $N$  muy chicos. La solución está dada por la regla de Cramer con estos coeficientes, la matriz  $A = (A_{i,j})_{i,j}$  y el vector  $b = (b_i)_i$ .

Pero también podemos resolver este sistema analíticamente para toda  $N$ .

En los próximos párrafos, nos conviene tomar el coeficiente  $A_{x,x}$  sin el  $(-1)$ . Para todo  $y$  y todo  $x$ , llamaremos  $a_{x,y}$  al coeficiente

$$a_{x,y} = \binom{N}{y} \left(\frac{x}{N}\right)^y \left(1 - \frac{x}{N}\right)^{N-y}.$$

Recordemos cual es el sistema del cual buscamos la única solución.

$$f(x) = \left(\frac{x}{N}\right)^N + \sum_{y=1}^{N-1} \binom{N}{y} \left(\frac{x}{N}\right)^y \left(1 - \frac{x}{N}\right)^{N-y} \cdot f(y),$$

$$f(x) = b_x + \sum_{y=1}^{N-1} a_{x,y} \cdot f(y).$$

Recordemos también que

$$a_{x,y} = \mathbb{P}(x, y) = \mathbb{P}\left(\text{Binomial}\left(N, \frac{x}{N}\right) = y\right)$$

$$b_x = \mathbb{P}(x, N) = \mathbb{P}\left(\text{Binomial}\left(N, \frac{x}{N}\right) = N\right)$$

Agreguemos la hipótesis de que  $f(N) = 1$  y  $f(0) = 0$ . Podemos agregar esta hipótesis ya que  $\rho_{\{N\}}(N) = 1$  y  $\rho_{\{N\}}(0) = 0$ , y cualquier solución del sistema es igual a  $\rho_{\{N\}}$ .

Con lo dicho en estos últimos párrafos, la parte derecha de la ecuación del sistema se puede escribir como la esperanza de una binomial evaluada en  $f$ ,

$$f(x) = b_x \cdot f(N) + \sum_{y=1}^{N-1} a_{x,y} \cdot f(y)$$

$$= f(N) \cdot \mathbb{P}\left(\text{Bin}\left(N, \frac{x}{N}\right) = N\right) + \sum_{y=1}^{N-1} f(y) \cdot \mathbb{P}\left(\text{Bin}\left(N, \frac{x}{N}\right) = y\right)$$

$$= \mathbb{E}\left[f\left(\text{Binomial}\left(N, \frac{x}{N}\right)\right)\right].$$

Al ser  $\mathbb{E}\left(\text{Binomial}\left(N, \frac{x}{N}\right)\right) = N \cdot \frac{x}{N}$  se puede ver fácilmente que  $f(y) = \frac{y}{N}$  es la solución.

$$f(x) = \sum_{y=1}^N \binom{N}{y} \left(\frac{x}{N}\right)^y \left(1 - \frac{x}{N}\right)^{N-y} \cdot \frac{y}{N}$$

lo cual es igual a

$$f(x) = \mathbb{E}\left(\text{Binomial}\left(N, \frac{x}{N}\right)\right) \cdot \frac{1}{N}$$

$$= N \frac{x}{N} \cdot \frac{1}{N} = \frac{x}{N}$$

lo cual prueba que  $f(y) = \frac{y}{N}$  es una solución al sistema. Como el sistema tiene una única solución y las probabilidades de absorción son una solución,  $f(y) = \frac{y}{N}$

es la probabilidad de absorción de la cadena al cerrado irreducible recurrente  $\{N\}$  empezando desde  $y$ .

$$\rho_{\{N\}}(x) = \frac{x}{N} \quad \forall x \in S_T.$$

Al ser  $\{N\}$  y  $\{0\}$  los únicos dos conjuntos cerrados, recurrentes e irreducibles en una cadena con una cantidad finita de estados,

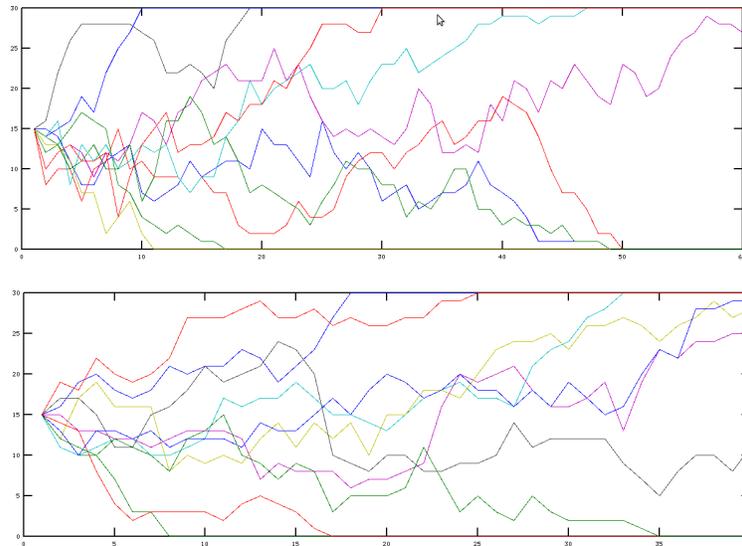
$$\rho_{\{0\}}(x) = 1 - \rho_{\{N\}}(x) = \frac{N - x}{N} \quad \forall x \in S_T.$$

### 2.1.3. Simulación

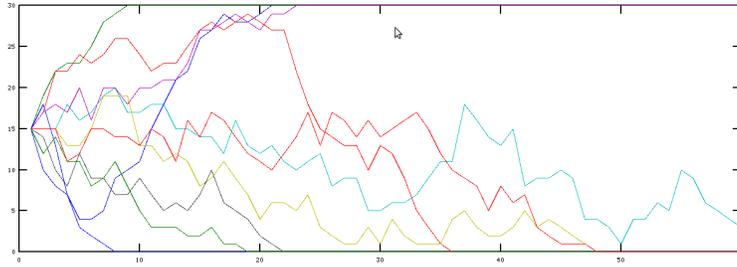
#### Población de 30 individuos a lo largo de 60 generaciones

En esta subsección agregamos imágenes de simulaciones de la cadena de Wright-Fisher de una población de 30 individuos, sin mutación, sin selección, con 2 tipos de alelos a lo largo de 60 generaciones, empezando desde diversos estados.

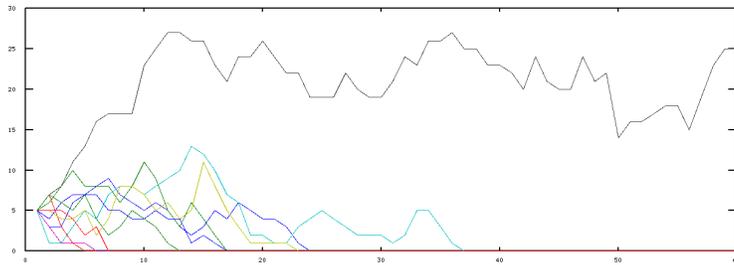
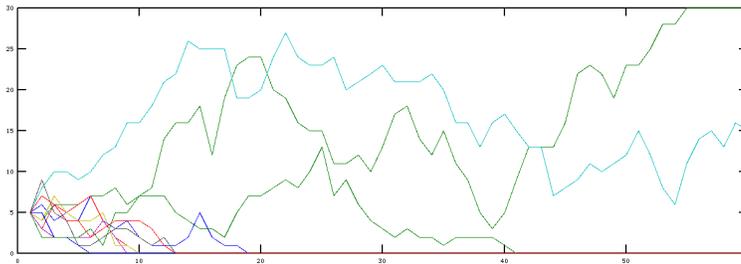
#### Empezando con 15 individuos de tipo $\{1\}$



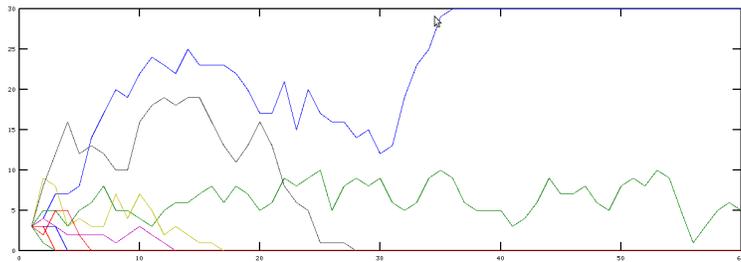
2.1. HAPLOIDE CON 2 ALELOS

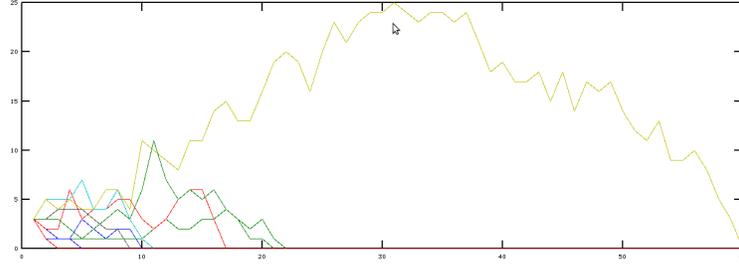


**Empezando con 5 individuos de tipo {1}**



**Empezando con 3 individuos de tipo {1}**





## 2.2. Modelo Haploide con r Alelos

Al vector cuyas entradas son todas 0, excepto por la  $i$ -ésima entrada, la cual es igual a 1, le llamamos  $e_i$ .

$x_n^i$  = número de  $n$ -individuos de tipo  $\{i\}$

$X_n = (x_1^n, x_2^n, \dots, x_r^n)$  = cantidades de individuos de cada tipo de la población.

$X_{n+1} \Big|_{X_n} \sim \text{Multinomial} \left( N, \left( \frac{x_1^n}{N}, \frac{x_2^n}{N}, \dots, \frac{x_r^n}{N} \right) \right)$

Espacio de estados de  $X_n = \{(a_1, \dots, a_r) \in \mathbb{N}^r \mid \sum_{i=1}^r a_i = N\}$

$$\mathbb{P} \left( X_{n+1} = (a_1, \dots, a_r) \mid X_n = (b_1, \dots, b_r) \right) = \binom{N}{a_1, \dots, a_r} \left( \frac{b_1}{N} \right)^{a_1} \cdots \left( \frac{b_r}{N} \right)^{a_r}$$

### 2.2.1. Clasificación de Estados

Los estados absorbentes de  $X_n$  son los vectores de la forma  $N \cdot e_i = (0, \dots, N, \dots, 0)$ , con una entrada  $N$  y el resto 0; estos son los estados en los que todos los individuos son del mismo tipo. Al ser absorbentes, son estados recurrentes positivos.

El resto de los estados son transitorios, ya que desde cualquier estado se puede acceder a un estado absorbente de tipo  $N \cdot e_i$  para algún  $i$ .

Tomemos un estado  $b = (b_1, \dots, b_r)$ . Como  $\sum_{i=1}^r b_i = N$ , algún  $b_s > 0$ , por lo tanto, se puede acceder desde  $b$  al estado  $N \cdot e_s$ , el estado en el que todos los individuos son de tipo  $s$ :

$$\mathbb{P} \left( X_{n+1} = (\overbrace{0, 0, \dots, 0}^{s-1}, N, 0, \dots, 0) \mid X_n = (b_1, \dots, b_r) \right) = \left( \frac{b_s}{N} \right)^N > 0$$

### 2.2.2. Probabilidades de Absorción

Como todos los estados comunican con algún estado absorbente  $N \cdot e_i$ , los únicos conjuntos irreducibles son los conjuntos  $\{N \cdot e_i\} = \{(0, \dots, N, \dots, 0)\}$ .

Calcularemos las probabilidades de absorción desde un estado transitorio  $b = (b_1, \dots, b_r)$  al estado  $N \cdot e_{i_0}$ , es decir, calcularemos  $\rho_{\{N e_{i_0}\}}(b)$ .

• Utilizando el Teorema 2.1.2 en la página 45, sabemos que  $(\rho_{\{N e_i\}}(b))_{b \in S_T}$  es la única solución al sistema de  $|S_T|$  ecuaciones

$$f(x) = \mathbb{P}(x, N \cdot e_{i_0}) + \sum_{y \in S_T} \mathbb{P}(x, y) \cdot f(y) \quad \forall x \in S_T \quad (2.5)$$

• En esencia, el modelo “W-F haploide con  $r$  alelos” se parece a lo que sería un conjunto de  $r$  modelos “W-F haploides con **dos** alelos” que se intersectan, donde los individuos son de ‘el’ tipo del modelo de dos alelos, o son de cualquier otro tipo. Esto se notará particularmente a continuación.

Sea  $B_j^k =$  el conjunto de estados que tienen exactamente  $k$  individuos en la  $j$ -ésima coordenada,

$$\begin{aligned} B_j^k &= \left\{ a \in \mathbb{N}^r \mid a_j = k, \sum_{i=1}^r a_i = N \right\} \\ &= \left\{ a = (a_1, \dots, a_{j-1}, k, a_{j+1}, \dots, a_r) \mid \sum_{i=1}^r a_i = N \right\} \end{aligned}$$

Dada una  $j$  fija, los conjuntos  $B_j^k$  son disjuntos, y su unión es el espacio de estados completo. Son disjuntos ya que dados dos conjuntos distintos  $B_j^s$  y  $B_j^r$  ( $s \neq r$ ), cualquier vector  $x \in B_j^s$  tiene como  $j$ -ésima coordenada al valor  $s$ , y por lo tanto,  $x \notin B_j^r$ , pues no tiene al valor  $r$  en su  $j$ -ésima coordenada; y su unión es el espacio de estados completo, pues cualquier vector  $x = (x_1, \dots, x_r) \in B_j^{x_j}$ .

$B_{i_0}^N = \{N \cdot e_{i_0}\}$  es justo el conjunto del cual buscamos las probabilidades de absorción, y es cerrado, recurrente e irreducible.

Si  $1 \leq k \leq N-1$  entonces cualquier estado en el conjunto  $B_{i_0}^k$  es transitorio.

Debido a estas características, la ecuación (2.5) se puede reescribir así:

$$\begin{aligned} f(x) &= \mathbb{P}(x, N \cdot e_{i_0}) + \sum_{k=0}^{N-1} \sum_{y \in B_{i_0}^k \cap S_T} \mathbb{P}(x, y) \cdot f(y) \\ &= \mathbb{P}(x, N \cdot e_{i_0}) + \sum_{k=1}^{N-1} \sum_{y \in B_{i_0}^k} \mathbb{P}(x, y) \cdot f(y) + \sum_{y \in B_{i_0}^0 \cap S_T} \mathbb{P}(x, y) \cdot f(y) \end{aligned}$$

$$\forall x \in S_T$$

- Ningún estado en  $B_{i_0}^0$  se comunica con  $B_{i_0}^N$ .  
Si  $a \in B_{i_0}^0$ , es decir  $a_{i_0} = 0$ , no se puede acceder a  $N \cdot e_{i_0}$  desde  $a$ :

En el estado “ $a$ ” ningún individuo es de tipo  $\{i_0\}$ , por lo que ningún hijo puede ser de tipo  $\{i_0\}$ , y entonces no se puede pasar a ningún estado donde el número de individuos de tipo  $\{i_0\}$  sea mayor que 0.

$$\begin{aligned} &\mathbb{P}\left(X_{n+1} = (b_1, \dots, b_r) \text{ con } b_{i_0} > 0 \mid X_n = (a_1, \dots, a_r) \text{ con } b_{i_0} = 0\right) \\ &= \binom{N}{b_1, \dots, b_r} \left(\frac{a_1}{N}\right)^{b_1} \cdots \left(\frac{0}{N}\right)^{b_{i_0}} \cdots \left(\frac{a_r}{N}\right)^{b_r} = 0 \end{aligned}$$

Se repetirá lo mismo para las siguientes generaciones, por lo que a partir de “ $a$ ” todos los estados tendrán 0 en su  $i_0$ -ésima coordenada. Lo cual implica que:

$$\rho_{\{N \cdot e_{i_0}\}}(a) = \mathbb{P}_a(T_{N \cdot e_{i_0}} < \infty) = 0.$$

Como la probabilidad de pasar en una generación de  $a_{i_0}$  a  $N \cdot e_{i_0}$  es positiva, sabemos que si  $a_{i_0} > 0$  para el estado  $a = (a_1, \dots, a_r)$ , se puede acceder a  $N \cdot e_{i_0}$  desde  $a$ , y por lo tanto  $a \in S_T$ .

Lo que hemos visto en la discusión de este último inciso, es que el conjunto de estados desde los cuales **no** se puede acceder a  $N \cdot e_{i_0}$ , coincide con  $B_{i_0}^0$ , el conjunto de estados cuya  $i_0$ -ésima coordenada es 0.

$$B_{i_0}^0 = \{a \in \mathbb{N}^r \mid a_{i_0} = 0\} = \{a \in \mathbb{N}^r \mid \rho_{\{N \cdot e_{i_0}\}}(a) = 0\}.$$

Como el sistema (2.5) tiene como única solución a las probabilidades de absorción a  $N \cdot e_{i_0}$ , es decir  $\rho_{\{N \cdot e_{i_0}\}}(x)$ , tenemos que  $f(x) = \rho_{\{N \cdot e_{i_0}\}}(x)$ . En particular, si  $x \in B_{i_0}^0$  entonces  $f(x) = 0$ . De esta forma el sistema (2.5) se reduce a :

$$\begin{aligned} f(x) &= \mathbb{P}(x, N \cdot e_{i_0}) + \sum_{k=1}^{N-1} \sum_{y \in B_{i_0}^k} \mathbb{P}(x, y) \cdot f(y) \\ &\forall x \in S_T \end{aligned} \tag{2.6}$$

• La cantidad de individuos de tipo  $i_0$  en la siguiente generación depende únicamente de la cantidad de individuos de tipo  $i_0$  en esta generación, por lo que  $\rho_{\{N \cdot e_{i_0}\}}(x)$  debería ser igual para todos los estados que tuvieran el mismo valor en la entrada  $i_0$  (cantidad de individuos tipo  $i_0$ ), es decir, debería ser constante sobre los conjuntos  $B_{i_0}^k$ . Si suponemos ésto, el sistema de ecuaciones se simplifica notablemente.

Sea  $y^{*k}$  un elemento de  $B_{i_0}^k$ , y **supongamos** que  $f(y) = f(y^{*k}) \forall y \in B_{i_0}^k$ .  
Tenemos que:

$$\begin{aligned}
& \sum_{y \in B_{i_0}^k} \binom{N}{y_1, \dots, k, \dots, y_r} \left(\frac{x_1}{N}\right)^{y_1} \cdots \left(\frac{x_r}{N}\right)^{y_r} \cdot f(y^{*k}) \\
&= f(y^{*k}) \cdot \sum_{y' \in \mathbb{N}^{r-1}} \binom{N-k}{y'_1, \dots, y'_{r-1}} \cdot \binom{N}{N-k, k} \left(\frac{x_1}{N}\right)^{y'_1} \cdots \left(\frac{x_r}{N}\right)^{y'_{r-1}} \cdot \left(\frac{x_{i_0}}{N}\right)^k \\
&= f(y^{*k}) \cdot \left(\frac{x_{i_0}}{N}\right)^k \cdot \binom{N}{N-k, k} \cdot \sum_{y' \in \mathbb{N}^{r-1}} \binom{N-k}{y'_1, \dots, y'_{r-1}} \left(\frac{x_1}{N}\right)^{y'_1} \cdots \left(\frac{x_r}{N}\right)^{y'_{r-1}} \\
&= f(y^{*k}) \cdot \left(\frac{x_{i_0}}{N}\right)^k \cdot \binom{N}{N-k, k} \cdot \left(\frac{x_1}{N} + \cdots + \frac{x_{i_0-1}}{N} + \frac{x_{i_0+1}}{N} + \cdots + \frac{x_r}{N}\right)^{N-k} \\
&= f(y^{*k}) \cdot \left(\frac{x_{i_0}}{N}\right)^k \cdot \binom{N}{N-k, k} \cdot \left(1 - \frac{x_{i_0}}{N}\right)^{N-k}
\end{aligned}$$

Este procedimiento es válido para todos los  $k$ , para todos los  $B_{i_0}^k$ .

Sustituyendo en el sistema (2.6) lo obtenido en el párrafo de ecuaciones anterior, obtenemos la ecuación:

$$\begin{aligned}
f(x) &= \mathbb{P}(x, N \cdot e_{i_0}) + \sum_{k=1}^{N-1} \sum_{y \in B_{i_0}^k} \mathbb{P}(x, y) \cdot f(y) \\
f(x) &= \left(\frac{x_{i_0}}{N}\right)^N \cdot f(y^{*N}) + \sum_{k=1}^{N-1} f(y^{*k}) \cdot \binom{N}{k} \cdot \left(\frac{x_{i_0}}{N}\right)^k \cdot \left(1 - \frac{x_{i_0}}{N}\right)^{N-k} \\
&= \sum_{k=1}^N \binom{N}{k} \cdot \left(\frac{x_{i_0}}{N}\right)^k \cdot \left(1 - \frac{x_{i_0}}{N}\right)^{N-k} \cdot f(y^{*k})
\end{aligned}$$

En el penúltimo renglón utilizamos el hecho de que  $f(y) = \rho_{\{N \cdot e_{i_0}\}}(y)$  es la probabilidad de absorción a  $\{N \cdot e_{i_0}\}$ , y entonces sabemos que  $f(y^{*N}) = 1$  pues  $y^{*N} \in B_{i_0}^N = \{N \cdot e_{i_0}\}$  y por lo tanto  $y^{*N} = N \cdot e_{i_0}$  que es el estado absorbente del cual estamos calculando la probabilidad de absorción ( $\rho_{\{N \cdot e_{i_0}\}}(N \cdot e_{i_0}) = 1$ ).

Este sistema es muy parecido al sistema de la sección anterior, al sistema del modelo de Wright-Fisher haploide con 2 alelos. La única diferencia es el número

de ecuaciones y el número de incógnitas. Por cada  $k$  tenemos  $|B_{i_0}^k|$  incógnitas (que hemos supuesto que son iguales) y  $|B_{i_0}^k|$  ecuaciones idénticas.

Inspirados en esta simetría proponemos  $f(y) = \frac{y_{i_0}}{N}$ . Esto también se puede escribir como  $f(y) = \frac{k}{N}$  para toda  $y \in B_{i_0}^k$ . Algo muy importante de esta proposición, es que respeta la suposición de que  $f(x)$  es constante sobre cada conjunto  $B_{i_0}^k$ . Evaluemos si esta función es solución:

$$\begin{aligned} f(x) &= \sum_{k=1}^N \binom{N}{k} \cdot \left(\frac{x_{i_0}}{N}\right)^k \cdot \left(1 - \frac{x_{i_0}}{N}\right)^{N-k} \cdot \frac{k}{N} \\ &= \mathbb{E}\left(\text{Binomial}\left(N, \frac{x_{i_0}}{N}\right)\right) \cdot \frac{1}{N} = N \cdot \frac{x_{i_0}}{N} \cdot \frac{1}{N} = \frac{x_{i_0}}{N}. \end{aligned}$$

Esta última ecuación nos dice que la definición que dimos de  $f(y)$  es una solución del sistema, y por lo tanto, las probabilidades de absorción están dadas por:

$$\rho_{\{N \cdot e_{i_0}\}}(x) = \frac{x_{i_0}}{N} \quad \forall \quad 1 \leq i_0 \leq r.$$

## 2.3. Modelo Haploide con 2 Alelos y Mutación

$M_n$  = número de  $n$ -individuos de tipo *uno*.

$\mu_{1,2}$  = Probabilidad de que un individuo de tipo *uno* tenga un hijo tipo *dos*.

Probabilidad de Mutar de tipo *uno* a tipo *dos*.

$\mu_{2,1}$  = Probabilidad de que un individuo de tipo *dos* tenga un hijo tipo *uno*.

Probabilidad de Mutar de tipo *dos* a tipo *uno*.

$p_n^*$  = Probabilidad de un  $(n+1)$ -individuo de ser de tipo *uno* dado  $M_n$ .

$$M_{n+1} \Big|_{M_n} \sim \text{Binomial}(N, p_n^*)$$

Espacio de estados de  $M_n = \{0, 1, 2, \dots, N\}$

$$p_n^* = (1 - \mu_{1,2}) \cdot \frac{M_n}{N} + \mu_{2,1} \cdot \left(1 - \frac{M_n}{N}\right) \quad (2.7)$$

$$\mathbb{P}\left(M_{n+1} = k \mid M_n\right) = \binom{N}{k} (p_n^*)^k \cdot (1 - p_n^*)^{N-k} \quad (2.8)$$

### 2.3.1. Caso uno, $\mu_{1,2}$ y $\mu_{2,1}$ mayores a 0

Empecemos con el caso en el que  $\mu_{1,2}$  y  $\mu_{2,1}$  son ambas mayores a 0.

En este caso  $M_n$  es una cadena *irreducible*, todos los estados de  $M_n$  se comunican entre sí. Es decir, para todo  $u, v \in \{0, 1, \dots, N\}$  tenemos que  $u$  es accesible desde  $v$ , y  $v$  es accesible desde  $u$ .

Esto se demuestra muy fácilmente debido a que la probabilidad de pasar directamente de cualquier estado a cualquier estado, resulta siempre mayor que cero.

$$\begin{aligned} \mathbb{P}\left(M_{n+1} = k \mid M_n = z\right) &= \binom{N}{k} \left( (1 - \mu_{1,2}) \cdot \frac{z}{N} + \mu_{2,1} \cdot \left(1 - \frac{z}{N}\right) \right)^k \\ &\quad \cdot \left( 1 - (1 - \mu_{1,2}) \cdot \frac{z}{N} - \mu_{2,1} \cdot \left(1 - \frac{z}{N}\right) \right)^{N-k} \\ &> 0. \end{aligned}$$

Sin importar como sea  $z$ , siempre tenemos alguna de las dos desigualdades,  $\frac{z}{N} > 0$  o  $(1 - \frac{z}{N}) > 0$ . También tenemos que  $((1 - \mu_{1,2}) \cdot \frac{z}{N} + \mu_{2,1} \cdot (1 - \frac{z}{N}))$  es siempre estrictamente menor a 1, por lo que ambos factores de la multiplicación son estrictamente mayores a 0.

Como el espacio de estados  $\{0, \dots, N\}$  es finito, y  $M_n$  es irreducible, de acuerdo al Corolario 3 de la sección 2.4 del libro de Hoel, Port y Stone [HPS], todos los estados de la cadena son estados recurrentes positivos.

De acuerdo al Teorema 5 de la sección 2.5 de “Hoel Port Stone”, debido a que el espacio de estados de  $M_n$  es irreducible,  $M_n$  tiene una única distribución estacionaria a la que llamaremos  $\pi$ ; y de acuerdo al Corolario 6 de la misma sección,  $\forall x \in \{0, \dots, N\}$ , con probabilidad 1 sucede que

$$\lim_{n \rightarrow \infty} \frac{N_n(x)}{n} = \pi(x) \quad \text{con} \quad N_n(x) = \sum_{i \leq n} \begin{cases} 1 & \text{si } M_i = x \\ 0 & \text{si } M_i \neq x. \end{cases}$$

### 2.3.2. Caso dos, $\mu_{1,2}$ mayor a 0 y $\mu_{2,1} = 0$

En este caso, el estado 0 es absorbente, y es el único, porque al ser  $\mu_{1,2} > 0$ , todos los estados, incluso el estado  $N$ , comunican con el estado 0.

Entonces tenemos que con probabilidad 1 la cadena se absorbe en el estado 0 en un tiempo finito, y también tenemos que, la única distribución estacionaria de la cadena es  $\delta_0$ .

## 2.4. Modelo Diploide con r Alelos, Selección Multiplicativa y Mutación

Este modelo tiene muchas ramas. Los parámetros de mutación determinan cómo se comunican los estados, y por lo tanto, determinan qué subconjuntos de estados serán recurrentes, cuáles serán absorbentes y si hay o no una única distribución estacionaria. En cambio, los parámetros de selección únicamente aumentan o disminuyen los valores de las probabilidades de absorción o de las probabilidades estacionarias, sin cambiar la estructura de la cadena.

Si todos los parámetros de mutación son 0, entonces, al igual que en la sección 2.2, los estados absorbentes son los estados  $(a_1, \dots, a_r)$  que tienen una única entrada distinta de 0, y los llamamos  $N \cdot e_i$ . Como el resto de los estados comunican con uno de estos estados absorbentes, entonces son transitorios. La cadena se absorbe en alguno de los estados absorbentes en un tiempo finito con probabilidad 1, y se podría calcular las probabilidades de absorción de cada uno de ellos.

Si todos los parámetros de mutación son mayores que 0, entonces, al igual que en la sección 2.3.1, la cadena es *irreducible*, todos los estados se comunican entre sí, son recurrentes y existe una única distribución estacionaria.

Analicemos brevemente qué sucede si algunos parámetros de mutación son mayores que 0 y otros son iguales a 0.

- Si para cierta  $s$  todos los parámetros  $\{\mu_{s,i}\}_i$  son 0, entonces el estado  $N \cdot e_s$  es un estado absorbente.
- Si para cierta  $s$  todos los parámetros  $\{\mu_{i,s}\}_i$  son 0, y algún parámetro  $\mu_{s,k}$  es mayor que 0, entonces el estado  $N \cdot e_s$  es un estado transitorio, al igual que cualquier estado que cuya coordenada  $s$  sea mayor que 0.
- Si el parámetro  $\mu_{i,j}$  es mayor a 0, entonces se puede acceder al estado  $N \cdot e_j$  desde el estado  $N \cdot e_i$ .
- Si hay un conjunto de parámetros de mutación de los cuales se podría decir que forman un conjunto cerrado e irreducible, por ejemplo, si  $\mu_{1,2}, \mu_{2,3}, \mu_{3,1}$  son todos mayores que 0, y todos los parámetros  $\{\mu_{1,i}\}_{i \neq 2}, \{\mu_{2,i}\}_{i \neq 3}, \{\mu_{3,i}\}_{i \neq 1}$  son iguales a 0, entonces el conjunto de vectores cuyas únicas entradas distintas de cero son las que aparecen en los subíndices de los parámetros de mutación, es un conjunto de estados cerrado e irreducible, y por lo tanto, recurrente.

En el caso del ejemplo, esto quiere decir que el conjunto

$$\{(a_1, a_2, a_3, 0, 0, \dots, 0) \mid a_1 + a_2 + a_3 = N\}$$

es un conjunto cerrado e irreducible de estados recurrentes.

#### 2.4. DIPLOIDE CON R ALELOS, SELECCIÓN MULTIPLICATIVA Y MUTACIÓN<sup>57</sup>

- En general se formarán varios subconjuntos cerrados e irreducibles y por lo tanto recurrentes. Y todos los subconjuntos cerrados, irreducibles y recurrentes se pueden describir como el conjunto de vectores cuyas entradas  $a, b, c, \dots$  son iguales a 0.



## Capítulo 3

# Convergencia Débil a un Proceso de Difusión

En muchos modelos hay cantidades importantes que son muy difíciles de encontrar, pero que se pueden aproximar por cantidades encontradas de una manera más sencilla, en modelos continuos muy parecidos al modelo original. Éste es el caso de la esperanza del tiempo de absorción del modelo de Wright-Fisher haploide con dos tipos de alelos.

Este capítulo está dedicado a construir un proceso en tiempo continuo que aproxime muy bien al modelo de Wright-Fisher haploide con dos tipos de alelos y sin mutación ni selección para poblaciones que tengan una gran cantidad de individuos.

En el capítulo 1 en la sección 1.1 definimos, para cualquier  $N$ , a la cadena  $Y$ , que representaba al modelo de Wright-Fisher haploide con dos tipos de alelos y una cantidad constante  $N$  de individuos. Vamos a utilizar la sucesión de cadenas de Markov que se forma a partir de estos modelos utilizando la  $N$  como índice de la sucesión, por lo tanto, necesitamos una nueva notación. Llamémosle  $X^N$  a la cadena de Wright-Fisher haploide con dos tipos de alelos y con una población de tamaño constante  $N$ .

Cualquier cadena de esta sucesión de cadenas  $\{X^N\}_{N \in \mathbb{N}}$  tiene la distribución cuyas probabilidades de transición están dadas por la ecuación (1.1.1) de la página 5 :

$$\mathbb{P}\left(X_{n+1}^N = j \mid X_n^N\right) = \binom{N}{j} \left(\frac{X_n^N}{N}\right)^j \left(1 - \frac{X_n^N}{N}\right)^{N-j} \quad j = 0, \dots, N.$$

La cadena que representa la proporción de individuos que son de tipo uno,  $\frac{X^N}{N}$ , contiene la misma información que la cadena  $X^N$ , sigue siendo una cadena

### 60CAPÍTULO 3. CONVERGENCIA DÉBIL A UN PROCESO DE DIFUSIÓN

de Markov, es Martingala y se mantiene acotada entre 0 y 1 cuando  $N$  tiende a infinito.

A esta nueva sucesión de cadenas  $\{\frac{X^N}{N}\}_{N \in \mathbb{N}}$  se le puede asociar una sucesión de procesos estocásticos en tiempo continuo que conserva, para cada  $N$ , la información de  $X^N$ , y además converge débilmente a un proceso de difusión cuando  $N$  tiende a infinito.

Dada esta convergencia, el proceso límite resulta ser una excelente aproximación continua al modelo  $\frac{X^N}{N}$  siempre que  $N$  sea suficientemente grande. Utilizaremos este límite para aproximar la esperanza del tiempo de absorción de las cadenas  $X^N$ .

Empezamos por definir los procesos continuos que preservan la información de  $X^N$ .

**Definición 3.0.1.** Sea  $X^N$  una cadena de Wright-Fisher haploide con dos tipos de alelos, definimos a  $Y^N$  como el proceso estocástico:

$$Y_t^N = \frac{X_{[Nt]}^N}{N} = \frac{X_n^N}{N} \quad \text{con } n = [Nt]$$

Definiendo a  $Y^N$  de esta forma,  $Y^N$  tiene trayectorias continuas por la derecha con límites por la izquierda, es constante en los pedazos de tiempo  $[aN, (a+1)N)$  con  $a \in \mathbb{N}$ , y en ellos toma el valor de su cadena asociada  $\frac{X_a^N}{N}$ . Cada una de sus realizaciones pertenece al espacio de funciones reales definidas sobre  $[0, \infty)$  y continuas por la derecha con límites por la izquierda (càdlag).

En las primeras tres secciones demostramos que la sucesión de procesos  $\{Y^N\}_{N \in \mathbb{N}}$  converge débilmente a un proceso de difusión, al que llamaremos  $Z$  y es generalmente conocido como el proceso de difusión de Wright-Fisher con dos tipos de alelos, sin selección y sin mutación.

En la primera sección se demuestra que la sucesión de procesos  $\{Y^N\}_{N \in \mathbb{N}}$ , que conforman una familia de procesos a la que llamaremos  $\mathcal{Y}$ , cumple el Criterio de Tensión de Aldous, y por tanto es relativamente compacta.

El hecho de que la sucesión sea relativamente compacta, quiere decir que toda subsucesión contiene una subsucesión que converge débilmente a cierto proceso, y si los límites de todas las subsucesiones son el mismo, entonces la sucesión completa converge débilmente.

Con el propósito de demostrar que las subsucesiones convergen al mismo límite, en la segunda sección calcularemos el generador de cualquiera de ellos en una amplia gama de funciones y para todo tiempo  $t > 0$  excepto un conjunto de medida 0. Como podremos observar, el generador resulta ser el mismo para todos los límites de las subsucesiones; obtendremos el generador  $\mathcal{L}(f)(x) = \frac{1}{2}f''(x) \cdot x(1-x)$ . Como conclusión de esta sección, tenemos que todos los límites de las subsucesiones satisfacen el Problema de Martingala para  $\mathcal{L}$  sobre las funciones 4 veces diferenciables.

En la tercera sección demostraremos que el Problema de Martingala para el generador  $\mathcal{L}$  (calculado en la segunda sección), sobre las funciones cuatro veces diferenciables, tiene una única solución, es decir, cualesquiera dos procesos que satisfagan el Problema de Martingala para  $\mathcal{L}$  sobre el conjunto de funciones  $C^4([0, 1])$  tienen la misma distribución. Esto demuestra que todas las subsucesiones convergen al mismo proceso.

De esta forma habremos demostrado que la sucesión entera  $\{Y^N\}_{N \in \mathbb{N}}$  converge débilmente al proceso cuyo generador es  $\mathcal{L}(f)(x) = \frac{1}{2}f''(x) \cdot x(1-x)$ , el cual es el generador de un proceso de difusión.

En la última sección utilizaremos el generador  $\mathcal{L}$  para calcular la esperanza del tiempo de absorción de este proceso de difusión que aproxima al modelo de Wright-Fisher haploide con dos tipos de alelos, y justificamos por qué es esta esperanza una buena aproximación a la esperanza de tiempo de absorción de la cadena de Wright-Fisher haploide con dos tipos de alelos.

A lo largo de este capítulo utilizaremos constantemente teoremas del libro [?] “Markov Processes, Characterization and Convergence” de los autores Stewart N. Ethier y Thomas Kurtz. (edición de 1986).

La estructura de éste capítulo fue tomada de la subsección 5.3.1 Neutral 2-allele Wright-Fisher Model de la sección 5.3 Diffusion approximation of Wright-Fisher del capítulo 5 Wright-Fisher processes de las notas “Preliminary Version Notes for Lectures on Stochastic Population Systems” del autor Donald Dawson. (edición July 7, 2009). Donde se definen las versiones continuas por la derecha con límites por la izquierda de las cadenas de Markov, y se enlistan los pasos que se tendrían que seguir para demostrar la convergencia al proceso de difusión de Wright-Fisher.

En la primera sección, donde demostramos que la sucesión  $\{Y^N\}_{N \in \mathbb{N}}$  es tensa y por lo tanto relativamente compacta, utilizamos ampliamente el libro [B] “Convergence of Probability Measures” del autor Patrick Billingsley. (edición de 1999).

### 3.1. La sucesión $\mathcal{Y} = \{Y^N\}_{N \in \mathbb{N}}$ es Tensa

Debido a la forma en que fue construido el proceso  $Y_t^N$ , sus trayectorias son continuas por la derecha con límites por la izquierda.

Existe un espacio métrico,  $D_\infty$ , separable, compacto y completo. Este espacio consiste en el conjunto de las funciones de  $[0, \infty)$  en  $\mathbb{R}$  continuas por la derecha con límites por la izquierda, con la métrica de Skorohod.

Las trayectorias de los procesos  $Y_t^N$  pertenecen a este espacio. Los procesos  $Y_t^N$  son cada uno, un elemento tomado aleatoriamente sobre este espacio  $D_\infty$ . Por lo tanto, podemos usar todos los resultados existentes sobre él.

Lo que demostraremos en esta sección, es que la sucesión de procesos  $\{Y_t^N\}_{N \in \mathbb{N}}$  converge débilmente a algún proceso al que llamaremos  $W_t$ .

Tomaremos la definición de Convergencia Débil del libro de Billingsley [B] en la página 7.

**Definición 3.1.1.** *Sea  $S$  un espacio métrico.*

*Sean  $P, \{P_n\}_n$  medidas de probabilidad sobre  $S$ .*

*Si para toda función real  $f$  sobre  $S$  ( $f : S \rightarrow \mathbb{R}$ ), continua y acotada tenemos la convergencia en los reales de:*

$$\int f \delta P_n \rightarrow \int f \delta P \quad (3.1)$$

*entonces decimos que la sucesión  $\{P_n\}_n$  converge débilmente a  $P$ .*

La convergencia débil en espacios como  $D_\infty$  o  $C_\infty$  (el espacio de funciones continuas de  $\mathbb{R}$  en  $\mathbb{R}$ ) es mucho más complicada que la convergencia débil en espacios más chicos, como  $\mathbb{R}^\infty$ , el espacio de sucesiones de números reales, de donde se escogen aleatoriamente las cadenas de Markov.

En este último espacio ( $\mathbb{R}^\infty$ ), basta probar la convergencia de las probabilidades  $P_n$  únicamente sobre los “conjuntos de dimensión finita” ( $\mathbb{R}_f^\infty$ ), conjuntos de la forma  $\pi_k^{-1}(H)$ ,  $H \in \mathbb{R}^k$ , los cuales tienen en  $k$  coordenadas un subconjunto propio de  $\mathbb{R}$ , y el espacio entero  $\mathbb{R}$  en las demás coordenadas. En el ámbito de las cadenas de Markov, la convergencia de las probabilidades sobre  $\mathbb{R}_f^\infty$ , equivale a la convergencia de las funciones de distribución conjuntas sobre  $k$  tiempos.

En el espacio de funciones continuas,  $C_\infty$ , la convergencia no es tan fácil de probar. Se puede construir un conjunto denso numerable de funciones continuas, es decir, un conjunto que separa al espacio  $C_\infty$ , sin embargo, la convergencia en este conjunto de funciones no implica la convergencia débil global, a diferencia del espacio  $\mathbb{R}^\infty$ .

En el espacio  $\mathbb{R}^\infty$  la clase  $\mathbb{R}_f^\infty$ , que es la clase de los conjuntos que tienen sólo un número finito de coordenadas distintas de  $\mathbb{R}$  completo (relacionado con las distribuciones finito-dimensionales), separa a  $\mathbb{R}^\infty$ , y es un conjunto que determina la convergencia de los elementos de  $\mathbb{R}^\infty$ . (Libro del autor Billingsley [B], página 19).

Al ser  $C_\infty \subseteq D_\infty$  la convergencia en  $D_\infty$  es igual o más difícil de probar. Pero ya hay un camino establecido.

**Teorema 3.1.1** ( Teorema 2.6, Billingsley [B], página 20 ). *Para cualquier espacio métrico  $S$ , una condición necesaria y suficiente para que una sucesión de probabilidades (sobre  $S$ )  $\{P_n\}_{n \in \mathbb{N}}$  converja a  $P$  es que cada subsucesión  $\{P_{n_i}\}_{i \in \mathbb{N}}$  contenga otra subsucesión  $\{P_{n_{i_k}}\}_{k \in \mathbb{N}}$  que converja débilmente a  $P$ .*

Esta última caracterización es muy parecida a otra característica a la que se le llama “**Compacidad Relativa**”. La enorme diferencia es que la compacidad relativa no asegura que la  $P$  a la que convergen las subsucesiones sea la misma.

**Definición 3.1.2** ( Billingsley [B], página 57 ). *Una familia  $\Pi$  de probabilidades en  $S$  es **relativamente compacta** si para cualquier sucesión de probabilidades de*

$\Pi$ , es decir, para cualquier  $\{P_n\}_{n \in \mathbb{N}}$ , existe una subsucesión  $\{P_{n_i}\}_i$  y una medida de probabilidad  $Q$  (dependiente completamente de la subsucesión  $\{P_{n_i}\}_i$ ) tal que  $P_{n_i} \Rightarrow_i Q$ .

Como la compacidad relativa de una sucesión muchas veces no es fácil de probar, se recurre a otro concepto muy relacionado, la tensión. Se recurre a él porque si una familia es tensa, entonces también es relativamente compacta.

**Definición 3.1.3.** Si una familia de probabilidades  $\Pi = \{P_\xi\}_{\xi \in \Xi}$  es tal que para todo  $\epsilon > 0$  existe un compacto  $K$  tal que :

$$\mathbb{P}_\xi(K) > 1 - \epsilon \quad \forall \xi \in \Xi \quad (3.2)$$

decimos que la familia es **tensa**.

El Teorema 5.1 en la página 59 del libro de Billingsley [B], prueba que:

**Teorema 3.1.2** ( Teorema 5.1, Billingsley [B], página 59 ). Si una familia es **tensa** entonces también es **relativamente compacta**.

Cada elemento aleatorio está directamente relacionado con la medida de probabilidad que induce sobre el espacio sobre el cual toma valores. A partir del proceso se induce una única medida, y a partir de la medida inducida se puede recuperar el proceso. Se pueden recuperar muchos más procesos a partir de la medida, pero todos serán equivalentes al original, tendrán la misma “Ley”, la misma función de distribución. La definición de que una familia de procesos sea relativamente compacta, sea tensa, o converge débilmente, consiste en que la familia de medidas asociadas sea relativamente compacta, tensa, o converge débilmente. Y por lo tanto para cada teorema que hemos dado, enunciado en términos de probabilidades, acerca de compacidad relativa, tensión o convergencia débil, se cumple el teorema análogo enunciado en términos de los procesos.

La familia de procesos que nos interesa es la familia

$$\mathcal{Y} = \{Y^N\}_{N \in \mathbb{N}} = \left\{ \frac{X_{\lfloor Nt \rfloor}^N}{N} \right\}_{N \in \mathbb{N}}$$

A continuación vamos a demostrar que la familia  $\mathcal{Y}$  es tensa. Con esto habremos demostrado que también es una familia relativamente compacta, y lo único que nos faltará para probar la convergencia débil utilizando el **teorema 3.1.1** es que las probabilidades  $Q$  a las que convergen las subsucesiones son la misma.

Para probar la tensión, tenemos un criterio muy fuerte:

**Teorema 3.1.3** ( **Criterio de Tensión de Aldous**, Teorema 16.10, Billingsley [B], página 178 ). Si para toda  $n$ ,  $X^n$  es un elemento tomado aleatoriamente del espacio de funciones de  $\mathbb{R}$  en  $\mathbb{R}$  continuas por la derecha con límites por la izquierda

64CAPÍTULO 3. CONVERGENCIA DÉBIL A UN PROCESO DE DIFUSIÓN

, es decir, si  $X^n \in D_\infty$  y se cumplen las siguientes dos condiciones, entonces  $\{X^n\}_n$  es una sucesión tensa.

**Condición 1** Para toda  $m \in \mathbb{R}^+$

$$\lim_{a \rightarrow \infty} \limsup_n \mathbb{P}(\|X^n\|_m \geq a) = 0 \quad (3.3)$$

**Condición 2** Para cada  $\epsilon, \eta, m$  existe una  $\delta_0$  y  $n_0$  tales que, si  $\delta \leq \delta_0$  y  $n_0 \leq n$ , y si  $\tau$  es cualquier tiempo de paro de  $X^n$  con  $\tau \leq m$ , se cumple la desigualdad :

$$\mathbb{P}(|X_{\tau+\delta}^n - X_\tau^n| \geq \epsilon) \leq \eta \quad (3.4)$$

Donde  $\|X^n\|_m = \sup_{0 \leq s \leq m} |X_s^n|$

**Lema 3.1.1.** La sucesión de procesos  $\{Y^N\}_N$  satisface las condiciones del criterio de Aldous, y por lo tanto, es tensa.

*Demostración.* Los procesos  $Y^N \in D_{[0,1]}[0, \infty) \subset D_\infty$ , es decir, son funciones de  $\mathbb{R}$  en  $\mathbb{R}$  continuas por la derecha con límites por la izquierda (càdlag).

**Condición 1.**

Como  $0 \leq Y_t^N \leq 1$  para toda  $t$  por construcción, entonces para toda  $m$ :

$$\|Y^N\|_m = \sup_{0 \leq s \leq m} |Y_s^N| \leq 1$$

y por lo tanto tenemos que, para cualquier  $N$  fija:

$$\mathbb{P}(\|Y^N\|_m \geq 2) = 0$$

y entonces :

$$\lim_{a \rightarrow \infty} \limsup_N \mathbb{P}(\|Y^N\|_m \geq a) = 0$$

**Condición 2.**

Primero recordemos la desigualdad de Chebyshev:

$$\mathbb{P}(|X - \mathbb{E}(X)| \geq \epsilon) \leq \frac{\text{Var}(X)}{\epsilon^2} \quad (3.5)$$

Recordemos también que  $Y_t^N = \frac{1}{N} X_{[Nt]}^N$ , que la cadena  $\{\frac{X_n^N}{N}\}_{n \in \mathbb{N}}$  es fuertemente de Markov, es Martingala y sus probabilidades de transición son  $\frac{1}{N} \text{Binomial}\left(N, \frac{X_{n-1}^N}{N}\right)$ .

Queremos demostrar que se cumple la Condición 2 del criterio de Aldous, es decir, queremos demostrar que para cada  $\epsilon, \eta, m$ , existen  $\delta_0$  y  $n_0$  tales que para

toda pareja  $\delta, N$  que cumpla que si  $\delta \leq \delta_0$  y  $N \geq n_0$  entonces se cumple la siguiente desigualdad para cualquier  $\tau$  tiempo de paro de  $Y_t^N$ .

$$\mathbb{P}\left(\left|Y_{\tau+\delta}^N - Y_{\tau}^N\right| \geq \epsilon\right) \leq \eta.$$

Como hablaremos mucho de  $Y_{\tau+\delta}^N = \frac{X_{\lfloor N(\tau+\delta) \rfloor}^N}{N}$  y de  $Y_{\tau}^N = \frac{X_{\lfloor N\tau \rfloor}^N}{N}$ , para simplificar las expresiones tomemos  $a = \lfloor N(\tau + \delta) \rfloor$  y tomemos  $k = \lfloor N(\tau + \delta) \rfloor - \lfloor N\tau \rfloor$ .

Notemos que  $k$  no es una variable aleatoria que dependa del tiempo de paro  $\tau$ ,  $k$  solo depende del número real  $\delta$ .

Demostramos rápidamente que  $\frac{X_n^N}{N}$  es Martingala.

$$\mathbb{E}\left(\frac{X_{n+1}^N}{N} \mid \frac{X_n^N}{N}\right) = \frac{1}{N} \mathbb{E}\left(X_{n+1}^N \mid \frac{X_n^N}{N}\right) = \frac{1}{N} \cdot N \cdot \frac{X_n^N}{N} = \frac{X_n^N}{N}$$

Empezaremos por calcular:

$$\begin{aligned} \mathbb{E}\left[Y_{\tau+\delta}^N \mid Y_{\tau}^N = \frac{i}{N}\right] &= \mathbb{E}\left[\frac{X_{a+k}^N}{N} \mid X_a^N = i\right] \\ &= \mathbb{E}\left[\frac{X_k^N}{N} \mid X_0^N = i\right] \\ &= \mathbb{E}\left[\frac{X_0^N}{N} \mid X_0^N = i\right] \\ &= \frac{i}{N} \end{aligned}$$

$$\begin{aligned}
 \mathbb{E} \left[ \left( Y_{\tau+\delta}^N \right)^2 \mid Y_{\tau}^N = \frac{i}{N} \right] &= \mathbb{E} \left[ \left( \frac{X_{a+k}^N}{N} \right)^2 \mid X_a^N = i \right] \\
 &= \mathbb{E} \left[ \left( \frac{X_k^N}{N} \right)^2 \mid X_0^N = i \right] \\
 &= \frac{1}{N^2} \cdot \mathbb{E} \left[ (X_k^N)^2 \mid X_0^N = i \right] \\
 &= \frac{1}{N^2} \cdot \mathbb{E} \left[ \mathbb{E} \left[ (X_k^N)^2 \mid X_0^N = i, X_{k-1}^N \right] \mid X_0^N = i \right] \\
 &= \frac{1}{N^2} \cdot \mathbb{E} \left[ N \cdot \frac{X_{k-1}^N}{N} \cdot \left( 1 - \frac{X_{k-1}^N}{N} \right) + \left( N \cdot \frac{X_{k-1}^N}{N} \right)^2 \mid X_0^N = i \right] \\
 &= \frac{1}{N^2} \cdot \mathbb{E} \left[ X_{k-1}^N + (X_{k-1}^N)^2 \cdot \left( 1 - \frac{1}{N} \right) \mid X_0^N = i \right] \\
 &= \frac{1}{N^2} \cdot \mathbb{E} \left[ X_{k-1}^N \mid X_0^N = i \right] + \left( 1 - \frac{1}{N} \right) \cdot \mathbb{E} \left[ \left( \frac{X_{k-1}^N}{N} \right)^2 \mid X_0^N = i \right] \\
 &= \frac{1}{N^2} \cdot i + \left( 1 - \frac{1}{N} \right) \cdot \mathbb{E} \left[ \left( \frac{X_{k-1}^N}{N} \right)^2 \mid X_0^N = i \right]
 \end{aligned}$$

Como podemos ver, el último término es igual al que queremos encontrar, pero con  $k-1$  en lugar de  $k$ , por lo que podemos proceder de manera recursiva. Obtenemos de esta forma

$$= \frac{1}{N} \cdot \frac{i}{N} \cdot \left( \sum_{i=0}^{k-1} \left( 1 - \frac{1}{N} \right)^i \right) + \left( \frac{i}{N} \right)^2 \cdot \left( 1 - \frac{1}{N} \right)^k$$

Y entonces tenemos que

$$\begin{aligned}
 \text{Var} \left[ \frac{X_{a+k}^N}{N} \mid X_a = i \right] &= \frac{1}{N} \cdot \frac{i}{N} \cdot \left( \sum_{i=0}^{k-1} \left( 1 - \frac{1}{N} \right)^i \right) + \left( \frac{i}{N} \right)^2 \cdot \left[ \left( 1 - \frac{1}{N} \right)^k - 1 \right] \\
 &\leq \frac{k}{N} \cdot \frac{i}{N} \leq \frac{k}{N}
 \end{aligned} \tag{3.6}$$

La desigualdad se sigue porque  $0 \leq i \leq N$  y  $\left( 1 - \frac{1}{N} \right)^k - 1 \leq 0$ .

Usando la desigualdad de Chebyshev ( 3.5 ) con la probabilidad condicional  $\mathbb{P}(\bullet \mid X_a^N = i) = \mathbb{P}_{\{X_a^N = i\}}(\bullet)$  obtenemos la siguiente desigualdad

$$\mathbb{P}_{\{X_a^N = i\}} \left( \left| \frac{X_{a+k}^N}{N} - \frac{X_a^N}{N} \right| \geq \epsilon \right) \leq \frac{\text{Var}_{\{X_a^N = i\}} \left( \frac{X_{a+k}^N}{N} \right)}{\epsilon^2} \leq \frac{k}{N \cdot \epsilon^2}$$

Llegar a una desigualdad como esta, pero con la probabilidad original (no la condicional) es muy sencillo:

$$\begin{aligned} \mathbb{P} \left( \left| \frac{X_{a+k}^N}{N} - \frac{X_a^N}{N} \right| \geq \epsilon \right) &= \sum_{i=0}^N \mathbb{P}_{\{X_a^N = i\}} \left( \left| \frac{X_{a+k}^N}{N} - \frac{X_a^N}{N} \right| \geq \epsilon \right) \cdot \mathbb{P}(X_a^N = i) \\ &\leq \sum_{i=0}^N \left( \frac{k}{N\epsilon^2} \right) \cdot \mathbb{P}(X_a^N = i) = \left( \frac{k}{N\epsilon^2} \right) \cdot 1 \end{aligned}$$

Regresemos a las variables  $Y_{\tau+\delta}^N$ . Habíamos tomado  $a = \lfloor N\tau \rfloor$ , y también habíamos tomamos  $k = \lfloor N(\tau + \delta) \rfloor - \lfloor N\tau \rfloor$ .

Lo que la ecuación anterior nos dice en términos de  $Y_{\tau+\delta}^N$  y  $Y_{\tau}^N$  es:

$$\mathbb{P} \left( \left| Y_{\tau+\delta}^N - Y_{\tau}^N \right| \geq \epsilon \right) = \mathbb{P} \left( \left| \frac{X_{a+k}^N}{N} - \frac{X_a^N}{N} \right| \geq \epsilon \right) \leq \frac{k}{N \cdot \epsilon^2}$$

Y además tenemos la siguiente desigualdad:

$$k = \lfloor N(\tau + \delta) \rfloor - \lfloor N\tau \rfloor \leq N\tau + N\delta - \lfloor N\tau \rfloor \leq N\delta + 1$$

y entonces:

$$\mathbb{P} \left( \left| Y_{\tau+\delta}^N - Y_{\tau}^N \right| \geq \epsilon \right) \leq \frac{N\delta + 1}{N\epsilon^2} = \frac{\delta}{\epsilon^2} + \frac{1}{N\epsilon^2}$$

Tomando  $\delta_0 \leq \eta \cdot \epsilon^2 \cdot \frac{1}{2}$  y tomando  $n_0 \geq \frac{1}{\eta \cdot \epsilon^2}$  se cumple la Condición 2:

Para cada  $\epsilon, \eta, m$ , tomemos  $\delta_0 = \frac{1}{2} \cdot \eta \cdot \epsilon^2$  y tomemos  $n_0 = \frac{1}{2} \cdot \frac{1}{\eta \cdot \epsilon^2}$ . Con estas  $\delta_0$  y  $n_0$  tenemos que si  $\delta \leq \delta_0$  y  $n_0 \leq N$ , para cualquier tiempo de paro  $\tau$  se cumple la desigualdad :

$$\mathbb{P} \left( \left| Y_{\tau+\delta}^N - Y_{\tau}^N \right| \geq \epsilon \right) \leq \frac{\delta}{\epsilon^2} + \frac{1}{N\epsilon^2} \leq \eta$$

□

**En resumen** tenemos que la familia de procesos  $\{Y^N\}_N$  es tensa (gracias al criterio de Aldous) y por lo tanto es relativamente compacta. Es decir, toda subsucesión  $\{Y^{N_w}\}_w$  tiene una subsucesión  $\{Y^{N_{a_k}}\}_k$  que converge débilmente, con un abuso de notación diremos que converge a  $Z^{w_k}$ .

Una vez que hemos demostrado que la familia es tensa, para demostrar que la sucesión completa converge débilmente, solo necesitamos probar que todos los límites de las subsucesiones de las subsucesiones,  $Z^{w_k}$ , son el mismo.

Para demostrar esto, utilizaremos el generador de los procesos límite de las subsucesiones de las subsucesiones,  $Z^{w_k}$ .

El único detalle, es que para que un proceso tenga un generador necesita ser un proceso de Markov, y no hemos demostrado que los límites de nuestras subsucesiones lo son. Sin embargo, ésto está demostrado en el libro [?] de Ethier y Kurtz.

Suponiendo que los procesos límite son procesos de Markov, (se puede encontrar la prueba en [?]) calcularemos este generador para casi todos los tiempos  $t > 0$ , y con esto podremos demostrar que cada uno de los límites  $Z^{w_k}$  satisface el Problema de Martingala para ese generador extendido de a todo  $t > 0$  sobre el conjunto de funciones  $C^4([0, 1])$ . Después demostraremos que el Problema de Martingala para ese generador sobre  $C^4([0, 1])$  tiene una única solución, y por lo tanto, todos los límites  $Z^{w_k}$  tienen la misma distribución, y entonces la sucesión entera converge débilmente a ese mismo proceso, la única solución del generador.

## 3.2. El generador de la subsucesiones convergentes

En esta sección, suponiendo que los procesos límite son procesos de Markov (consultar [?]), calcularemos para casi toda  $t \geq 0$  el generador del proceso límite de las subsucesiones que convergen débilmente, extenderemos el generador a todos los tiempos  $t \geq 0$  de forma continua, y demostraremos que los procesos límite satisfacen el Problema de Martingala de este generador sobre el conjunto de funciones  $C^4([0, 1])$ .

**Definición 3.2.1.** *El generador de un proceso estocástico  $X$  es un operador  $\mathcal{L}$  cuyo dominio es el conjunto de funciones continuas de  $\mathbb{R}$  en  $\mathbb{R}$  para las que el siguiente límite existe para todos los posibles valores del proceso  $X_t$ , y cuya regla de correspondencia consiste en asignar a cada una de estas funciones  $f$ , la correspondiente función límite  $\mathcal{L}(f)$ .*

$$\begin{aligned} \mathcal{L} : \text{Dominio}(\mathcal{L}) \subseteq C_{\mathbb{R}}(\mathbb{R}) &\longrightarrow C_{\mathbb{R}}(\mathbb{R}) \\ \mathcal{L}(f) &= g \end{aligned}$$

### 3.2. EL GENERADOR DE LA SUBSUCESIONES CONVERGENTES 69

$$\begin{aligned} \mathcal{L}(f) : \text{Rango}(X) &\longrightarrow \mathbb{R} \\ \mathcal{L}(f)(X_t) &= \lim_{h \rightarrow 0} \frac{\mathbb{E} \left[ f(X_{t+h}) - f(X_t) \mid \mathcal{F}_t^X \right]}{h} = g(X_t) \end{aligned}$$

En esta sección calcularemos para casi todo tiempo  $t$  cada subsucesión de subsucesión  $\{Y^{N_{w_k}}\}_{k \in \mathbb{N}}$  que converge débilmente a  $Z^{w_k} = \lim_{k \rightarrow \infty} Y^{N_{w_k}}$ , el generador del proceso límite  $Z^{w_k}$ . Es decir, calcularemos para casi todo  $t$ :

$$\mathcal{L}(f)(Z_t^{w_k}) = \lim_{h \rightarrow 0} \frac{1}{h} \cdot \mathbb{E} \left[ f(Z_{t+h}^{w_k}) - f(Z_t^{w_k}) \mid \mathcal{F}_t^{Z^{w_k}} \right] = g(Z_t^{w_k})$$

Y obtendremos como resultado:

**Teorema 3.2.1.** *Dada una subsucesión  $\{Y^{N_{w_k}}\}_{k \in \mathbb{N}}$  que converge débilmente a  $Z^{w_k}$ , y dada una función 3 veces diferenciable y con tercera derivada acotada, el generador  $\mathcal{L}$  del proceso  $Z^{w_k}$  evaluado en la función  $f$  es:*

$$\begin{aligned} \mathcal{L}(f)(Z_t^{w_k}) &= \frac{1}{2} \cdot f''(Z_t^{w_k}) \cdot \left[ Z_t^{w_k} (1 - Z_t^{w_k}) \right] \\ &= g(Z_t^{w_k}) = \frac{1}{2} \cdot f''(\bullet) \cdot \left[ \bullet \cdot (1 - \bullet) \right] (Z_t^{w_k}) \end{aligned}$$

#### 3.2.1. El generador y el Teorema de Taylor

Para funciones 3 veces diferenciables y con tercera derivada acotada se puede calcular este límite utilizando el Teorema de Taylor de grado 2, y eso es lo que haremos a continuación. Más adelante justificaremos por qué solo necesitamos encontrar el límite para estas funciones.

**Teorema 3.2.2** (Teorema de Taylor de grado 2). *Para cualquier función real  $f$  cuyas primeras 3 derivadas están definidas sobre  $[0, 1]$ , y para todo par de reales  $a$  y  $x$ , con  $a, x \in [0, 1]$  sucede que existe algún  $t^* \in (a, x)$  tal que*

$$f(x) - f(a) = f'(a) \cdot (x - a) + f''(a) \cdot \frac{1}{2} \cdot (x - a)^2 + f'''(t^*) \cdot \frac{1}{3!} \cdot (x - a)^3$$

Como  $0 \leq Z^{w_k} \leq 1$ , podemos reescribir al generador como:

$$\begin{aligned} \mathcal{L}(f)(Z_t^{w_k}) &= \lim_{h \rightarrow 0} \frac{1}{h} \cdot \mathbb{E} \left[ f'(Z_t^{w_k}) \cdot (Z_{t+h}^{w_k} - Z_t^{w_k}) + \frac{1}{2!} f''(Z_t^{w_k}) \cdot (Z_{t+h}^{w_k} - Z_t^{w_k})^2 \right. \\ &\quad \left. + \frac{1}{3!} \cdot f'''(T^*) \cdot (Z_{t+h}^{w_k} - Z_t^{w_k})^3 \mid \mathcal{F}_t^{Z^{w_k}} \right]. \end{aligned}$$

70CAPÍTULO 3. CONVERGENCIA DÉBIL A UN PROCESO DE DIFUSIÓN

para algún  $\mathbf{T}^*$  que se encuentra entre los reales  $Z_t^{w_k}$  y  $Z_{t+h}^{w_k}$ .

Como estamos condicionando con respecto a  $\mathcal{F}_t^{Z^{w_k}}$ , los términos que solo dependan de esta variable pueden salir de la esperanza.

$$\begin{aligned} \mathcal{L}(f)(Z_t^{w_k}) &= \lim_{h \rightarrow 0} \frac{1}{h} \cdot f'(Z_t^{w_k}) \cdot \mathbb{E} \left[ Z_{t+h}^{w_k} - Z_t^{w_k} \mid \mathcal{F}_t^{Z^{w_k}} \right] \\ &+ \frac{1}{h} \cdot \frac{1}{2!} f''(Z_t^{w_k}) \cdot \mathbb{E} \left[ (Z_{t+h}^{w_k} - Z_t^{w_k})^2 \mid \mathcal{F}_t^{Z^{w_k}} \right] \\ &+ \frac{1}{h} \cdot \frac{1}{3!} \cdot \mathbb{E} \left[ f'''(\mathbf{T}^*) \cdot (Z_{t+h}^{w_k} - Z_t^{w_k})^3 \mid \mathcal{F}_t^{Z^{w_k}} \right]. \end{aligned} \quad (3.7)$$

Lo que haremos es calcular los límites de las primeras dos esperanzas condicionales para casi todos los tiempos  $t$ ,

$$\begin{aligned} \lim_{h \rightarrow 0} \frac{1}{h} \cdot \frac{1}{3!} \cdot \mathbb{E} \left[ Z_{t+h}^{w_k} - Z_t^{w_k} \mid \mathcal{F}_t^{Z^{w_k}} \right] &\quad \text{en la subsección 3.2.2} \\ \lim_{h \rightarrow 0} \frac{1}{h} \cdot \frac{1}{3!} \cdot \mathbb{E} \left[ (Z_{t+h}^{w_k} - Z_t^{w_k})^2 \mid \mathcal{F}_t^{Z^{w_k}} \right] &\quad \text{en la subsección 3.2.2} \end{aligned}$$

y después demostrar en la subsección 3.2.2 que

$$\lim_{h \rightarrow 0} \frac{1}{h} \cdot \frac{1}{3!} \cdot \mathbb{E} \left[ (Z_{t+h}^{w_k} - Z_t^{w_k})^3 \mid \mathcal{F}_t^{Z^{w_k}} \right] = 0$$

(para casi todos los tiempos  $t$ ) con lo cual, para cualquier función  $f$  tres veces diferenciable y con tercera derivada acotada tenemos que

$$\lim_{h \rightarrow 0} \frac{1}{h} \cdot \frac{1}{3!} \cdot \mathbb{E} \left[ f'''(\mathbf{T}^*) (Z_{t+h}^{w_k} - Z_t^{w_k})^3 \mid \mathcal{F}_t^{Z^{w_k}} \right] = 0$$

(para casi todos los tiempos  $t$ ) y entonces, para estas funciones, su imagen bajo el operador  $\mathcal{L}$  consiste simplemente en sustituir los valores de los límites de las primeras dos esperanzas condicionales y de las primeras dos derivadas de la función (para casi todos los tiempos  $t$ ).

Obtener el generador  $\mathcal{L}$  sobre todos los tiempos  $t$  excepto un conjunto de medida 0 será suficiente para demostrar que los procesos  $Z^{w_k}$  satisfacen las condiciones del Problema de Martingala del generador de un proceso de difusión que tiene solución única. Con estos demostraremos que todos los procesos  $Z^{w_k}$  son el mismo, ya que si dos procesos son soluciones de un Problema de Martingala que tiene solución única, entonces los dos procesos tienen la misma distribución.

### 3.2.2. El Problema de Martingala y las esperanzas condicionales

#### Teorema 3.2.3.

*El límite débil de una sucesión de Martingalas es una Martingala.*

Del **teorema 5.1** del libro de Ethier y Kurtz [?] en la página 74, junto con la observación **5.2 Remark (b)** y el párrafo anterior al teorema, podemos concluir que se cumple el siguiente teorema:

**Teorema 3.2.4** (Teorema 5.1 y Remark 5.2, Ethier y Kurtz [?], página 74 ).

*Sea  $X$  una submartingala càdlag (continua por la derecha) y acotada por abajo.*

*Si existe un proceso  $A$ , adaptado, **continuo**, con trayectorias no decrecientes, con  $A_0 = 0$  y tal que  $M_t = X_t - A_t$  es una Martingala, tenemos que este proceso es el único que cumple con estas características.*

**Teorema 3.2.5** ([?]). *Si  $M_t$  es una Martingala y  $f(x)$  es una función convexa, entonces  $f(M_t)$  es una submartingala.*

Notemos que  $f(x) = x^2$  y  $f(x) = |x|^3$  son funciones convexas.

**Teorema 3.2.6** ( **Teorema de Martingala**, Proposición 1.7, Ethier y Kurtz [?], página 162 ).

*Sea  $X$  un proceso estocástico con generador  $\mathcal{L}$ . Para toda función  $f \in \text{Dominio}(\mathcal{L})$ , si  $g(x)$  es la imagen de  $f(x)$  bajo  $\mathcal{L}$ , entonces el siguiente proceso es una Martingala:*

$$M_t = f(X_t) - \int_0^t g(X_s) ds$$

**Definición 3.2.2** ( Solución al Problema de Martingala, Ethier y Kurtz [?], página 173 ). *Decimos que un proceso estocástico  $X$ , que toma valores en el espacio métrico  $E$  y está definido sobre el espacio de probabilidad  $(\Omega, \mathcal{F}, \mathbb{P})$ , es una **solución al problema de Martingala**  $(\mathbf{B}, \mu)$ , ( $B$  siendo un generador y  $\mu$  una distribución inicial) si se cumple que*

$$f(X_t) - \int_0^t B(f)(X_s) ds$$

*es una Martingala para cualquier  $f$  en el dominio del generador  $B$ , y la distribución inicial del proceso está dada por  $\mu$*

$$\mathbb{P}(X_0^{-1}) = \mu.$$

**La primera esperanza condicional**

En esta subsección calcularemos la esperanza condicional:

$$\mathbb{E} \left[ Z_{t+h}^{w_k} - Z_t^{w_k} \mid \mathcal{F}_t^{Z^{w_k}} \right] = 0.$$

Para cualquier  $N$ , el proceso  $Y^N$  es una Martingala; entonces el límite de cualquier subsección de  $\{Y^N\}_{N \in \mathbb{N}}$  que converja débilmente también es una Martingala. Por lo tanto,  $Z_t^{w_k}$  es una Martingala, y entonces,

$$\lim_{h \rightarrow 0} \frac{1}{h} \cdot \mathbb{E} \left[ Z_{t+h}^{w_k} - Z_t^{w_k} \mid \mathcal{F}_t^{Z^{w_k}} \right] = \lim_{h \rightarrow 0} \frac{1}{h} \cdot \left[ Z_t^{w_k} - Z_t^{w_k} \right] = 0$$

**La segunda esperanza condicional**

En esta subsección calculamos el límite de la segunda esperanza condicional:

$$\lim_{h \rightarrow 0} \frac{1}{h} \cdot \mathbb{E} \left[ (Z_{t+h}^{w_k} - Z_t^{w_k})^2 \mid \mathcal{F}_t^{Z^{w_k}} \right] = Z_t^{w_k} \cdot (1 - Z_t^{w_k}).$$

**Lema 3.2.1.** *El generador  $\mathcal{L}$  del proceso  $Z^{w_k}$  evaluado en  $f(x) = x^2$ , quien por definición es  $\mathcal{L}(f)(Z^{w_k}) = \lim_{h \rightarrow 0} \frac{1}{h} \cdot \mathbb{E} \left[ (Z_{t+h}^{w_k})^2 - (Z_t^{w_k})^2 \mid \mathcal{F}_t^{Z^{w_k}} \right]$ , resulta ser igual a*

$$\mathcal{L}(f)(Z^{w_k}) = \lim_{h \rightarrow 0} \frac{1}{h} \cdot \mathbb{E} \left[ (Z_{t+h}^{w_k} - Z_t^{w_k})^2 \mid \mathcal{F}_t^{Z^{w_k}} \right].$$

*Demostración.*

En la subsección 3.2.1, en la página 70 demostramos que:

$$\begin{aligned} \mathcal{L}(f)(Z_t^{w_k}) &= \lim_{h \rightarrow 0} \frac{1}{h} \cdot f'(Z_t^{w_k}) \cdot \mathbb{E} \left[ Z_{t+h}^{w_k} - Z_t^{w_k} \mid \mathcal{F}_t^{Z^{w_k}} \right] \\ &\quad + \frac{1}{h} \cdot \frac{1}{2!} f''(Z_t^{w_k}) \cdot \mathbb{E} \left[ (Z_{t+h}^{w_k} - Z_t^{w_k})^2 \mid \mathcal{F}_t^{Z^{w_k}} \right] \\ &\quad + \frac{1}{h} \cdot \frac{1}{3!} f'''(Z_t^{w_k}) \cdot \mathbb{E} \left[ (Z_{t+h}^{w_k} - Z_t^{w_k})^3 \mid \mathcal{F}_t^{Z^{w_k}} \right]. \end{aligned}$$

Si evaluamos esta ecuación en la función  $f(x) = x^2$ , y utilizamos el hecho de que  $2 \cdot Z^{w_k} \cdot \mathbb{E} \left[ Z_{t+h}^{w_k} - Z_t^{w_k} \mid \mathcal{F}_t^{Z^{w_k}} \right] = 0$  y  $f'''(x) = 0$ , obtenemos lo que

### 3.2. EL GENERADOR DE LA SUBSUCESIONES CONVERGENTES 73

buscamos:

$$\mathcal{L}(f)(Z_t^{w_k}) = \lim_{h \rightarrow 0} \frac{1}{h} \cdot \mathbb{E} \left[ (Z_{t+h}^{w_k} - Z_t^{w_k})^2 \mid \mathcal{F}_t^{Z^{w_k}} \right]$$

□

**Teorema 3.2.7.** *Sea  $Z^{w_k}$  el límite débil de una subsucesión de  $\{Y^N\}_{N \in \mathbb{N}}$  que converge débilmente. El siguiente proceso es una Martingala:*

$$\left( Z_t^{w_k} \right)^2 + \int_0^t Z_s^{w_k} \cdot (1 - Z_s^{w_k}) ds.$$

*Demostración.* Llamemos  $v = \lfloor Nt \rfloor$  y llamemos  $k = \lfloor Nh \rfloor$ . Tenemos las siguientes igualdades:

$$\left( Y_t^N \right)^2 = \left( \frac{X_{\lfloor Nt \rfloor}^N}{N} \right)^2 = \left( \frac{X_0^N}{N} \right)^2 + \sum_{i=0}^{v-1} \left( \frac{X_{i+1}^N}{N} \right)^2 - \left( \frac{X_i^N}{N} \right)^2$$

$$\left( Y_t^N \right)^2 - \left( Y_h^N \right)^2 = \sum_{i=k}^{v-1} \left( \frac{X_{i+1}^N}{N} \right)^2 - \left( \frac{X_i^N}{N} \right)^2$$

$$\mathbb{E} \left[ \left( Y_t^N \right)^2 - \left( Y_h^N \right)^2 \mid Y_h^N \right] = \mathbb{E} \left[ \frac{1}{N^2} \cdot \sum_{i=k}^{v-1} \left( X_{i+1}^N \right)^2 - \left( X_i^N \right)^2 \mid X_k^N \right]$$

$$\mathbb{E}_{Y_h^N} \left[ \left( Y_t^N \right)^2 \right] - \left( Y_h^N \right)^2 = \mathbb{E}_{X_k^N} \left[ \frac{1}{N^2} \cdot \sum_{i=k}^{v-1} \mathbb{E}_{X_k^N, X_i^N} \left[ \left( X_{i+1}^N \right)^2 - \left( X_i^N \right)^2 \right] \right]$$

$$\mathbb{E}_{Y_h^N} \left[ \left( Y_t^N \right)^2 \right] - \mathbb{E}_{X_0^N, \dots, X_k^N} \left[ \frac{1}{N^2} \cdot \sum_{i=k}^{v-1} N \cdot \frac{X_i^N}{N} \left( 1 - \frac{X_i^N}{N} \right) \right] = \left( Y_h^N \right)^2$$

Si buscamos un tiempo  $s$  tal que  $i = \lfloor Ns \rfloor$ , entonces podemos tomar  $s = \frac{i}{N}$ , ya que  $\lfloor N \cdot \frac{i}{N} \rfloor = i$ . Con lo cual tenemos que  $Y_{\frac{i}{N}}^N = \frac{X_i^N}{N} = \frac{X_{\lfloor N \cdot \frac{i}{N} \rfloor}^N}{N}$ .

$$\begin{aligned} \mathbb{E}_{Y_h^N} \left[ \left( Y_t^N \right)^2 \right] - \mathbb{E}_{\{Y_0^N, \dots, Y_{\frac{k}{N}}^N\}} \left[ \sum_{i=k}^{v-1} \frac{1}{N} \cdot Y_{\frac{i}{N}}^N \left( 1 - Y_{\frac{i}{N}}^N \right) \right] &= \sum_{i=0}^{k-1} \frac{1}{N} \cdot Y_{\frac{i}{N}}^N \left( 1 - Y_{\frac{i}{N}}^N \right) \\ &= \left( Y_h^N \right)^2 - \sum_{i=0}^{k-1} \frac{1}{N} \cdot Y_{\frac{i}{N}}^N \left( 1 - Y_{\frac{i}{N}}^N \right) \end{aligned}$$

74CAPÍTULO 3. CONVERGENCIA DÉBIL A UN PROCESO DE DIFUSIÓN

Recordemos que  $Y_h^N = \frac{X_{\lfloor Nh \rfloor}^N}{N} = \frac{X_k^N}{N} = Y_{\frac{k}{N}}^N = Y_{\frac{\lfloor Nh \rfloor}{N}}^N$ .

$$\mathbb{E}_{\{Y_0^N, \dots, Y_h^N\}} \left[ (Y_t^N)^2 - \sum_{i=0}^{v-1} \frac{1}{N} \cdot Y_{\frac{i}{N}}^N \left(1 - Y_{\frac{i}{N}}^N\right) \right] = (Y_h^N)^2 - \sum_{i=0}^{k-1} \frac{1}{N} \cdot Y_{\frac{i}{N}}^N \left(1 - Y_{\frac{i}{N}}^N\right)$$

Como podemos ver, esta última ecuación demuestra que el siguiente proceso, al cual llamaremos  $M_t^N$ , es Martingala.

$$M_t^N = (Y_t^N)^2 - \sum_{i=0}^{\lfloor Nt \rfloor - 1} \frac{1}{N} \cdot Y_{\frac{i}{N}}^N \left(1 - Y_{\frac{i}{N}}^N\right)$$

Debido a que tomamos una subsucesión  $\{Y^{N_{w_k}}\}_{k \in \mathbb{N}}$  que converge débilmente a  $Z^{w_k}$ , y con ella construimos la sucesión  $\{M^{N_{w_k}}\}_{k \in \mathbb{N}}$ , tenemos que esta sucesión converge débilmente a la Martingala:

$$M_t = \left(Z_t^{N_{w_k}}\right)^2 - \int_0^t Z_s^{N_{w_k}} \left(1 - Z_s^{N_{w_k}}\right) ds$$

Es decir, para cada subsucesión que converge débilmente

$$\lim_{N_{w_k} \rightarrow \infty} Y^{N_{w_k}} = Z^{w_k}$$

tenemos que el siguiente proceso es Martingala:

$$\left(Z_t^{w_k}\right)^2 - \int_0^t Z_s^{w_k} \left(1 - Z_s^{w_k}\right) ds,$$

que era lo que queríamos demostrar.  $\square$

Ahora reunamos estos dos resultados para demostrar el límite de esta sección.

**Teorema 3.2.8.**

$$\lim_{h \rightarrow 0} \frac{1}{h} \cdot \mathbb{E} \left[ \left(Z_{t+h}^{w_k} - Z_t^{w_k}\right)^2 \mid \mathcal{F}_t^{Z^{w_k}} \right] = Z_t^{w_k} \cdot \left(1 - Z_t^{w_k}\right).$$

*Demostración.*

En el teorema anterior demostramos que el siguiente proceso es una Martingala,

$$\left(Z_t^{w_k}\right)^2 - \int_0^t Z_s^{w_k} \cdot \left(1 - Z_s^{w_k}\right) ds.$$

Gracias al teorema 3.2.6 (página 71), tenemos que el siguiente proceso también es una Martingala,

### 3.2. EL GENERADOR DE LA SUBSUCESIONES CONVERGENTES 75

$$\begin{aligned} & \left( Z_t^{w_k} \right)^2 - \int_0^t \mathcal{L}(f) \left( Z_s^{w_k} \right) ds \\ = & \left( Z_t^{w_k} \right)^2 - \int_0^t \lim_{h \rightarrow 0} \frac{1}{h} \cdot \mathbb{E} \left[ \left( Z_{t+h}^{w_k} - Z_t^{w_k} \right)^2 \mid \mathcal{F}_t^{Z^{w_k}} \right] ds. \end{aligned}$$

llamando  $f$  a la función  $f(x) = x^2$ , y  $\mathcal{L}$  al generador del proceso  $Z^{w_k}$ .

Aquí estamos suponiendo implícitamente que el límite  $\lim_{h \rightarrow 0} \frac{1}{h} \cdot \mathbb{E} \left[ \left( Z_{t+h}^{w_k} - Z_t^{w_k} \right)^2 \mid \mathcal{F}_t^{Z^{w_k}} \right]$

existe. Se puede calcular el límite directamente demostrando de esta forma que en efecto el límite existe, sin embargo la cantidad de cuentas es excesiva, rebasa 20 páginas de puras cuentas, y por lo tanto no las anexamos a esta tesis.

Ambas integrales cumplen con ser continuas, adaptadas e iguales a 0 cuando  $t = 0$ . Además, sus integrandos siempre son no negativos, y entonces tienen trayectorias no decrecientes. También sucede que  $(Z^{w_k})^2$  es submartingala, porque  $Z^{w_k}$  es martingala. Es decir, ambas martingalas cumplen con todas las hipótesis del teorema **3.2.4** ( página 71 ), y por lo tanto son iguales para todo  $t$ .

$$\int_0^t \mathcal{L}(f) \left( Z_s^{w_k} \right) ds = \int_0^t Z_s^{w_k} \cdot \left( 1 - Z_s^{w_k} \right) ds.$$

Entonces, para casi toda  $s$ ,

$$\mathcal{L}(f) \left( Z_s^{w_k} \right) = Z_s^{w_k} \cdot \left( 1 - Z_s^{w_k} \right).$$

Con el primer lema de esta subsección, esto significa que para casi toda  $s$ ,

$$\lim_{h \rightarrow 0} \frac{1}{h} \cdot \mathbb{E} \left[ \left( Z_{t+h}^{w_k} - Z_t^{w_k} \right)^2 \mid \mathcal{F}_t^{Z^{w_k}} \right] = Z_s^{w_k} \cdot \left( 1 - Z_s^{w_k} \right).$$

□

#### La tercera esperanza condicional

En esta subsubsección calcularemos el límite de la tercera esperanza condicional:

$$\lim_{h \rightarrow 0} \frac{1}{h} \cdot \mathbb{E} \left[ \left( Z_{t+h}^{w_k} - Z_t^{w_k} \right)^3 \mid \mathcal{F}_t^{Z^{w_k}} \right] = 0.$$

**Lema 3.2.2.** Tomando  $f$  de la forma  $f(x) = x^3$ , tenemos que:

$$\mathcal{L}(f)(Z_t^{w_k}) = 3 \cdot (Z_t^{w_k})^2 - 3 \cdot (Z_t^{w_k})^3 + \lim_{h \rightarrow 0} \frac{1}{h} \cdot \mathbb{E} \left[ (Z_{t+h}^{w_k} - Z_t^{w_k})^3 \mid \mathcal{F}_t^{Z^{w_k}} \right].$$

*Demostración.*

En la sección 3.2.1 en la página 70, teníamos la ecuación (3.7):

$$\begin{aligned} \mathcal{L}(f)(Z_t^{w_k}) &= \lim_{h \rightarrow 0} \frac{1}{h} \cdot f'(Z_t^{w_k}) \cdot \mathbb{E} \left[ Z_{t+h}^{w_k} - Z_t^{w_k} \mid \mathcal{F}_t^{Z^{w_k}} \right] \\ &\quad + \frac{1}{h} \cdot \frac{1}{2!} f''(Z_t^{w_k}) \cdot \mathbb{E} \left[ (Z_{t+h}^{w_k} - Z_t^{w_k})^2 \mid \mathcal{F}_t^{Z^{w_k}} \right] \\ &\quad + \frac{1}{h} \cdot \frac{1}{3!} \cdot \mathbb{E} \left[ f'''(T^*) \cdot (Z_{t+h}^{w_k} - Z_t^{w_k})^3 \mid \mathcal{F}_t^{Z^{w_k}} \right]. \end{aligned}$$

Si sustituimos lo que hemos obtenido en las dos subsecciones anteriores acerca de las primeras dos esperanzas condicionales, y evaluamos esta expresión en la función  $f(x) = x^3$  obtenemos, para casi todo tiempo  $t$ :

$$\begin{aligned} \mathcal{L}(f)(Z_t^{w_k}) &= \lim_{h \rightarrow 0} \frac{1}{2} \cdot 6 \cdot Z_t^{w_k} \cdot \left[ Z_t^{w_k} \cdot (1 - Z_t^{w_k}) \right] \\ &\quad + \frac{1}{h} \cdot \frac{1}{3!} \cdot \mathbb{E} \left[ 3! \cdot (Z_{t+h}^{w_k} - Z_t^{w_k})^3 \mid \mathcal{F}_t^{Z^{w_k}} \right], \end{aligned}$$

a partir de esta ecuación se obtiene directamente la que buscamos en este lema.  $\square$

**Teorema 3.2.9.** Sea  $Z^{w_k}$  el límite débil de una subsucesión de  $\{Y^N\}_{N \in \mathbb{N}}$  que converge débilmente. El siguiente proceso es una Martingala

$$\left( Z_t^{w_k} \right)^3 - \int_0^t 3 \left( Z_s^{w_k} \right)^2 - 3 \left( Z_s^{w_k} \right)^3 ds.$$

*Demostración.*

Recordemos que  $Y_t^N = Y_{\lfloor \frac{Nt}{N} \rfloor}^N = \frac{X_{\lfloor Nt \rfloor}^N}{N}$ .

Sea  $v = \lfloor Nt \rfloor$  y sea  $k = \lfloor Nh \rfloor$ .

$$\begin{aligned} \mathbb{E}_{Y_h^N} \left[ \left( Y_t^N \right)^3 - \left( Y_h^N \right)^3 \right] &= \mathbb{E}_{Y_h^N} \left[ \sum_{i=k}^{v-1} \left( Y_{\frac{i+1}{N}}^N \right)^3 - \left( Y_{\frac{i}{N}}^N \right)^3 \right] \\ &= \mathbb{E}_{Y_{\frac{k}{N}}^N} \left[ \sum_{i=k}^{v-1} \mathbb{E}_{Y_{\frac{i+1}{N}}^N, Y_{\frac{i}{N}}^N} \left[ \left( Y_{\frac{i+1}{N}}^N \right)^3 - \left( Y_{\frac{i}{N}}^N \right)^3 \right] \right] \end{aligned}$$

### 3.2. EL GENERADOR DE LA SUBSUCESIONES CONVERGENTES 77

Ahora, utilizando la ecuación (A.3) de la página 90 del Anexo, y recordando que  $Y_t^N = \frac{X_{\lfloor Nt \rfloor}^N}{N} = Y_{\frac{\lfloor Nt \rfloor}{N}}^N$ , tenemos que cada uno de los términos de la suma es:

$$\begin{aligned} & \mathbb{E}_{Y_{\frac{k}{N}}^N} \left[ \mathbb{E}_{Y_{\frac{k}{N}}^N, Y_{\frac{i}{N}}^N} \left[ \left( Y_{\frac{i+1}{N}}^N \right)^3 - \left( Y_{\frac{i}{N}}^N \right)^3 \right] \right] = \\ & \mathbb{E}_{Y_{\frac{k}{N}}^N} \left[ \frac{N^2 - 3N + 2}{N^2} \cdot \left( Y_{\frac{i}{N}}^N \right)^3 + 3 \cdot \frac{1}{N} \cdot \frac{N-1}{N} \cdot \left( Y_{\frac{i}{N}}^N \right)^2 + \frac{1}{N^2} \cdot Y_{\frac{i}{N}}^N - \left( Y_{\frac{i}{N}}^N \right)^3 \right] \\ & = \mathbb{E}_{Y_{\frac{k}{N}}^N} \left[ \frac{1}{N} \left[ 3 \left( Y_{\frac{i}{N}}^N \right)^2 - 3 \left( Y_{\frac{i}{N}}^N \right)^3 \right] + \frac{1}{N^2} \left[ 2 \left( Y_{\frac{i}{N}}^N \right)^3 - 3 \left( Y_{\frac{i}{N}}^N \right)^2 + \left( Y_{\frac{i}{N}}^N \right) \right] \right] \end{aligned}$$

y entonces, recordando que  $v = \lfloor Nt \rfloor$  y que  $k = \lfloor Nh \rfloor$

$$\begin{aligned} \mathbb{E}_{Y_n^N} \left[ \left( Y_t^N \right)^3 - \left( Y_h^N \right)^3 \right] & = \sum_{i=\lfloor Nh \rfloor}^{\lfloor Nt \rfloor - 1} \frac{1}{N} \cdot \mathbb{E}_{Y_{\frac{k}{N}}^N} \left[ 3 \left( Y_{\frac{i}{N}}^N \right)^2 - 3 \left( Y_{\frac{i}{N}}^N \right)^3 \right] \\ & + \frac{1}{N} \cdot \sum_{i=\lfloor Nh \rfloor}^{\lfloor Nt \rfloor - 1} \frac{1}{N} \cdot \mathbb{E}_{Y_{\frac{k}{N}}^N} \left[ 2 \left( Y_{\frac{i}{N}}^N \right)^3 - 3 \left( Y_{\frac{i}{N}}^N \right)^2 + \left( Y_{\frac{i}{N}}^N \right) \right] \end{aligned}$$

Esta ecuación nos dice que el siguiente proceso es una Martingala, a la que llamaremos en esta subsección  $M^N$ :

$$\begin{aligned} M_t^N & = \left( Y_t^N \right)^3 - \sum_{i=0}^{\lfloor Nt \rfloor - 1} \frac{1}{N} \left[ 3 \left( Y_{\frac{i}{N}}^N \right)^2 - 3 \left( Y_{\frac{i}{N}}^N \right)^3 \right] \\ & + \frac{1}{N} \cdot \sum_{i=0}^{\lfloor Nt \rfloor - 1} \frac{1}{N} \left[ 2 \left( Y_{\frac{i}{N}}^N \right)^3 - 3 \left( Y_{\frac{i}{N}}^N \right)^2 + \left( Y_{\frac{i}{N}}^N \right) \right]. \end{aligned}$$

El último renglón converge a 0, debido a que  $0 \leq Y_t^N \leq 1$ , porque cada sumando está entre  $\frac{-3}{N}$  y  $\frac{3}{N}$ , y la suma entera está acotada entre  $\frac{-3}{N} \cdot (\lfloor Nt \rfloor - \lfloor Nh \rfloor)$  y  $\frac{3}{N} \cdot (\lfloor Nt \rfloor - \lfloor Nh \rfloor)$ , haciendo que el multiplicando  $\frac{1}{N}$  haga tender a toda la suma a 0.

Si tomamos una subsucesión  $\{Y^{N_{w_k}}\}_{k \in \mathbb{N}}$  que converge débilmente a  $Z^{w_k}$ , y con ella construimos la sucesión de Martingalas  $\{M^{N_{w_k}}\}_{k \in \mathbb{N}}$ , tenemos que esta sucesión converge débilmente a la Martingala:

$$M_t = \left( Z_t^{N_{w_k}} \right)^3 - \int_0^t 3 \cdot \left( Z_s^{N_{w_k}} \right)^2 - 3 \cdot \left( Z_s^{N_{w_k}} \right)^3 ds$$

Y lo que queríamos demostrar en este teorema era que este proceso es una Martingala.  $\square$

**Teorema 3.2.10.**

$$\lim_{h \rightarrow 0} \frac{1}{h} \cdot \mathbb{E} \left[ (Z_{t+h}^{w_k} - Z_t^{w_k})^3 \mid \mathcal{F}_t^{Z^{w_k}} \right] = 0.$$

*Demostración.*

En el teorema anterior demostramos que el siguiente proceso es una Martingala,

$$(Z_t^{w_k})^3 - \int_0^t 3(Z_s^{w_k})^2 - 3(Z_s^{w_k})^3 ds.$$

Gracias al teorema 3.2.6 ( página 71 ), tenemos que el siguiente proceso también es un Martingala,

$$(Z_t^{w_k})^3 - \int_0^t \mathcal{L}(f)(Z_s^{w_k}) ds$$

tomando  $f$  de la forma  $f(x) = x^3$ , y  $\mathcal{L}$  al generador del proceso  $Z^{w_k}$ .

Entonces tenemos que la diferencia de ambos procesos también es martingala,

$$\begin{aligned} (Z_t^{w_k})^3 - \int_0^t 3(Z_s^{w_k})^2 - 3(Z_s^{w_k})^3 ds &= (Z_t^{w_k})^3 + \int_0^t \mathcal{L}(f)(Z_s^{w_k}) ds \\ &= \int_0^t \mathcal{L}(f)(Z_s^{w_k}) ds - \int_0^t 3(Z_s^{w_k})^2 - 3(Z_s^{w_k})^3 ds = M_t, \end{aligned}$$

a la que llamaremos en esta subsección  $M_t$ .

Sustituyendo  $\mathcal{L}(f)(Z_s^{w_k})$ , obtenemos que

$$\begin{aligned} M_t &= \int_0^t 3 \cdot (Z_s^{w_k})^2 - 3 \cdot (Z_s^{w_k})^3 + \lim_{h \rightarrow 0} \frac{1}{h} \cdot \mathbb{E} \left[ (Z_{s+h}^{w_k} - Z_s^{w_k})^3 \mid \mathcal{F}_s^{Z^{w_k}} \right] ds \\ &= \int_0^t 3(Z_s^{w_k})^2 - 3(Z_s^{w_k})^3 ds \\ &= \int_0^t \lim_{h \rightarrow 0} \frac{1}{h} \cdot \mathbb{E} \left[ (Z_{s+h}^{w_k} - Z_s^{w_k})^3 \mid \mathcal{F}_s^{Z^{w_k}} \right] ds \end{aligned}$$

es martingala, y por lo tanto, para  $r < t$ ,  $\mathbb{E} [ M_t \mid \mathcal{F}_r^{Z^{w_k}} ] = M_r$ ,

$$\begin{aligned} \mathbb{E} \left[ \int_0^t \lim_{h \rightarrow 0} \frac{1}{h} \cdot \mathbb{E} \left[ (Z_{s+h}^{w_k} - Z_s^{w_k})^3 \mid \mathcal{F}_s^{Z^{w_k}} \right] ds \mid \mathcal{F}_r^{Z^{w_k}} \right] \\ = \int_0^r \lim_{h \rightarrow 0} \frac{1}{h} \cdot \mathbb{E} \left[ (Z_{s+h}^{w_k} - Z_s^{w_k})^3 \mid \mathcal{F}_s^{Z^{w_k}} \right] ds \end{aligned}$$

### 3.2. EL GENERADOR DE LA SUBSUCESIONES CONVERGENTES 79

y también sucede que  $\mathbb{E} [ M_t \mid \mathcal{F}_r^{Z^{w_k}} ] = M_r = \mathbb{E} [ M_r \mid \mathcal{F}_r^{Z^{w_k}} ]$ ,

$$\begin{aligned} & \int_0^t \lim_{h \rightarrow 0} \frac{1}{h} \cdot \mathbb{E} \left[ \mathbb{E} \left[ (Z_{s+h}^{w_k} - Z_s^{w_k})^3 \mid \mathcal{F}_s^{Z^{w_k}} \right] \mid \mathcal{F}_r^{Z^{w_k}} \right] ds \\ = & \int_0^r \lim_{h \rightarrow 0} \frac{1}{h} \cdot \mathbb{E} \left[ \mathbb{E} \left[ (Z_{s+h}^{w_k} - Z_s^{w_k})^3 \mid \mathcal{F}_s^{Z^{w_k}} \right] \mid \mathcal{F}_r^{Z^{w_k}} \right] ds \end{aligned}$$

y entonces tenemos que

$$\int_r^t \lim_{h \rightarrow 0} \frac{1}{h} \cdot \mathbb{E} \left[ \mathbb{E} \left[ (Z_{s+h}^{w_k} - Z_s^{w_k})^3 \mid \mathcal{F}_s^{Z^{w_k}} \right] \mid \mathcal{F}_r^{Z^{w_k}} \right] ds = 0$$

Y esto sucede para cualquier  $t$  y cualquier  $r < t$ , es decir, la integral del límite que buscamos es 0 sobre cualquier intervalo sobre el que integremos, por lo tanto, para casi toda  $s$ ,

$$\lim_{h \rightarrow 0} \frac{1}{h} \cdot \mathbb{E} \left[ (Z_{s+h}^{w_k} - Z_s^{w_k})^3 \mid \mathcal{F}_s^{Z^{w_k}} \right] = 0.$$

□

Estamos suponiendo implícitamente que el límite  $\lim_{h \rightarrow 0} \frac{1}{h} \cdot \mathbb{E} \left[ (Z_{s+h}^{w_k} - Z_s^{w_k})^3 \mid \mathcal{F}_s^{Z^{w_k}} \right]$

existe. Se puede calcular el límite directamente demostrando de esta forma que en efecto el límite existe, sin embargo la cantidad de cuentas es excesiva, rebasa 20 páginas de puras cuentas, y por lo tanto no las anexamos a esta tesis.

También se puede demostrar este límite de la misma forma que en la subsección anterior, demostrando que ambas integrales,  $\int_0^t 3(Z_s^{w_k})^2 - 3(Z_s^{w_k})^3 ds$  y  $\int_0^t \mathcal{L}(f)(Z_s^{w_k}) ds$  satisfacen las hipótesis del teorema **3.2.4**, y por lo tanto son iguales, haciendo que el límite de la esperanza condicional sea 0.

- Como  $f(x) = |x|^3$  es convexa y  $Z_t^{w_k}$  solo toma valores entre 0 y 1, tenemos que  $(Z_t^{w_k})^3$  es una submartingala acotada por abajo, continua por la derecha (por ser el límite débil de procesos continuos por la derecha).
- Ambas integrales son adaptadas, continuas (por ser integrales) y valen 0 en el tiempo  $t = 0$ .
- $(Z_t^{w_k})^3 - \int_0^t 3(Z_s^{w_k})^2 - 3(Z_s^{w_k})^3 ds$  es una Martingala.
- $(Z_t^{w_k})^3 - \int_0^t \mathcal{L}(f)(Z_s^{w_k}) ds$  es una Martingala.

- Como  $Z_t^{w_k}$  siempre está entre 0 y 1, sucede que  $3(Z_s^{w_k})^2 - 3(Z_s^{w_k})^3$  siempre es mayor que 0, y por lo tanto, su integral  $\int_0^t 3(Z_s^{w_k})^2 - 3(Z_s^{w_k})^3 ds$  tiene trayectorias no decrecientes.
- Como  $(Z_t^{w_k})^3$  es una submartingala,  $\mathbb{E}[(Z_{s+h}^{w_k})^3 | \mathcal{F}_s^{Z^{w_k}}] - Z_s^{w_k} \geq 0$  y si multiplicamos por  $\frac{1}{h}$ , el límite cuando  $h \rightarrow 0^+$  sigue siendo no negativo, por lo que  $\mathcal{L}(f)(Z_s^{w_k}) \geq 0$  y entonces  $\int_0^t \mathcal{L}(f)(Z_s^{w_k}) ds$  tiene trayectorias no decrecientes.

Como conclusión del teorema 3.2.4 obtenemos que

$$\int_0^t 3(Z_s^{w_k})^2 - 3(Z_s^{w_k})^3 ds = \int_0^t \mathcal{L}(f)(Z_s^{w_k}) ds$$

y por lo tanto, para casi todo tiempo  $t$ ,

$$\begin{aligned} 3 \cdot (Z_s^{w_k})^2 - 3 \cdot (Z_s^{w_k})^3 + \lim_{h \rightarrow 0} \frac{1}{h} \cdot \mathbb{E} \left[ (Z_{s+h}^{w_k} - Z_s^{w_k})^3 \mid \mathcal{F}_s^{Z^{w_k}} \right] ds \\ = 3(Z_s^{w_k})^2 - 3(Z_s^{w_k})^3 \end{aligned}$$

y entonces, para casi todo tiempo  $t$ ,

$$\lim_{h \rightarrow 0} \frac{1}{h} \cdot \mathbb{E} \left[ (Z_{s+h}^{w_k} - Z_s^{w_k})^3 \mid \mathcal{F}_s^{Z^{w_k}} \right] ds = 0. \quad \square$$

### El Problema de Martingala

Ya que hemos calculado todas las esperanzas condicionales de la ecuación (3.7) de la página 70 para casi todo tiempo  $t$ ,

$$\begin{aligned} \mathcal{L}(f)(Z_t^{w_k}) &= \lim_{h \rightarrow 0} \frac{1}{h} \cdot f'(Z_t^{w_k}) \cdot \mathbb{E} \left[ Z_{t+h}^{w_k} - Z_t^{w_k} \mid \mathcal{F}_t^{Z^{w_k}} \right] \\ &\quad + \frac{1}{h} \cdot \frac{1}{2!} f''(Z_t^{w_k}) \cdot \mathbb{E} \left[ (Z_{t+h}^{w_k} - Z_t^{w_k})^2 \mid \mathcal{F}_t^{Z^{w_k}} \right] \\ &\quad + \frac{1}{h} \cdot \frac{1}{3!} f'''(Z_t^{w_k}) \cdot \mathbb{E} \left[ (Z_{t+h}^{w_k} - Z_t^{w_k})^3 \mid \mathcal{F}_t^{Z^{w_k}} \right]. \end{aligned}$$

tenemos que el generador del límite de cualquier subsucesión convergente débilmente,  $Z^{w_k}$ , está dado para todo tiempo  $s$  excepto un conjunto de medida de Lebesgue 0 por

$$\mathcal{L}(f)(Z_s^{w_k}) = \frac{1}{2} \cdot f''(Z_s^{w_k}) \cdot [Z_s^{w_k} \cdot (1 - Z_s^{w_k})].$$

### 3.2. EL GENERADOR DE LA SUBSUCESIONES CONVERGENTES 81

Al generador completo (en todos los tiempos) del proceso  $Z^{w_k}$  le llamaremos  $\mathcal{L}_{Z^{w_k}}$ .

En estos términos, lo que hemos demostrado es que  $\mathcal{L}_{Z^{w_k}}(f)(Z_s^{w_k}) = \frac{1}{2} \cdot f''(Z_s^{w_k}) \cdot [Z_s^{w_k} \cdot (1 - Z_s^{w_k})]$  para todos los tiempos  $s$  excepto un conjunto de medida de Lebesgue 0.

Por el Teorema 3.2.6, tenemos que para cada  $Z^{w_k}$  y para toda función 3 veces diferenciable y con tercera derivada continua, el siguiente proceso es una Martingala,

$$f(Z_t^{w_k}) - \int_0^t \mathcal{L}_{Z^{w_k}}(f)(Z_s^{w_k}) ds.$$

Como  $\mathcal{L}_{Z^{w_k}}(f)(Z_s^{w_k}) = \frac{1}{2} \cdot f''(Z_s^{w_k}) \cdot [Z_s^{w_k} \cdot (1 - Z_s^{w_k})]$  para todos los tiempos  $s$  excepto un conjunto de medida de Lebesgue 0, al integrar a  $\mathcal{L}_{Z^{w_k}}(Z_s^{w_k})$  en la ecuación anterior, nos queda exactamente lo mismo que al integrar a  $\frac{1}{2} \cdot f''(Z_s^{w_k}) \cdot [Z_s^{w_k} \cdot (1 - Z_s^{w_k})]$ . Esto implica que para todos los procesos  $Z^{w_k}$ , el siguiente proceso es una Martingala,

$$f(Z_t^{w_k}) - \int_0^t \frac{1}{2} \cdot f''(Z_s^{w_k}) \cdot [Z_s^{w_k} \cdot (1 - Z_s^{w_k})] ds.$$

Es decir, todos los procesos  $Z^{w_k}$  son solución al Problema de Martingala para el generador  $\mathcal{L}(Z^{w_k}) = \frac{1}{2} \cdot f''(Z_s^{w_k}) \cdot [Z_s^{w_k} \cdot (1 - Z_s^{w_k})]$ .

Recordando la Definición 3.2.2 de lo que es una solución a un Problema de Martingala de la página 71,

**Definición 3.2.3.** Decimos que un proceso estocástico  $X$  que toma valores en el espacio métrico  $E$  y está definido sobre el espacio de probabilidad  $(\Omega, \mathfrak{F}, \mathbb{P})$  es una **solución al problema de Martingala**  $(\mathbf{A}, \mu)$ , ( $A$  siendo un generador y  $\mu$  una distribución inicial) si se cumple que

$$f(X_t) - \int_0^t A(f)(X_s) ds$$

es una Martingala para cualquier  $f$  en el dominio del generador  $A$ , y la distribución inicial del proceso está dada por  $\mu$

$$\mathbb{P}(X_0^{-1}) = \mu.$$

Hemos demostrado el siguiente Teorema:

**Teorema 3.2.11.** *Todos los procesos límite  $Z^{w_k}$  son solución al Problema de Martingala del generador  $\mathcal{L}(f)(Z_s^{w_k}) = \frac{1}{2} \cdot f''(Z_s^{w_k}) \cdot [Z_s^{w_k} \cdot (1 - Z_s^{w_k})]$ , sobre el conjunto de funciones 3 veces derivables y con tercera derivada acotada. Es decir, el siguiente proceso es una Martingala,*

$$f(Z_t^{w_k}) - \int_0^t \frac{1}{2} \cdot f''(Z_s^{w_k}) \cdot [Z_s^{w_k} \cdot (1 - Z_s^{w_k})] ds.$$

Ahora demostraremos que cualesquiera dos procesos que sean solución a ese Problema de Martingala, con ese generador y sobre ese conjunto de funciones (o uno más chico), tienen la misma distribución. Por lo tanto, demostraremos que la sucesión entera  $\{Y^N\}_{N \in \mathbb{N}}$  converge débilmente.

### 3.3. Unicidad de la solución al Problema de Martingala de $\mathcal{L}$

**Definición 3.3.1.** *Decimos que un problema de Martingala  $(A, \mu)$  está **bien expuesto**, si tiene una única solución. Decimos que la solución es **única**, si cualesquiera dos soluciones tienen las mismas distribuciones finitodimensionales.*

El generador  $\mathcal{L}(f)(Z_s^{w_k}) = \frac{1}{2} \cdot f''(Z_s^{w_k}) \cdot [Z_s^{w_k} \cdot (1 - Z_s^{w_k})]$  es el generador de un proceso de difusión.

La unicidad de la solución a un Problema de Martingala del generador de una difusión se puede demostrar con el siguiente teorema:

**Teorema 3.3.1** ( Teorema 2.3, Ethier y Kurtz [?], página 372 ).

Sean  $-\infty \leq r_0 < r_1 \leq \infty$ ,  $I = [r_0, r_1] \cap \mathbb{R}$ .

Tomemos el generador

$$G f = \frac{1}{2} \cdot a(t, x) \cdot \frac{d^2}{dx^2} f + b(t, x) \cdot \frac{d}{dx} f$$

con  $a : [0, \infty) \times I \rightarrow \mathbb{R}^+ \cup \{0\}$  y  $b : [0, \infty) \times I \rightarrow \mathbb{R}$  funciones localmente acotadas y Borel-medibles, tales que  $a$  y  $b$  son funciones Lipschitz continuas sobre los conjuntos  $\{x \mid |x| \leq n\}_{n \in \mathbb{N}}$  y uniformemente continuas en los conjuntos  $\{t \mid 0 \leq t \leq t_0\}_{t_0 \in \mathbb{R}^+}$ .

### 3.3. UNICIDAD DE LA SOLUCIÓN AL PROBLEMA DE MARTINGALA DE L83

Si el generador  $G$  satisface para toda  $x \in I$  y para toda  $t_0 > 0$  las ecuaciones

$$\begin{aligned} a(t_0, r_i) &= 0 \leq b(t_0, r_i) \cdot (-1)^i \\ \text{para } i = 0, i = 1 \quad \text{si } |r_i| < \infty, \end{aligned}$$

y existe  $K$  tal que para todo  $x \in I$  y todo  $t > 0$

$$\begin{aligned} |a(t, x)| &\leq K \cdot (1 + |x|^2) \\ x \cdot b(t, x) &\leq K \cdot (1 + |x|^2) \end{aligned}$$

entonces el Problema de Martingala para  $\{(f, Gf) : f \in C_c^\infty(I)\}$  tiene solución única (está bien expuesto).

**Lema 3.3.1.** El generador

$$\mathcal{L}(f)(X_t) = \frac{1}{2} \cdot f''(X_t) \cdot [X_t \cdot (1 - X_t)]$$

satisface las ecuaciones e hipótesis del teorema anterior, y por lo tanto, su Problema de Martingala tiene solución única.

*Demostración.*

Para nosotros  $I = [0, 1]$ .

Como los coeficientes de nuestro generador ( $a(t, x) = x(1 - x)$ ,  $b(t, x) = 0$ ) no dependen del tiempo, las ecuaciones se simplifican mucho.

Las funciones  $a(t, x) = x(1 - x) : [0, \infty) \times [0, 1] \rightarrow \mathbb{R}^+ \cup \{0\}$  y  $b(t, x) = 0 : [0, \infty) \times [0, 1] \rightarrow \mathbb{R}$  son funciones acotadas y Borel-medibles. Además son Lipschitz continuas sobre el conjunto entero  $[0, 1]$  y por lo tanto sobre todos los conjuntos de las hipótesis del teorema. También son constantes con respecto al tiempo y por lo tanto son uniformemente continuas en  $[0, t_0]$  para todo  $t_0$ . Las funciones también satisfacen las ecuaciones del teorema:

Para toda  $x \in [0, 1]$  y  $t_0 > 0$  tenemos que

$$\begin{aligned} a(t, 0) &= 0 \cdot (1 - 0) = 0 \\ b(t, 0) &= 0 \geq 0 \\ a(t, 1) &= 1 \cdot (1 - 1) = 0 \\ b(t, 1) \cdot (-1) &= 0 \geq 0 \end{aligned}$$

y con  $K = 1$  tenemos que

$$\begin{aligned} |a(t, x)| &= x(1 - x) \leq 1 \leq (1 + |x|^2) \\ x \cdot b(t, x) &= 0 \leq (1 + |x|^2) \end{aligned}$$

□

Por lo tanto, el Problema de Martingala para  $\{ (f, \mathcal{L}f) \text{ tales que } f \in C_c^\infty([0, 1]) \}$  tiene solución única (está bien expuesto).

Y como  $C_c^\infty([0, 1])$  es un subconjunto de  $C^4([0, 1])$ , entonces el Problema de Martingala para  $\{ (f, \mathcal{L}f) \text{ tales que } f \in C^4([0, 1]) \}$  tiene solución única.

Como vimos en las secciones anteriores, todos los límites débiles de las subsubsecciones  $Z^{w_k}$  son solución del Problema de Martingala para  $\{ (f, \mathcal{L}f) \text{ tales que } f \in C^4([0, 1]) \}$ , y como la solución es única, entonces todos los límites  $Z^{w_k}$  tienen la misma distribución, y por lo tanto, la sucesión entera  $\{Y^N\}_{N \in \mathbb{N}}$  converge débilmente a este proceso de difusión con generador  $\mathcal{L}$ .

Le llamaremos  $Z$  al límite débil de la sucesión de procesos  $\{Y^N\}_{N \in \mathbb{N}}$ . Es generalmente conocido como el proceso de difusión de Wright-Fisher con dos tipos de alelos, sin selección y sin mutación.

### 3.4. La Esperanza del Tiempo de Absorción

En esta sección calcularemos la esperanza del tiempo de absorción de  $Z$  (el proceso de difusión de Wright-Fisher con dos tipos de alelos, sin selección y sin mutación) utilizando su generador  $\mathcal{L}(f)(Z_t) = \frac{1}{2} \cdot f''(Z_t) \cdot Z_t(1 - Z_t)$ .

Dado un proceso de difusión, se tiene asociado un semigrupo de Feller, el cual toma la forma

$$T_s f(Z_t) = \mathbb{E} \left[ f(Z_{t+s}) \mid \mathcal{F}_t^Z \right].$$

El semigrupo de un proceso junto con su distribución inicial determinan las distribuciones finito dimensionales del proceso. Además, a partir de ellos, derivando con respecto a 's', se obtiene el generador del proceso.

No profundizaremos en las propiedades de los semigrupos de los procesos, ya que no las necesitaremos.

Para agilizar la notación, cuando tomemos al proceso  $Z$  con distribución inicial  $\delta_{p_0}$ , para referirnos a  $T_t(f)(Z_0) = \mathbb{E}[f(Z_t) \mid \mathcal{F}_0^Z]$  utilizaremos la notación  $T_t(f)(p_0)$ .

Esta es una buena notación porque como la distribución inicial es  $\delta_{p_0}$ ,  $\mathbb{P}(Z_0 = p_0) = 1$ , y tenemos que

$$T_t(f)(Z_0) = \mathbb{E}[f(Z_t) \mid Z_0] = \mathbb{E}[f(Z_t) \mid Z_0 = p_0] = \mathbb{E}_{p_0}(f(Z_t)).$$

**Teorema 3.4.1.** Sea  $\tau$  el tiempo de absorción de  $Z$  al conjunto  $\{1, 0\}$ . Sea  $\delta_{p_0}$  la distribución inicial con  $0 < p_0 < 1$ . Entonces,

$$\mathbb{E}(\tau) = -2 \cdot \left[ p_0 \cdot \log(p_0) + (1 - p_0) \cdot \log(1 - p_0) \right].$$

Como  $\mathbb{P}(Z_0 = p_0) = 1$ , tenemos que

$$\mathbb{E}(\tau) = \mathbb{E}(\mathbb{E}_{Z_0}(\tau)) = \mathbb{E}_{Z_0=p_0}(\tau) = \int_0^\infty \mathbb{P}_{p_0}(\tau > t) dt$$

y como  $\{0\}$  y  $\{1\}$  son los estados absorbentes de los cuales  $\tau$  es el tiempo de absorción,

$$= \int_0^\infty \mathbb{P}_{p_0}(Z_t \notin \{0, 1\}) dt = \int_0^\infty \mathbb{E}_{p_0}(\mathbb{I}_{\{Z_t \in (0,1)\}}) dt,$$

ahora, recordando que  $T_t(f)(Z_0) = \mathbb{E}[f(Z_t) | \mathcal{F}_0^Z] = \mathbb{E}_{p_0}[f(Z_t)]$ ,

$$= \int_0^\infty T_t(\mathbb{I}_{\{x \in (0,1)\}})(Z_0) dt = g(Z_0).$$

A la función  $\mathbb{I}_{\{x \in (0,1)\}}$  le llamaremos  $h(x)$  y a la función  $\int_0^\infty T_t(h)(Z_0) dt$  le llamaremos  $g(Z_0)$ . Si  $Z_0$  se distribuye como  $\delta_x$ , a  $g(Z_0)$  la denotaremos por  $g(x)$ .

A continuación demostraremos que la función  $g(x)$  satisface la siguiente ecuación diferencial, y después la resolveremos para encontrar la forma explícita de  $\mathbb{E}(\tau) = g(p_0)$ .

$$\begin{aligned} \mathcal{L} \left( \int_0^\infty T_s h(x) ds \right) &= \mathcal{L}(g(x)) = \frac{1}{2} \cdot x \cdot (1 - x) \cdot \frac{d^2}{(dx)^2} g(x) = -1 \\ \text{para todo } x \in [0, 1], \text{ y} & \qquad \qquad \qquad g(0) = 0 \qquad (3.8) \\ & \qquad \qquad \qquad g(1) = 0. \end{aligned}$$

En el libro de Ethier y Kurtz [?] podemos encontrar el siguiente teorema

**Teorema 3.4.2** ( Proposición 1.5, Ethier y Kurtz [?], página 9 ).

Sea  $\{T_t\}$  un semigrupo fuertemente continuo sobre  $B([0, 1])$  con generador  $A$ .

Si  $f \in B([0, 1])$  y  $t \geq 0$ , entonces  $\int_0^t T_s(f) ds \in \text{Dominio}(A)$  y

$$A \left( \int_0^t T_s f ds \right) = T_t f - f.$$

Lo que obtenemos de este teorema es que para nuestra función  $h(x)$ , el semigrupo fuertemente continuo generado por el proceso de difusión  $Z$  y su correspondiente generador  $\mathcal{L}$ , sucede que

$$\mathcal{L} \left( \int_0^t T_s h \, ds \right) = T_t h - h. \quad (3.9)$$

**Lema 3.4.1.**

$$\lim_{t \rightarrow \infty} \mathcal{L} \left( \int_0^t T_s h \, ds \right) (x) = \mathcal{L} \left( \int_0^\infty T_s h \, ds \right) (x)$$

*Demostración.*

Para toda  $p_0 \in [0, 1]$ ,

$$\begin{aligned} & \lim_{t \rightarrow \infty} \mathcal{L} \left( \int_0^t T_s h \, ds \right) (p_0) \\ &= \lim_{t \rightarrow \infty} \lim_{b \rightarrow 0} \frac{1}{b} \cdot \left[ T_b \left( \int_0^t T_s h \, ds \right) (p_0) - \left( \int_0^t T_s h \, ds \right) (p_0) \right] \\ &= \lim_{b \rightarrow 0} \frac{1}{b} \cdot \left[ \lim_{t \rightarrow \infty} \mathbb{E}_{\mathbf{p}_0} \left( \int_0^t T_s h \, ds (Z_b) \right) - \int_0^\infty T_s h(p_0) \, ds \right] \\ &^* = \lim_{b \rightarrow 0} \frac{1}{b} \cdot \left[ \mathbb{E}_{\mathbf{p}_0} \left( \lim_{t \rightarrow \infty} \int_0^t T_s h(Z_b) \, ds \right) - \int_0^\infty T_s h(p_0) \, ds \right] \\ &= \lim_{b \rightarrow 0} \frac{1}{b} \cdot \left[ T_b \left( \int_0^\infty T_s h \, ds \right) (p_0) - \int_0^\infty T_s h \, ds (p_0) \right] \\ &= \mathcal{L} \left( \int_0^\infty T_s h \, ds \right) (p_0). \end{aligned}$$

La igualdad marcada con \* se da por el teorema de convergencia monótona. La sucesión de funciones  $\{\int_0^k T_s h(Z_b) \, ds\}_{k \in \mathbb{N}}$  es no decreciente para cualquier valor de  $Z_b$ , porque  $T_s h(Z_b)$  es siempre mayor o igual que 0, debido a que  $T_s h(Z_b) = \mathbb{E} \left( h(Z_{b+s}) \mid \mathcal{F}_b^Z \right)$  y  $h(x)$  es una función indicadora que solo toma los valores 1 y 0. Entonces, por el teorema de convergencia monótona, sucede que

$$\lim_{t \rightarrow \infty} \mathbb{E} \left( \int_0^t T_s h(Z_b) \, ds \right) = \mathbb{E} \left( \lim_{t \rightarrow \infty} \int_0^t T_s h(Z_b) \, ds \right).$$

□

**Lema 3.4.2.**

$$\mathcal{L} \left( \int_0^\infty T_s h \, ds \right) (p_0) = \lim_{t \rightarrow \infty} T_t h(p_0) - h(p_0) = -h(p_0).$$

*Demostración.*

Debido al Lema anterior, y a la ecuación (3.9) tenemos que

$$\mathcal{L} \left( \int_0^\infty T_s h \, ds \right) (p_0) = \lim_{t \rightarrow \infty} \mathcal{L} \left( \int_0^t T_s h \, ds \right) (p_0) = \lim_{t \rightarrow \infty} T_t h(p_0) - h(p_0)$$

Ahora vamos a demostrar que  $\lim_{t \rightarrow \infty} T_t h(p_0) = 0$ .

Tenemos que

$$\lim_{t \rightarrow \infty} \mathbb{P}(Z_t \in (0, 1)) = 0$$

porque  $Z$  es el límite débil de  $\{Y^N\}_{N \in \mathbb{N}}$ , y para todos los procesos  $Y^N$ , los estados  $\{1\}$  y  $\{0\}$  son los únicos estados absorbentes, todos los demás estados son transitorios y por lo tanto,

$$\lim_{t \rightarrow \infty} \mathbb{P}(Y_t^N \in (0, 1)) = 0.$$

Entonces,

$$\begin{aligned} 0 &\leq \lim_{t \rightarrow \infty} T_t h(p_0) = \lim_{t \rightarrow \infty} \mathbb{E}_{p_0}(h(Z_t)) \\ &= \lim_{t \rightarrow \infty} 1 \cdot \mathbb{P}_{p_0}(Z_t \in (0, 1)) + 0 \cdot \mathbb{P}_{p_0}(Z_t \in \{1, 0\}) \\ &\leq \lim_{t \rightarrow \infty} \mathbb{P}(Z_t \in (0, 1)) \leq 0. \end{aligned}$$

Es decir,

$$\lim_{t \rightarrow \infty} T_t h(p_0) = 0.$$

Con lo cual se concluye que

$$\mathcal{L} \left( \int_0^\infty T_s h \, ds \right) (p_0) = \lim_{t \rightarrow \infty} T_t h(p_0) - h(p_0) = -h(p_0).$$

□

88CAPÍTULO 3. CONVERGENCIA DÉBIL A UN PROCESO DE DIFUSIÓN

Regresando a la ecuación diferencial (3.8),

$$\mathcal{L}\left(\int_0^\infty T_s h(x) ds\right) = \mathcal{L}(g(x)) = \frac{1}{2} \cdot x \cdot (1-x) \cdot \frac{d^2}{(dx)^2} g(x) = -1$$

para todo  $x \in [0, 1]$ , y

$$\begin{aligned} g(0) &= 0 \\ g(1) &= 0. \end{aligned}$$

Como  $h(x)$  es igual a 1 para todo  $x \in (0, 1)$ ,  $h(0) = 0$  y  $h(1) = 0$ , el lema anterior demuestra que  $g(x) = \int_0^\infty T_s h(x) ds$  satisface la ecuación diferencial anterior. Por lo tanto,

$$\begin{aligned} -1 &= \frac{1}{2} \cdot p \cdot (1-p) \cdot \frac{d^2}{(dp)^2} g(p) \\ \frac{d^2}{(dp)^2} g(p) \cdot \frac{1}{2} &= -\frac{1}{p(1-p)} = -\frac{1}{p} - \frac{2}{1-p} \\ \frac{d}{dp} g(p) \cdot \frac{1}{2} &= -1 - \log(p) + 1 + \log(1-p) + C \\ g(p) \cdot \frac{1}{2} &= -p \cdot \log(p) - (1-p) \cdot \log(1-p) + C \cdot p + K \end{aligned}$$

$$0 = g(0) = -2 \cdot [0 \cdot \log(0) + (1-0) \cdot \log(1-0) + C \cdot 0 + K]$$

$$0 = -2 \cdot K$$

$$0 = g(1) = -2 \cdot [1 \cdot \log(1) + (1-1) \cdot \log(1-1) + C \cdot 1 + 0]$$

$$0 = -2 \cdot C$$

Y entonces podemos concluir que la esperanza del tiempo de absorción de  $Z$  al conjunto  $\{1, 0\}$  cuando el proceso empieza en  $Z_0 = p_0$  es

$$\mathbb{E}_{p_0}(\tau) = g(p_0) = -2 \cdot [p_0 \cdot \log(p_0) + (1-p_0) \cdot \log(1-p_0)].$$

Se puede demostrar, se hace en el libro de Ethier y Kurtz [?] en el Corolario 2.6 página 419, que la sucesión  $\{a_n\}_{n \in \mathbb{N}}$  en donde cada  $a_n$  es la esperanza del tiempo de absorción de la cadena  $\frac{X^N}{N}$  converge a la esperanza del tiempo de absorción de  $Z$  cuando  $N$  tiende a infinito.

## Apéndice A

# Momentos y otros cálculos de la variable $X_{n+1}$ distribuida como una $Binomial(N, \frac{X_n}{N})$

### A.1. Momentos.

Si  $X_{n+1}^N$  se distribuye como una variable aleatoria  $Binomial(N, \frac{X_n^N}{N})$  conocemos ya todos sus momentos.

Para facilitar la notación, llamaremos  $X$  a  $X_{n+1}^N$  y llamaremos  $p$  a  $\frac{X_n^N}{N}$ .

La función generadora de momentos de una variable aleatoria  $X \sim Binomial(N, p)$  y sus primeros tres momentos son:

$$\begin{aligned}
 M(t) &= (p \cdot e^t + 1 - p)^N \\
 M'(t) &= N \cdot p \cdot e^t \cdot (p \cdot e^t + 1 - p)^{N-1} \\
 M'(0) &= N \cdot p \\
 M''(t) &= N \cdot (N-1) \cdot p^2 \cdot e^{2t} \cdot (p \cdot e^t + 1 - p)^{N-2} \\
 &\quad + N \cdot p \cdot e^t \cdot (p \cdot e^t + 1 - p)^{N-1} \\
 M''(0) &= N \cdot (N-1) \cdot p^2 + N \cdot p \\
 M'''(t) &= N \cdot (N-1) \cdot (N-2) \cdot p^2 \cdot e^{2t} \cdot (p \cdot e^t + 1 - p)^{N-3} \\
 &\quad + 3 \cdot N \cdot (N-1) \cdot p^2 \cdot e^{2t} \cdot (p \cdot e^t + 1 - p)^{N-1} \\
 &\quad + N \cdot p \cdot e^t \cdot (p \cdot e^t + 1 - p)^{N-1} \\
 M'''(0) &= N \cdot (N-1) \cdot (N-2) \cdot p^2 + 3 \cdot N \cdot (N-1) \cdot p^2 + N \cdot p
 \end{aligned}$$

A partir de los momentos de  $X_{n+1}^N$  se calculan directamente los momentos de la variable  $\frac{X_{n+1}^N}{N}$ .

$$\begin{aligned}
 \mathbb{E}_p \left( \frac{X}{N} \right) &= \frac{1}{N} \cdot \mathbb{E}_p(X) = \frac{1}{N} \cdot N \cdot p \\
 &= p
 \end{aligned} \tag{A.1}$$

$$\begin{aligned}
 \mathbb{E}_p \left( \left( \frac{X}{N} \right)^2 \right) &= \frac{1}{N^2} \cdot \mathbb{E}_p(X^2) = \frac{1}{N^2} \left( N(N-1)p^2 + Np \right) \\
 &= \frac{N-1}{N} \cdot p^2 + \left( \frac{1}{N} \right) \cdot p
 \end{aligned} \tag{A.2}$$

$$\begin{aligned}
 \mathbb{E}_p \left( \left( \frac{X}{N} \right)^3 \right) &= \frac{1}{N^3} \cdot \mathbb{E}_p(X^3) \\
 &= \frac{1}{N^3} \left( N(N-1)(N-2)p^3 + 3N(N-1)p^2 + Np \right) \\
 &= \frac{N-1}{N} \cdot \frac{N-2}{N} \cdot p^3 + 3 \cdot \left( \frac{1}{N} \right) \cdot \frac{N-1}{N} \cdot p^2 \\
 &\quad + \left( \frac{1}{N^2} \right) \cdot p.
 \end{aligned} \tag{A.3}$$

## A.2. La esperanza condicional del incremento en un paso al cuadrado.

Ahora calculamos la esperanza  $\mathbb{E}_{x_n^N} \left[ \left( \frac{X_{n+1}^N}{N} - \frac{X_n^N}{N} \right)^3 \right]$ .

## A.2. LA ESPERANZA CONDICIONAL DEL INCREMENTO EN UN PASO AL CUADRADO.91

Le seguiremos llamando  $X$  a  $X_{n+1}^N$  y  $p$  a  $\frac{X_n^N}{N}$ .

$$\mathbb{E}_p \left[ \left( \frac{X}{N} - p \right)^3 \right] = \mathbb{E}_p \left[ \left( \frac{X}{N} \right)^3 - 3p \left( \frac{X}{N} \right)^2 + 3p^2 \frac{X}{N} - p^3 \right]$$

Sustituyendo en ésta última ecuación los valores obtenidos en (A.1), (A.2), (A.3), con la función generadora de momentos, y ordenándolos como polinomios respecto a  $p$ .

$$\begin{aligned} \mathbb{E}_p \left[ \left( \frac{X}{N} - p \right)^3 \right] &= \mathbb{E}_{\frac{X_n^N}{N}} \left[ \left( \frac{X_{n+1}^N}{N} - \frac{X_n^N}{N} \right)^3 \right] \\ &= \frac{N-1}{N} \cdot \frac{N-2}{N} \cdot p^3 - 3 \cdot p \cdot \frac{N-1}{N} \cdot p^2 + 3 \cdot p^2 \cdot p - p^3 \\ &\quad + 3 \cdot \left( \frac{1}{N} \right) \cdot \frac{N-1}{N} \cdot p^2 - 3 \cdot p \cdot \left( \frac{1}{N} \right) \cdot p \\ &\quad + \left( \frac{1}{N^2} \right) \cdot p \\ &= p^3 \cdot \left( \frac{N-1}{N} \cdot \frac{N-2}{N} - 3 \cdot \frac{N-1}{N} + 3 - 1 \right) \\ &\quad + 3 \cdot p^2 \cdot \left( \frac{1}{N} \right) \cdot \left( \frac{N-1}{N} - 1 \right) \\ &\quad + \left( \frac{1}{N^2} \right) \cdot p \\ &= p^3 \cdot \left( \frac{N^2 - 3N + 2 - 3 \cdot (N^2 - N) + 3 \cdot N^2 - N^2}{N^2} \right) \\ &\quad + 3 \cdot p^2 \cdot \frac{-1}{N^2} + p \cdot \frac{1}{N^2} \\ &= p^3 \cdot \frac{2}{N^2} - 3 \cdot p^2 \cdot \frac{1}{N^2} + p \cdot \frac{1}{N^2} \\ &= \frac{1}{N^2} \cdot (2p^3 - 3p^2 + p) \end{aligned}$$

Es decir,

$$\mathbb{E}_{\frac{X_n^N}{N}} \left[ \left( \frac{X_{n+1}^N}{N} - \frac{X_n^N}{N} \right)^3 \right] = \frac{1}{N^2} \cdot \left( 2 \cdot \left( \frac{X_n^N}{N} \right)^3 - 3 \cdot \left( \frac{X_n^N}{N} \right)^2 + \frac{X_n^N}{N} \right). \quad (\text{A.4})$$

Esta expresión tiende a 0 con orden  $\mathcal{O}(N^2)$  cuando  $N \rightarrow \infty$ . Como consecuencia tenemos que:

$$\lim_{N \rightarrow \infty} N \cdot \mathbb{E}_{\frac{X_n^N}{N}} \left( \left( \frac{X_{n+1}^N}{N} - \frac{X_n^N}{N} \right)^3 \right) = 0.$$

### A.3. La esperanza condicional del incremento en un paso al cubo

Ahora calculamos la esperanza  $\mathbb{E}_{X_n^N} \left[ \left( \frac{X_{n+1}^N}{N} - \frac{X_n^N}{N} \right)^3 \right]$ .

Le seguiremos llamando  $X$  a  $X_{n+1}^N$  y  $p$  a  $\frac{X_n^N}{N}$ .

Tenemos que

$$\mathbb{E}_p \left[ \left( \frac{X}{N} - p \right)^2 \right] = \mathbb{E}_p \left[ \left( \frac{X}{N} \right)^2 - 2p \left( \frac{X}{N} \right) + p^2 \right].$$

Sustituyendo en ésta ecuación los valores obtenidos en (A.1) y (A.2) con la función generadora de momentos obtenemos:

$$\begin{aligned} \mathbb{E}_p \left[ \left( \frac{X}{N} - p \right)^2 \right] &= \mathbb{E}_{\frac{X_n^N}{N}} \left[ \left( \frac{X_{n+1}^N}{N} - \frac{X_n^N}{N} \right)^2 \right] \\ &= \frac{N-1}{N} \cdot p^2 + \left( \frac{1}{N} \right) \cdot p - 2 \cdot p \cdot p + p^2 \\ &= p^2 \cdot \frac{-1}{N} + \frac{1}{N} \cdot p \\ &= \frac{1}{N} \cdot p \cdot (1 - p) \end{aligned}$$

Es decir,

$$\mathbb{E}_{\frac{X_n^N}{N}} \left[ \left( \frac{X_{n+1}^N}{N} - \frac{X_n^N}{N} \right)^3 \right] = \frac{1}{N} \cdot \frac{X_n^N}{N} \cdot \left( 1 - \frac{X_n^N}{N} \right). \quad (\text{A.5})$$

## Apéndice B

### Notación.

Si estamos estudiando a la cadena  $X_n$  :

$$\mathbb{P}^n(x, y) = \mathbb{P}(X_n = y | X_0 = x).$$

$$\mathbb{P}_x(A) = \mathbb{P}(A | X_0 = x), \mathbb{E}_x(A) = \mathbb{E}(A | X_0 = x).$$

$T_A$  = tiempo de llegada al conjunto  $A$ .

$$T_A = \min\{n \in \mathbb{N} | X_n \in A\}.$$

$\rho_{xy} = \mathbb{P}(X_n \text{ llegue en algún tiempo en el futuro al estado } y \text{ partiendo en el presente del estado } x) = \mathbb{P}(X_m = y \text{ para alguna } m | X_0 = x).$

$$\rho_{xy} = \mathbb{P}_x(T_y < \infty).$$

Llamamos “**recurrente**” a un estado  $y$  si  $\rho_{yy} = 1$ ; y lo llamaremos “**transitivo**” si  $\rho_{yy} < 1$ .

$N(y)$  = la cantidad de veces que la cadena  $X_n$  está en el estado  $y$  a lo largo del tiempo.

$$N(y) = \sum_{i=1}^{\infty} \mathbb{I}_y(X_i) \quad \text{donde } \mathbb{I}_y(x) = \delta_{y,x} = \{1 \text{ siempre que } x = y, \quad 0 \text{ si no}\}.$$

$$G(x, y) = \mathbb{E}_x(N(y)) = \sum_{n=1}^{\infty} \mathbb{P}^n(x, y).$$

$G(x, y)$  = número esperado de visitas a  $y$  empezando en  $x$ .

Decimos que **un estado  $y$  es accesible desde  $x$**  si  $\rho_{xy} > 0$ . También decimos que **se puede acceder a  $x$  desde  $y$**  si  $\rho_{xy} > 0$ .

Un conjunto es **cerrado** si para cada  $x \in C$  solo se puede acceder a elementos de  $C$ .

Un conjunto es **irreducible** si para cada  $y, x \in C$  se puede acceder a  $x$  desde  $y$ .

$\rho_C(x) = \mathbb{P}_x(T_C < \infty)$  con  $C$  un conjunto cerrado e irreducible de estados recurrentes.

$$\rho_C(x) = \mathbb{P}(X_n \text{ llegue en algún momento al conjunto } C | \text{ empezamos en } x)$$



# Bibliografía

- [EK] Stewart N. Ethier y Thomas Kurtz (1986) “*Markov Processes, Characterization and Convergence*” John Wiley & sons, Inc.
- [B] Patrick Billingsley (1999) “*Convergence of Probability Measures*” John Wiley & sons, Inc.
- [D] Donald Dawson (2009) “*Preliminary Version Notes for Lectures on Stochastic Population Systems*” PIMS-UBC 2009 Summer School.
- [HPS] Paul G. Hoel, Sidney C. Port y Charles J. Stone (1972) “*Introduction to Stochastic Processes*” Houghton Mifflin Company.