



UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

FACULTAD DE CIENCIAS



EL PROCESO DE LINEAS DE DESCENDENCIA EN
EL ESTUDIO GENETICO DE POBLACIONES

T E S I S
QUE PARA OBTENER EL TITULO DE
MATEMATICO
PRESENTA:

PAULA SANCHEZ ROMEU



FACULTAD DE CIENCIAS
UNAM

DIRECTOR DE TESIS:
DRA. ELIANE REGINA RODRIGUES CALONI



2002

FACULTAD DE CIENCIAS
SECCION ESCOLAR



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.



UNIVERSIDAD NACIONAL
AUTÓNOMA DE
MÉXICO

M. EN C. ELENA DE OTEYZA DE OTEYZA
Jefa de la División de Estudios Profesionales de la
Facultad de Ciencias
Presente

Comunicamos a usted que hemos revisado el trabajo escrito:

El proceso de líneas de descendencia en el estudio genético de poblaciones.

realizado por Paula Sánchez Romeu

con número de cuenta 9350413-9, quien cubrió los créditos de la carrera de: Matemáticas

Dicho trabajo cuenta con nuestro voto aprobatorio.

A t e n t a m e n t e

Director de Tesis

Propietario Dra. Eliane Regina Rodríguez Caloni *Eliane Regina Rodríguez*

Propietario Dra. Laura Ortiz Bobadilla *Laura Ortiz B.*

Propietario Dr. Pablo Padilla Longoria *Pablo Padilla L.*

Suplente M. en C. Beatriz Eugenia Rodríguez Fernández *Beatriz Rodríguez F.*

Suplente M. en A.P. María del Pilar Alonso Reyes *María del Pilar Alonso Reyes*

Consejo Departamental de Ma



JCAJ
M. en C. José Antonio **FACULTAD DE CIENCIAS**
CONSEJO DEPARTAMENTAL
DE
MATEMÁTICAS

Agradecimientos

Quiero agradecer primeramente a mis padres, por el cariño, apoyo y consejos que me han dado siempre, y que me han hecho ser la persona que soy.

Al Dr. David Romeu, gracias por ser siempre una guía, y por darme perspectiva y claridad en tantos momentos de mi vida.

A la Dra. Laura Ortiz y al Dr. Ernesto Rosales, un millón de gracias por su confianza, su apoyo incondicional y por valorar mi trabajo. Gracias por reiterarme cada día el valor del esfuerzo y empujarme siempre hacia adelante.

A la Dra. Eliane Rodrigues, gracias por hacer este trabajo posible y por todas esas horas de aprendizaje en lo personal y lo académico.

Al Dr. Luis Gorostiza Ortega, gracias por todos sus consejos y orientación, y por el apoyo que me otorgó a través del SNI.

A la Maestra Beatriz Rodríguez, gracias por inculcarme el amor a la Probabilidad y los Procesos Estocásticos, a través de la agradable experiencia de tus clases.

A mis maestros: Laura Ortiz y Ernesto Rosales, Beatriz Rodríguez, Pilar Alonso, Jefferson King, Laura Pastrana, Flor de María Aceff, Pablo Barrera, gracias por hacer la diferencia, por cursos que realmente valieron la pena y por sembrar una semilla en mí.

Al Instituto de Matemáticas de la UNAM, gracias por ser un segundo hogar y por todo el apoyo brindado a lo largo de mi proyecto académico.

A mis compañeros del Instituto de Matemáticas, gracias por los buenos ratos de convivencia y compañerismo.

Agradezco el apoyo del proyecto PAPIIT no. IN113800 proporcionado a través del Dr. Ernesto Rosales.

Índice General

Introducción	v
1 Cadenas de Markov en tiempo continuo	1
1.1 Resultados preliminares	1
1.2 Probabilidades límite	7
1.3 Ecuaciones diferenciales de Kolmogorov	10
1.4 Proceso de nacimiento y muerte	15
1.4.1 Generalidades	15
1.4.2 Ecuaciones diferenciales de Kolmogorov	16
1.4.3 Probabilidades límite	21
2 Proceso de difusión	25
2.1 Resultados preliminares	25
2.2 Ecuaciones diferenciales	29
2.3 Distribución estacionaria	31
2.3.1 Cálculo de la distribución estacionaria	32
2.4 Comportamiento en la frontera	34
2.4.1 Clasificación de fronteras	36
2.4.2 Fronteras de entrada y la distribución estacionaria	38
2.5 Convergencia a una difusión	39
3 El modelo de Wright-Fisher y el modelo de Moran	43
3.1 Modelo de Wright-Fisher	44
3.1.1 Descripción del modelo	44
3.2 Modelo de Moran	46
3.2.1 Descripción del modelo	46
3.2.2 Distribución límite	49

3.3	Aproximación por difusión	55
3.3.1	Difusión para el modelo de Wright-Fisher	56
3.3.2	Difusión para el modelo de Moran	63
4	El proceso n-coalescente de Kingman y el proceso de líneas de descendencia	67
4.1	El proceso n-coalescente de Kingman	68
4.2	El proceso de líneas de descendencia	76
4.2.1	Generalidades	76
4.2.2	Descripción del proceso	79
4.2.3	Teorema principal	81
A	Apéndice A	87
B	Apéndice B	91

Introducción

Una de las mayores áreas de investigación en la actualidad es la referente al genoma humano. A mediados de los años 80 se inicia el llamado *Proyecto Genoma* como una combinación de investigación bioquímica, genética y biológica-molecular, con el objetivo de determinar la totalidad de la secuencia de componentes químicos que integran el genoma humano (así como el de animales y plantas). El conocimiento de esta secuencia permitirá identificar los genes relevantes y saber cómo están regulados, con lo cual será posible, por ejemplo, combatir varias de las enfermedades incurables más severas del hombre, que provienen de defectos genéticos (por ejemplo, las mutaciones). ¹

De los avances en esta decodificación del genoma humano, ha surgido la posibilidad de medir la evolución de la especie humana, rastreando hacia atrás en el tiempo las líneas ancestrales de la población mundial actual a partir de la información de los genes de individuos que la conforman. Esto ha sido posible gracias a estudios que se basan en:

- (1) la propiedad de los genes de presentarse en diferentes versiones o formas alternativas, llamadas *alelos*,
- (2) la hipótesis de que las mutaciones genéticas se acumulan en las moléculas de ADN a una tasa gradual y constante,

lo cual ha permitido comparar secuencias de ADN de cualesquiera dos individuos y obtener una medida de la relación que éstos guardan entre sí, así como a qué tiempo pasado pertenece su ancestro común. ²

El uso de modelos matemáticos para explicar la evolución de una población, permite describir la dinámica de reproducción de la misma; incorporar

¹Ver [19].

²Ver [4] y [23].

factores tales como mutación, selección ³ y migración, entre otros; estudiar ciertas características de la población y el cambio que sufren a través del tiempo; y hacer predicciones sobre el comportamiento de la población en el futuro.

De los modelos matemáticos para describir la evolución de una población genética, los más simples y utilizados son el *modelo de Moran* y el *modelo de Wright-Fisher*. En esta tesis se utilizan versiones de ambos modelos que tienen su propia dinámica de reproducción, pero que comparten la dinámica de mutación.

El primer propósito de esta tesis es utilizar los dos modelos anteriores para:

- (1) describir la evolución de una población genética en la que un gene tiene dos tipos de alelos (A_1 y A_2), y se permiten mutaciones entre ambos tipos.
- (2) observar el cambio en la composición de la población respecto a los tipos alélicos, según va evolucionando en el tiempo.

Basándose en una población genética que evoluciona de acuerdo al modelo de Moran, y tomando de ella una muestra en algún momento de su evolución, el segundo propósito de esta tesis es presentar dos procesos que describen relaciones entre ancestros de los genes de dicha muestra. El primero de los procesos es el *proceso n-coalescente de Kingman*, que rastrea el ancestro común de la muestra; y el segundo es el *proceso de líneas de descendencia*, que permite obtener la distribución de tipos alélicos de los genes de la muestra, así como el ancestro común de los que poseen el mismo tipo alélico.

En el **Capítulo 1** se presentan propiedades y resultados de cadenas de Markov en tiempo continuo, y en particular del proceso de nacimiento y muerte, que servirán de fundamento para describir el modelo de Moran y el de Wright-Fisher, así como el proceso n-coalescente de Kingman y el de líneas de descendencia.

En el **Capítulo 2** se introducen los procesos de difusión y sus propiedades, el cálculo de la distribución estacionaria asociada y el comportamiento

³Superioridad de supervivencia de unos individuos respecto a otros.

de estos procesos en la frontera de su espacio de estados. Posteriormente se aborda la convergencia de procesos estocásticos a un proceso de difusión. Este capítulo da la base teórica de la relación existente entre el modelo de Wright-Fisher y el modelo de Moran.

En el **Capítulo 3** se describen y comparan el modelo de Wright-Fisher y el modelo de Moran para una población genética en la que se observa, a medida que evoluciona, el cambio en su composición respecto a los genes tipo A_1 . Para el modelo de Moran se calcula su distribución límite. Para ambos modelos se demuestra la convergencia a una difusión del proceso que registra el cambio en la composición de la población respecto a los genes tipo A_1 , y se calcula la distribución estacionaria de dicha difusión.

En el **Capítulo 4** se describen el proceso n -coalescente de Kingman y el proceso de líneas de descendencia para una población que evoluciona de acuerdo al modelo de Moran (dada la simplicidad de este modelo y la correspondencia que guarda con el de Wright-Fisher). Y se observa que, en ausencia de mutación, el proceso n -coalescente de Kingman es un caso particular del proceso de líneas de descendencia.

En el **Apéndice A** se presentan resultados relacionados con la convergencia de un proceso estocástico a una difusión, y en el **Apéndice B**, resultados requeridos en la Sección 3.2.2.

Capítulo 1

Cadenas de Markov en tiempo continuo

En este capítulo se presentan resultados de cadenas de Markov en tiempo continuo que serán utilizados a lo largo de la presente tesis.

En la Sección 1.1 se muestran resultados básicos de la Teoría de Probabilidad. En la Sección 1.2 se definen las condiciones para la existencia de las probabilidades límite de una cadena de Markov en tiempo continuo y se da la expresión de dichas probabilidades. En la Sección 1.3 se da la expresión de las probabilidades de transición de una cadena de Markov en tiempo continuo y las ecuaciones diferenciales que satisfacen estas probabilidades. En la Sección 1.4 se describe el proceso de nacimiento y muerte; se da la expresión de las probabilidades de transición, las ecuaciones diferenciales que éstas satisfacen y la expresión de las probabilidades límite.

1.1 Resultados preliminares

A continuación se presentan algunos resultados básicos de la Teoría de Probabilidad que pueden encontrarse descritos más ampliamente en Ash(1972), Bartle(1991), Feller(1973), Grimmet y Stirzaker(1992) y Rudin(1976), entre otros.

Definición 1. El conjunto de todos los posibles resultados de un fenómeno aleatorio estudiado, se denomina *espacio muestral* y se denota usualmente por Ω .

Definición 2. Sea \mathcal{F} una colección de subconjuntos de un conjunto Ω . Se dice que \mathcal{F} es una σ -álgebra si y sólo si:

- (a) $\Omega \in \mathcal{F}$
- (b) si $A \in \mathcal{F}$, entonces $A^c \in \mathcal{F}$
- (c) si $A_1, A_2, \dots \in \mathcal{F}$, entonces $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$,

donde A^c denota el complemento del conjunto A :

Definición 3. Sea Ω un conjunto y \mathcal{F} una σ -álgebra de subconjuntos de Ω . Una *medida* sobre una σ -álgebra \mathcal{F} es una función $\mu : \mathcal{F} \rightarrow [0, \infty]$ tal que:

- (a) $\mu(\emptyset) = 0$
- (b) $\mu(A) \geq 0$, para todo $A \in \mathcal{F}$
- (c) si $\{A_i\}_{i \in T}$ es una secuencia de conjuntos disjuntos en \mathcal{F} , con T un conjunto de índices finito o infinito numerable, entonces:

$$\mu \left(\bigcup_{i \in T} A_i \right) = \sum_{i \in T} \mu(A_i).$$

Si $\mu(\Omega) = 1$, μ es llamada *medida de probabilidad* y entonces $\mu : \mathcal{F} \rightarrow [0, 1]$, y usualmente se denota por P .

Definición 4. Un *espacio de medida* es una terna de la forma $(\Omega, \mathcal{F}, \mu)$, donde Ω es un conjunto, \mathcal{F} es una σ -álgebra de subconjuntos de Ω y μ es una medida en \mathcal{F} . Si μ es una medida de probabilidad, entonces $(\Omega, \mathcal{F}, \mu)$ se denomina *espacio de probabilidad*, y usualmente se denota por (Ω, \mathcal{F}, P) .

Definición 5. Sean Ω_1, Ω_2 dos conjuntos, y sea \mathcal{F}_j una σ -álgebra de subconjuntos de Ω_j , $j = 1, 2$. Sea $h : \Omega_1 \rightarrow \Omega_2$. Se dice que h es *medible* respecto a las σ -álgebras \mathcal{F}_1 y \mathcal{F}_2 , si y sólo si $h^{-1}(A) \in \mathcal{F}_1$ para todo $A \in \mathcal{F}_2$. Si $\Omega_2 = \mathbb{R}$ (o \mathbb{C}) y \mathcal{F}_2 es la σ -álgebra generada por una familia de abiertos en \mathbb{R} (o \mathbb{C}), se dice simplemente que h es medible respecto a \mathcal{F}_1 .

Definición 6. Una *variable aleatoria* X en un espacio de probabilidad (Ω, \mathcal{F}, P) es una función de Ω en \mathbb{R} que es medible respecto a \mathcal{F} .

Definición 7. Sea (Ω, \mathcal{F}, P) un espacio de probabilidad. Si $A \subset \Omega$ es tal que $A \in \mathcal{F}$, se dice que A es un *evento*.

Definición 8. Sea (Ω, \mathcal{F}, P) un espacio de probabilidad y sea A un evento. Se define la *probabilidad del evento* A como:

$$P(A) = \sum_{\omega \in A} P(\omega).$$

A continuación se presentan resultados relacionados con las cadenas de Markov en tiempo continuo. Mayores detalles de estos resultados pueden hallarse en Feller(1973), Grimmet y Stirzaker(1992), Karlin y Taylor(1975) y Ross(1983), entre otros.

Definición 9. Un *proceso estocástico* es una familia de variables aleatorias $X = \{X_t : t \in T\}$ indexadas por el conjunto ordenado de índices T , definidas en un mismo espacio de probabilidad (Ω, \mathcal{F}, P) y con valores en un mismo conjunto $S \subset \mathbb{R}$, que se denomina espacio de estados de X .

Definición 10. Un *proceso de Markov* en tiempo continuo con espacio de estados S , es un proceso estocástico $X = \{X_t : t \in T\}$ con valores en S , $T \subset \mathbb{R}$ (se tomará en adelante $T = [0, \infty)$) y con la propiedad de Markov, es decir, dados $t, s \in T$, $s < t$, la distribución ¹ de X_t dado $\{X_u : u \leq s\}$ es igual a la distribución de X_t dado X_s . Formalmente, un proceso estocástico $X = \{X_t : t \in [0, \infty)$ es un proceso de Markov en tiempo continuo si para $x_u \in S$, $u \leq s$, $A \subset S$,

$$P\{X_t \in A \mid X_u = x_u, u \leq s\} = P\{X_t \in A \mid X_s = x_s\}.$$

Cuando el espacio de estados S es finito o infinito numerable, entonces X es llamado *cadena de Markov*. La función $P\{X_t = j \mid X_s = i\}$ es llamada *probabilidad de transición de X* .

Observación 1. Un proceso de Markov es un proceso estocástico tal que la probabilidad de cualquier comportamiento futuro del proceso, dado que se conoce exactamente el estado presente del mismo, no se altera si se conoce su comportamiento pasado.

¹La función de distribución de una variable aleatoria Y que toma valores en un conjunto $S \subset \mathbb{R}$, se define como $F_Y(x) = P\{Y \leq y\}$, $y \in S$.

Definición 11. Un proceso estocástico $X = \{X_t : t \geq 0\}$ con espacio de estados S finito o infinito numerable, es llamado *proceso de semi-Markov* si, dado que X está en el estado $i \in S$:

- (a) el siguiente estado al que accede es el estado $j \in S$ con probabilidad p_{ij} , donde $\sum_{j \in S} p_{ij} = 1$.
- (b) dado que el siguiente estado al que accede es $j \in S$, el tiempo hasta que ocurre la transición de i a j , tiene distribución F_{ij} .

Una definición alternativa de cadena de Markov en tiempo continuo en términos de procesos de semi-Markov es la siguiente.

Definición 12. Una cadena de Markov en tiempo continuo $X = \{X_t : t \geq 0\}$ con espacio de estados S , es un proceso de semi-Markov tal que:

- (a) dado que X está en el estado $i \in S$, el siguiente estado que visita será el estado $j \in S$, $j \neq i$, con probabilidad p_{ij} , donde $\sum_{j \neq i} p_{ij} = 1$.
- (b) F_{ij} es una distribución exponencial con parámetro $\nu_i > 0$, es decir, si τ_i es el tiempo que permanece la cadena en i antes de pasar a j , entonces $P\{\tau_i > t\} = \exp(-\nu_i t)$.

Observación 2. (a) Desde esta óptica, se puede ver a una cadena de Markov en tiempo continuo como un proceso estocástico que pasa de un estado a otro de acuerdo a una cadena de Markov en tiempo discreto, pero tal que el tiempo que permanece en cada estado tiene distribución exponencial.

- (b) ν_i se interpreta como la *tasa a la que el proceso deja el estado i* . De este modo:
 - si $\nu_i = \infty$, se dice que el estado i es instantáneo.
 - si $\nu_i = 0$, se dice que el estado i es absorbente.
- (c) En adelante, salvo que se especifique lo contrario, se tomará $0 \leq \nu_i < \infty$, $i \in S$. Todo estado $i \in S$ que satisface esta condición se denomina *estado estable*.

Definición 13. La cadena de Markov en tiempo continuo $X = \{X_t : t \geq 0\}$ con espacio de estados S , tal que para todo $t, s \in [0, \infty)$ se tiene que:

$$p_{ij}(s, s+t) = P\{X_{t+s} = j \mid X_s = i\},$$

se dice que es *homogénea en el tiempo* si $p_{ij}(s, s+t) = p_{ij}(0, t)$, es decir, si $p_{ij}(s, s+t)$ sólo depende de la longitud del intervalo $[s, s+t]$ y no del tiempo s . En este caso se denota por $p_{ij}(t)$ a $p_{ij}(0, t)$.

Observación 3. (a) Se denotará por $P(t)$ a la matriz cuyos elementos son las probabilidades de transición $p_{ij}(t)$, $i, j \in S$ para toda $t \geq 0$.

(b) En adelante sólo se considerarán cadenas de Markov homogéneas en el tiempo.

Proposición 1. Sea $X = \{X_t : t \geq 0\}$ una cadena de Markov en tiempo continuo con espacio de estados S y probabilidades de transición $p_{ij}(t)$, $i, j \in S$, $t \geq 0$. Entonces, $p_{ij}(t)$ satisface las siguientes propiedades para toda $t \geq 0$ e $i, j \in S$:

$$(a) \quad p_{ij}(0) = \delta_{ij} = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases}$$

$$(b) \quad \sum_{j \in S} p_{ij}(t) = 1.$$

Demostración. (a) Es inmediata pues, dado que la cadena X está en el estado $i \in S$, en un lapso de tiempo igual a cero, sólo puede estar en el estado i (es decir, con probabilidad 1 está en el estado i), y no puede estar en ningún otro estado $j \neq i$ (es decir, con probabilidad 0 estaría en cualquier estado $j \neq i$).

(b) Sea $i \in S$. Supóngase que i es el estado inicial de la cadena X , es decir, $X_0 = i$. Considérense los siguientes eventos:

$$A_j = \{X_t = j \mid X_0 = i\}, \quad j \in S, t \geq 0.$$

Obsérvese que estos eventos son ajenos entre sí y que $\bigcup_{j \in S} A_j = S$. Como el espacio de probabilidad es de medida 1, es decir, $P(S) = 1$, entonces:

$$1 = P(S) = P\left(\bigcup_{j \in S} A_j\right) = \sum_{j \in S} p_{ij}(t), \quad t \geq 0.$$

□

Definición 14. Sea $X = \{X_t : t \geq 0\}$ un proceso de semi-Markov con espacio de estados S . Se denotará por $\{X_n^s : n = 0, 1, \dots\}$ a la sucesión de variables aleatorias donde X_n^s registra el estado visitador por X en el n -ésimo cambio de estado. El proceso estocástico $X^s = \{X_n^s : n = 0, 1, \dots\}$ es una cadena de Markov en tiempo discreto con espacio de estados S y probabilidades de transición p_{ij} , $i, j \in S$ dadas por la Definición 11; es decir, para $i, j, i_0, \dots, i_{n-1} \in S$, $n \geq 0$, se tiene:

$$P\{X_{n+1}^s = j \mid X_0^s = i_0, \dots, X_{n-1}^s = i_{n-1}, X_n^s = i\} = P\{X_{n+1}^s = j \mid X_n^s = i\} \\ = p_{ij}$$

La cadena X^s es llamada *cadena anidada de X* o *cadena de saltos de X* .

Definición 15. Sea $X = \{X_t : t \geq 0\}$ una cadena de Markov en tiempo continuo con espacio de estados S y probabilidades de transición $p_{ij}(t)$, $i, j \in S$, $t \geq 0$. Sea $X^s = \{X_n^s : n = 0, 1, \dots\}$ la cadena de saltos de X . Se dice que X es *irreducible* si la cadena X^s es irreducible, es decir, para toda $i, j \in S$ existe un entero $m \geq 0$ tal que:

$$P\{X_{n+m}^s = j \mid X_n^s = i\} = p_{ij}^{(m)} > 0.$$

Definición 16. Sea $X^s = \{X_n^s : n = 0, 1, \dots\}$ la cadena de saltos de una cadena de Markov en tiempo continuo $X = \{X_t : t \geq 0\}$, con espacio de estados S y probabilidades de transición p_{ij} , $i, j \in S$, $i \neq j$. Decimos que X es *recurrente positiva* si para todo $i \in S$, dado que $X_0^s = i$ y

$$R_{ii} = \min\{n \geq 1, X_n^s = i\}$$

es el tiempo de primer retorno a i , se tiene que:

$$m_i = E[R_{ii} \mid X_0^s = i] < \infty.^2$$

Definición 17. Sea $X = \{X_t : t \geq 0\}$ una cadena de Markov en tiempo continuo con espacio de estados S como se describió en la Definición 12. Se define la *tasa de transición de i a j* , $i, j \in S$, $j \neq i$, denotada por q_{ij} , como:

$$q_{ij} = \nu_i p_{ij}.$$

² $E[\cdot]$ denota la esperanza matemática de una variable aleatoria.

Definición 13. La cadena de Markov en tiempo continuo $X = \{X_t : t \geq 0\}$ con espacio de estados S , tal que para todo $t, s \in [0, \infty)$ se tiene que:

$$p_{ij}(s, s+t) = P\{X_{t+s} = j \mid X_s = i\},$$

se dice que es *homogénea en el tiempo* si $p_{ij}(s, s+t) = p_{ij}(0, t)$, es decir, si $p_{ij}(s, s+t)$ sólo depende de la longitud del intervalo $[s, s+t]$ y no del tiempo s . En este caso se denota por $p_{ij}(t)$ a $p_{ij}(0, t)$.

Observación 3. (a) Se denotará por $P(t)$ a la matriz cuyos elementos son las probabilidades de transición $p_{ij}(t)$, $i, j \in S$ para toda $t \geq 0$.

(b) En adelante sólo se considerarán cadenas de Markov homogéneas en el tiempo.

Proposición 1. Sea $X = \{X_t : t \geq 0\}$ una cadena de Markov en tiempo continuo con espacio de estados S y probabilidades de transición $p_{ij}(t)$, $i, j \in S$, $t \geq 0$. Entonces, $p_{ij}(t)$ satisface las siguientes propiedades para toda $t \geq 0$ e $i, j \in S$:

$$(a) \quad p_{ij}(0) = \delta_{ij} = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases}$$

$$(b) \quad \sum_{j \in S} p_{ij}(t) = 1.$$

Demostración. (a) Es inmediata pues, dado que la cadena X está en el estado $i \in S$, en un lapso de tiempo igual a cero, sólo puede estar en el estado i (es decir, con probabilidad 1 está en el estado i), y no puede estar en ningún otro estado $j \neq i$ (es decir, con probabilidad 0 estaría en cualquier estado $j \neq i$).

(b) Sea $i \in S$. Supóngase que i es el estado inicial de la cadena X , es decir, $X_0 = i$. Considérense los siguientes eventos:

$$A_j = \{X_t = j \mid X_0 = i\}, \quad j \in S, t \geq 0.$$

Obsérvese que estos eventos son ajenos entre sí y que $\bigcup_{j \in S} A_j = S$. Como el espacio de probabilidad es de medida 1, es decir, $P(S) = 1$, entonces:

$$1 = P(S) = P\left(\bigcup_{j \in S} A_j\right) = \sum_{j \in S} p_{ij}(t), \quad t \geq 0.$$

□

Observación 4. Nótese que por definición de tasa de transición de i a j , $i, j \in S$, $i \neq j$, se tiene que:

$$\sum_{j \neq i} q_{ij} = \sum_{j \neq i} \nu_i p_{ij} = \nu_i \sum_{j \neq i} p_{ij} = \nu_i.$$

Proposición 2 (Ecuaciones de Chapman-Kolmogorov). Sea $X = \{X_t : t \geq 0\}$ una cadena de Markov en tiempo continuo con espacio de estados S y probabilidades de transición $p_{ij}(t)$, $i, j \in S$, $t \geq 0$. Entonces, para $t \geq 0$ e $i, j \in S$ se satisface:

$$p_{ij}(t) = \sum_{k \in S} p_{ik}(s) p_{kj}(t-s)$$

para $0 \leq s \leq t$.

Demostración. Sea $t \geq 0$ y sea $0 \leq s \leq t$. De la Definición 13 se puede escribir:

$$p_{ij}(t) = P\{X_t = j \mid X_0 = i\}.$$

Ahora, por la regla de la probabilidad total y la propiedad de Markov de X :

$$P\{X_t = j \mid X_0 = i\} = \sum_{k \in S} P\{X_t = j \mid X_s = k\} P\{X_s = k \mid X_0 = i\}.$$

Entonces:

$$\begin{aligned} p_{ij}(t) &= \sum_{k \in S} P\{X_t = j \mid X_s = k\} P\{X_s = k \mid X_0 = i\} \\ &= \sum_{k \in S} p_{ik}(s) p_{kj}(t-s) \quad (\text{por la Def. 13}) \end{aligned}$$

□

1.2 Probabilidades límite

En esta sección se presentan las condiciones bajo las cuales existe el límite

$$\lim_{t \rightarrow \infty} P\{X_t = j \mid X_0 = i\},$$

y se da la forma explícita del mismo.

Definición 18. Sea $X^s = \{X_n^s : n = 0, 1, \dots\}$ la cadena de saltos de una cadena de Markov en tiempo continuo $X = \{X_t : t \geq 0\}$, con espacio de estados S y probabilidades de transición p_{ij} , $i, j \in S$, $i \neq j$. La *distribución estacionaria* de X^s es una colección de números $\pi^s(i)$, $i \in S$, tal que para toda $i \in S$ las siguientes condiciones son satisfechas:

$$(a) \pi^s(i) \geq 0$$

$$(b) \sum_{i \in S} \pi^s(i) = 1$$

$$(c) \pi^s(i) = \sum_{k \in S} \pi^s(k) p_{ki} = \sum_{k \neq i} \pi^s(k) p_{ki}.$$

Definición 19. Se dice que una variable aleatoria Y es *lattice* si existe un entero $d \geq 0$ tal que

$$\sum_{n=0}^{\infty} P\{Y = nd\} = 1,$$

es decir, Y es *lattice* si toma solamente valores que son múltiplos de d .

Teorema 1. Sea $X = \{X_t : t \geq 0\}$ un proceso de semi-Markov con espacio de estados S . Sea $X^s = \{X_n^s : n = 0, 1, \dots\}$ la cadena de saltos de X y supóngase que X^s es irreducible, recurrente positiva y con distribución estacionaria $\pi^s(i)$, $i \in S$. Considérese a R_{ii} con distribución no *lattice* para toda $i \in S$. Sea μ_i el tiempo medio de permanencia de X en $i \in S$ antes de que salte a un estado $j \in S$, $j \neq i$, es decir, $E[\tau_i] = \mu_i$. Entonces, la distribución límite de X , denotada por $P(i)$, está dada por:

$$P(i) = \lim_{t \rightarrow \infty} P\{X_t = i \mid X_0 = j\} = \frac{\pi^s(i) \mu_i}{\sum_{j \in S} \pi^s(j) \mu_j}, \quad i \in S. \quad (1.2.1)$$

Demostración. Ver [21]. □

Corolario. Sea $X = \{X_t : t \geq 0\}$ una cadena de Markov en tiempo continuo con espacio de estados S y probabilidades de transición $p_{ij}(t)$, $i, j \in S$, $t \geq 0$. Sea $X^s = \{X_n^s : n = 0, 1, \dots\}$ la cadena de saltos de X , irreducible y recurrente positiva con distribución estacionaria $\pi^s(i)$, $i \in S$. Entonces la distribución límite de X está dada por:

$$P(i) = \frac{\pi^s(i) / \nu_i}{\sum_{j \in S} \pi^s(j) / \nu_j}, \quad i \in S. \quad (1.2.2)$$

Demostración. Como una cadena de Markov en tiempo continuo es un proceso de semi-Markov con distribución F_{ij} , $i, j \in S$, igual a la distribución exponencial con parámetro $\nu_i > 0$, entonces $\mu_i = 1/\nu_i$, $i \in S$ y el resultado sigue por el Teorema 1. \square

Observación 5. Por las condiciones (b) y (c) de la Definición 18 y por (1.2.2) se desprende lo siguiente:

$$\nu_j P(j) = \sum_{i \in S} \nu_i P(i) p_{ij}, \quad \sum_{j \in S} P(j) = 1, \quad j \in S.$$

O bien, como $q_{ij} = \nu_i p_{ij}$ (por la Definición 17),

$$\nu_j P(j) = \sum_{i \in S} P(i) q_{ij}, \quad \sum_{j \in S} P(j) = 1, \quad j \in S. \quad (1.2.3)$$

La interpretación de (1.2.3) se da a continuación:

La expresión $\nu_j P(j)$ representa la tasa a la que el proceso deja el estado j , y la expresión $\sum_{i \in S} P(i) q_{ij}$ representa la tasa a la que el proceso entra al estado j . Por tanto, de (1.2.3) se sigue que la tasa de entrada y la de salida del proceso X al estado j son iguales. Por esta razón, a las ecuaciones (1.2.3) se les denomina *ecuaciones de balance total*.

Teorema 2. Sea $X = \{X_t : t \geq 0\}$ una cadena de Markov en tiempo continuo con espacio de estados S y probabilidades de transición $p_{ij}(t)$, $i, j \in S$, $t \geq 0$. Supóngase que X es homogénea e irreducible y que satisface las proposiciones 1 y 2. Entonces:

- (1) si existe una distribución estacionaria Π , ésta es única y $p_{ij}(t) \rightarrow \Pi(t)$ conforme $t \rightarrow \infty$, para toda $i, j \in S$.
- (2) si no existe distribución estacionaria, entonces $p_{ij}(t) \rightarrow 0$ conforme $t \rightarrow \infty$, para toda $i, j \in S$.

Demostración. Ver [21]. \square

1.3 Ecuaciones diferenciales de Kolmogorov

Teorema 3. Sea $X = \{X_t : t \geq 0\}$ una cadena de Markov en tiempo continuo con espacio de estados S y probabilidades de transición $p_{ij}(t)$, $i, j \in S$, $t \geq 0$. Entonces:

(a) para cada $i \in S$ se tiene que $\lim_{h \rightarrow 0^+} \frac{1 - p_{ii}(h)}{h} = \nu_i$ (tasa de salida de i).

(b) para todo $i, j \in S$, $i \neq j$, se tiene que $\lim_{h \rightarrow 0^+} \frac{p_{ij}(h)}{h} = q_{ij}$ (tasa de transición de i a j).

Demostración. Ver, por ejemplo, [15]. □

Observación 6. Si S es finito, entonces ν_i , $i \in S$, no puede tomar el valor ∞ , pues como

$$1 = p_{ii}(h) + \sum_{j \in S, j \neq i} p_{ij}(h),$$

dividiendo por h en ambos lados de la igualdad y tomando el límite conforme $h \rightarrow 0^+$, por el Teorema 3 se tiene que $\nu_i = \sum_{j \in S, j \neq i} q_{ij} < \infty$.

Definición 20. Se dice que una cadena de Markov en tiempo continuo $X = \{X_t : t \geq 0\}$ con espacio de estados S , tasas de transición q_{ij} , $i \neq j$, y tasa de salida ν_i , $i \in S$, es *conservativa* si $\sum_{j \in S, j \neq i} q_{ij} = \nu_i < \infty$ para toda $i \in S$. En adelante, salvo que se especifique lo contrario, se considerarán únicamente cadenas conservativas.

Observación 7. Nótese que por el Teorema 3 se tiene para toda $h \geq 0$, $i, j \in S$:

$$p_{ij}(h) = P\{X_{t+h} = j \mid X_t = i\} \begin{cases} q_{ij}h + o(h) & \text{si } j \neq i \\ 1 - \nu_i h + o(h) & \text{si } j = i, \end{cases}$$

donde $\lim_{h \rightarrow 0^+} \frac{o(h)}{h} = 0$.

Teorema 4 (Ecuaciones retrógradas de Kolmogorov). Sea $X = \{X_t : t \geq 0\}$ una cadena de Markov en tiempo continuo con espacio de estados S , probabilidades de transición $p_{ij}(t)$, $i, j \in S$, $t \geq 0$, tasa de transición q_{ij} , $i \neq j$, y tasa de salida ν_i . Entonces:

$$\frac{d}{dt} p_{ij}(t) = p'_{ij}(t) = \sum_{k \in S, k \neq i} q_{ik} p_{kj}(t) - \nu_i p_{ij}(t). \quad (1.3.1)$$

Demostración. Por definición se tiene que:

$$\frac{d}{dt} p_{ij}(t) = \lim_{h \rightarrow 0} \frac{p_{ij}(t+h) - p_{ij}(t)}{h},$$

y por la Proposición 2 se tiene:

$$p_{ij}(t+h) = \sum_{k \in S} p_{ik}(h) p_{kj}(t).$$

Entonces:

$$\begin{aligned} p_{ij}(t+h) - p_{ij}(t) &= \sum_{k \in S} p_{ik}(h) p_{kj}(t) - p_{ij}(t) \\ &= \sum_{k \in S, k \neq i} p_{ik}(h) p_{kj}(t) - [1 - p_{ii}(h)] p_{ij}(t). \end{aligned}$$

Dividiendo por h , se tiene que:

$$\frac{p_{ij}(t+h) - p_{ij}(t)}{h} = \sum_{k \in S, k \neq i} \frac{p_{ik}(h)}{h} p_{kj}(t) - \frac{1 - p_{ii}(h)}{h} p_{ij}(t), \quad (1.3.2)$$

y por el Teorema 3, para toda $i \in S$ es válido:

$$\lim_{h \rightarrow 0} \frac{1 - p_{ii}(h)}{h} = \nu_i.$$

Para que exista

$$\lim_{h \rightarrow 0} \frac{p_{ij}(t+h) - p_{ij}(t)}{h},$$

es necesario demostrar que existe

$$\lim_{h \rightarrow 0} \sum_{k \in S, k \neq i} \frac{p_{ik}(h)}{h} p_{kj}(t). \quad (1.3.3)$$

y es igual a

$$\sum_{k \in S, k \neq i} q_{ik} p_{kj}(t). \quad (1.3.4)$$

Si S es finito, la existencia de (1.3.3) y la igualdad con (1.3.4) es inmediata, ya que por el Teorema 3,

$$\lim_{h \rightarrow 0} \frac{p_{ik}(h)}{h}$$

existe y es igual a q_{ik} . Y como en (1.3.3) se tiene una suma finita (al ser S finito), se utiliza el hecho de que si $(a_n^{(k)})_{n \geq 0}$, $k = 1, 2, \dots, N$, son sucesiones tales que para cada k existe $\lim_{n \rightarrow \infty} a_n^{(k)}$, entonces $\lim_{n \rightarrow \infty} \sum_{k=1}^N a_n^{(k)}$ también existe.

Se demostrará ahora que para S infinito numerable también se tiene:

$$\lim_{h \rightarrow 0} \sum_{k \in S, k \neq i} \frac{p_{ik}(h)}{h} p_{kj}(t) = \sum_{k \in S, k \neq i} q_{ik} p_{kj}(t).$$

Se usará el hecho de que por definición de límite se tiene:

$$\liminf_{h \rightarrow 0} a_h \leq \lim_{h \rightarrow 0} a_h \leq \limsup_{h \rightarrow 0} a_h,$$

y si

$$\limsup_{h \rightarrow 0} a_h = \liminf_{h \rightarrow 0} a_h,$$

entonces existe

$$\lim_{h \rightarrow 0} a_h \quad \text{y} \quad \liminf_{h \rightarrow 0} a_h = \lim_{h \rightarrow 0} a_h = \limsup_{h \rightarrow 0} a_h.$$

De esta forma, si se demuestra que:

$$(a) \quad \liminf_{h \rightarrow 0} \sum_{k \in S, k \neq i} \frac{p_{ik}(h)}{h} p_{kj}(t) \geq \sum_{k \in S, k \neq i} q_{ik} p_{kj}(t)$$

$$(b) \quad \limsup_{h \rightarrow 0} \sum_{k \in S, k \neq i} \frac{p_{ik}(h)}{h} p_{kj}(t) \leq \sum_{k \in S, k \neq i} q_{ik} p_{kj}(t),$$

el resultado sigue.

Fijese $N < \infty$ arbitraria. Entonces:

$$\begin{aligned} \liminf_{h \rightarrow 0} \sum_{k \in S, k \neq i} \frac{p_{ik}(h)}{h} p_{kj}(t) &\geq \liminf_{h \rightarrow 0} \sum_{k \in S, k \neq i, k < N} \frac{p_{ik}(h)}{h} p_{kj}(t) \\ &= \sum_{k \in S, k \neq i, k < N} p_{kj}(t) \liminf_{h \rightarrow 0} \frac{p_{ik}(h)}{h} \\ &= \sum_{k \neq i, k < N} q_{ik} p_{kj}(t). \end{aligned}$$

Como N es arbitraria, se sigue que:

$$\liminf_{h \rightarrow 0} \sum_{k \in S, k \neq i} \frac{p_{ik}(h)}{h} p_{kj}(t) \geq \sum_{k \in S, k \neq i} q_{ik} p_{kj}(t). \quad (1.3.5)$$

Por otro lado, se tiene para $i < N$:

$$\begin{aligned} \sum_{k \in S, k \neq i} p_{ik}(h) p_{kj}(t) &= \sum_{k \in S, k \neq i, k < N} p_{ik}(h) p_{kj}(t) + \sum_{k \in S, k \geq N} p_{ik}(h) p_{kj}(t) \\ &\stackrel{(1)}{\leq} \sum_{k \in S, k \neq i, k < N} p_{ik}(h) p_{kj}(t) + \sum_{k \in S, k \geq N} p_{ik}(h) \\ &\stackrel{(2)}{=} \sum_{k \in S, k \neq i, k < N} p_{ik}(h) p_{kj}(t) + 1 - \left(p_{ii}(h) + \sum_{k \in S, k \neq i, k < N} p_{ik}(h) \right). \end{aligned}$$

Entonces:

$$\begin{aligned} \limsup_{h \rightarrow 0} \sum_{k \in S, k \neq i} \frac{p_{ik}(h)}{h} p_{kj}(t) &\stackrel{(3)}{\leq} \limsup_{h \rightarrow 0} \sum_{k \in S, k \neq i, k < N} \frac{p_{ik}(h)}{h} p_{kj}(t) + \limsup_{h \rightarrow 0} \frac{1 - p_{ii}(h)}{h} \\ &\quad - \limsup_{h \rightarrow 0} \sum_{k \in S, k \neq i, k < N} p_{ik}(h) \\ &\stackrel{(4)}{=} \sum_{k \in S, k \neq i, k < N} q_{ik} p_{kj}(t) + \nu_i - \sum_{k \in S, k \neq i, k < N} q_{ik}, \end{aligned} \tag{1.3.6}$$

donde:

- (1) porque $p_{kj}(t) \leq 1$ para toda $k, j \in S$ y $t \geq 0$.
- (2) $\sum_{k \in S, k \geq N} p_{ik}(h) = 1 - \sum_{k \in S, k \leq N} p_{ik}(h)$.
- (3) por propiedades de límite.
- (4) pues las sumas son finitas, y $\lim_{h \rightarrow 0} (1 - p_{ii}(h))/h = \nu_i$ y $\lim_{h \rightarrow 0} (p_{ik}(h)/h) = q_{ik}$ existen.

Por la Definición 20 se supone que la cadena es conservativa, i.e. $\sum_{j \neq i} q_{ij} = \nu_i < \infty$. Entonces, haciendo N tender a infinito, de (1.3.6) se obtiene:

$$\limsup_{h \rightarrow 0} \sum_{k \in S, k \neq i} \frac{p_{ik}(h)}{h} p_{kj}(t) \leq \sum_{k \in S, k \neq i} q_{ik} p_{kj}(t) + \nu_i - \nu_i$$

y, por tanto,

$$\limsup_{h \rightarrow 0} \sum_{k \in S, k \neq i} \frac{p_{ik}(h)}{h} p_{kj}(t) \leq \sum_{k \in S, k \neq i} q_{ik} p_{kj}(t). \tag{1.3.7}$$

Por (1.3.5) y (1.3.7) se tiene que:

$$\begin{aligned} \sum_{k \in S, k \neq i} q_{ik} p_{kj}(t) &\leq \liminf_{h \rightarrow 0} \sum_{k \in S, k \neq i} \frac{p_{ik}(h)}{h} p_{kj}(t) \\ &\leq \limsup_{h \rightarrow 0} \sum_{k \in S, k \neq i} \frac{p_{ik}(h)}{h} p_{kj}(t) \leq \sum_{k \in S, k \neq i} q_{ik} p_{kj}(t), \end{aligned}$$

por lo que

$$\lim_{h \rightarrow 0} \sum_{k \in S, k \neq i} \frac{p_{ik}(h)}{h} p_{kj}(t)$$

existe y es igual a

$$\sum_{k \in S, k \neq i} q_{ik} p_{kj}(t).$$

Entonces, tomando el límite conforme $h \rightarrow 0$ en (1.3.2), se tiene que:

$$p'_{ij}(t) = \lim_{h \rightarrow 0} \frac{p_{ij}(t+h) - p_{ij}(t)}{h} = \sum_{k \in S, k \neq i} q_{ik} p_{kj}(t) - \nu_i p_{ij}(t).$$

□

Teorema 5 (Ecuaciones progresivas de Kolmogorov). Sea $X = \{X_t : t \geq 0\}$ una cadena de Markov en tiempo continuo con espacio de estados S y probabilidades de transición $p_{ij}(t)$, $i, j \in S$, $t \geq 0$. Entonces se tiene que:

$$\frac{d}{dt} p_{ij}(t) = p'_{ij}(t) = \sum_{k \in S, k \neq j} p_{ik}(t) q_{kj} - \nu_j p_{ij}(t). \quad (1.3.8)$$

Demostración. Aunque similar a la demostración del Teorema 4, esta demostración involucra otros supuestos para establecer su validez (ver [15] y [21], entre otros) y no se presentará aquí. □

Observación 8. La condición inicial para ambos tipos de ecuaciones es:

$$p_{ij}(0) = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j. \end{cases}$$

1.4 Proceso de nacimiento y muerte

En esta sección se presenta un caso particular de una cadena de Markov en tiempo continuo, el proceso de nacimiento y muerte, y se obtienen algunos resultados relevantes relacionados con este proceso.

1.4.1 Generalidades

Definición 21. Sea $X = \{X_t : t \geq 0\}$ una cadena de Markov en tiempo continuo con espacio de estados $S = \{0, 1, 2, \dots\}$ y para el cual la tasa de transición de i a j , $i \neq j$, $i, j \in S$, es tal que $q_{ij} = 0$ si $|i - j| > 1$. Tal proceso se conoce como *proceso de nacimiento y muerte*. Obsérvese que en este proceso las transiciones desde el estado i sólo son a sus estados vecinos $i - 1, i + 1$.

Observación 9. Salvo que se especifique lo contrario, a lo largo de esta sección se tomará $S = \{0, 1, 2, \dots\}$.

Definición 22. Sean λ_i, μ_i , $i \in S$, dadas por:

$$\begin{aligned} \lambda_i &= q_{i, i+1} \geq 0, & i &= 0, 1, 2, \dots \\ \mu_i &= q_{i, i-1} \geq 0, & i &= 1, 2, \dots \\ \mu_0 &= 0. \end{aligned}$$

Los valores $\{\lambda_i, i = 0, 1, \dots\}$ y $\{\mu_i, i = 1, 2, \dots\}$ se denominan, respectivamente, tasas de nacimiento y tasas de muerte de los procesos respectivos. Si S es finito, por ejemplo $S = \{0, 1, 2, \dots, N\}$, se incluye la condición $\lambda_N = 0$.

Observación 10. Por la Observación 4 se tiene que $\sum_{j \in S, j \neq i} q_{ij} = \nu_i$, $i \in S$, de donde se sigue para el proceso de nacimiento y muerte que:

$$\nu_i = \lambda_i + \mu_i, \quad i = 0, 1, 2, \dots$$

y, por la Definición 17, las probabilidades de transición de la cadena de saltos asociada al proceso de nacimiento y muerte están dadas por:

$$p_{ij} = \begin{cases} \frac{\lambda_i}{\lambda_i + \mu_i} & \text{si } j = i + 1, i = 0, 1, 2, \dots \\ \frac{\mu_i}{\lambda_i + \mu_i} & \text{si } j = i - 1, i = 1, 2, \dots \\ 0 & \text{en otro caso.} \end{cases}$$

Por lo anterior, así como por la Observación 7 y la Proposición 1, se tienen las siguientes probabilidades de transición del proceso de nacimiento y muerte:

$$p_{ij}(h) = \begin{cases} \lambda_i h + o(h) & \text{si } j = i + 1, i = 0, 1, \dots \\ \mu_i h + o(h) & \text{si } j = i - 1, i = 1, 2, \dots \\ 1 - (\lambda_i + \mu_i)h + o(h) & \text{si } j = i, i = 0, 1, \dots \\ 0 & \text{en otro caso} \end{cases}$$

$$\text{y } p_{ij}(0) = \delta_{ij}, i, j \in S.$$

1.4.2 Ecuaciones diferenciales de Kolmogorov

Teorema 6. Sea $X = \{X_t : t \geq 0\}$ un proceso de nacimiento y muerte con espacio de estados $S = \{0, 1, 2, \dots\}$ y probabilidades de transición $p_{ij}(t)$, $i, j \in S$. Las ecuaciones retrógradas y progresivas de Kolmogorov de X están dadas, respectivamente, por:

- Ecuaciones retrógradas

$$\begin{aligned} p'_{0j}(t) &= \lambda_0 p_{1j}(t) - \lambda_0 p_{0j}(t) \\ p'_{ij}(t) &= \lambda_i p_{i+1j}(t) + \mu_i p_{i-1j}(t) - (\lambda_i + \mu_i) p_{ij}(t) \quad i = 1, 2, \dots \end{aligned} \quad (1.4.1)$$

sujetas a la condición inicial $p_{ij}(0) = \delta_{ij}$.

- Ecuaciones progresivas

$$\begin{aligned} p'_{i0}(t) &= \mu_1 p_{i1}(t) - \lambda_0 p_{i0}(t) \\ p'_{ij}(t) &= \lambda_{j-1} p_{i,j-1}(t) + \mu_{j+1} p_{i,j+1}(t) - (\lambda_j + \mu_j) p_{ij}(t) \quad j = 1, 2, \dots \end{aligned} \quad (1.4.2)$$

sujetas a la condición inicial $p_{ij}(0) = \delta_{ij}$.

Demostración. Como un proceso de nacimiento y muerte es un caso particular de las cadenas de Markov en tiempo continuo, se pueden utilizar los teoremas 4 y 5. Entonces:

- (a) Del Teorema 4 se tiene que:

$$p'_{ij}(t) = \sum_{k \in S, k \neq i} q_{ik} p_{kj}(t) - \nu_i p_{ij}(t),$$

donde q_{ik} es la tasa de transición de i a k , $i \neq k$, $k \in S$, y ν_i es la tasa de salida de i , $i \in S$, para una cadena de Markov en tiempo continuo. Para el caso de un proceso de nacimiento y muerte se tiene que:

$$q_{ik} = \begin{cases} q_{i+1} = \lambda_i & \text{si } k = i + 1, i = 0, 1, \dots \\ q_{i-1} = \mu_i & \text{si } k = i - 1, i = 1, 2, \dots \\ 0 & \text{en otro caso,} \end{cases}$$

y para todo $i \in S$,

$$\nu_i = \lambda_i + \mu_i.$$

(i) Tomando $i = 0$ se tiene para $t > 0$ que:

$$\begin{aligned} p'_{0j}(t) &= q_{01}p_{1j}(t) - \nu_0p_{0j}(t) \\ &= \lambda_0p_{1j}(t) - \lambda_0p_{0j}(t). \end{aligned}$$

(ii) Para $i = 1, 2, \dots$ se tiene para $t > 0$ que:

$$\begin{aligned} p'_{ij}(t) &= q_{i+1}p_{i+1j}(t) + q_{i-1}p_{i-1j}(t) - \nu_i p_{ij}(t) \\ &= \lambda_i p_{i+1j}(t) + \mu_i p_{i-1j}(t) - (\lambda_i + \mu_i) p_{ij}(t). \end{aligned}$$

(iii) Para $t = 0$ y para toda $i, j = 0, 1, 2, \dots$ se tiene que:

$$p_{ij}(0) = \delta_{ij} = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j. \end{cases}$$

De (i), (ii) y (iii) se sigue que las ecuaciones retrógradas para un proceso de nacimiento y muerte están dadas por:

$$p'_{0j}(t) = \lambda_0 p_{1j}(t) - \lambda_0 p_{0j}(t)$$

$$p'_{ij}(t) = \lambda_i p_{i+1j}(t) + \mu_i p_{i-1j}(t) - (\lambda_i + \mu_i) p_{ij}(t) \quad i = 1, 2, \dots$$

sujetas a la condición inicial $p_{ij}(0) = \delta_{ij}$.

(b) Del Teorema 5 se tiene que:

$$p'_{ij}(t) = \sum_{k \in S, k \neq j} p_{ik}(t) q_{kj} - \nu_j p_{ij}(t),$$

donde q_{kj} es la tasa de transición de k a j , $k \neq j$, $k \in S$, y ν_j es la tasa de salida de j , $j \in S$, para una cadena de Markov en tiempo continuo. Para el caso de un proceso de nacimiento y muerte se tiene lo siguiente:

(i) Tomando $j = 0$ se tiene para $t > 0$ que:

$$\begin{aligned} p'_{i0}(t) &= p_{i1}(t)q_{10}(t) - \nu_0 p_{i0}(t) \\ &= p_{i1}(t)\mu_1 - \lambda_0 p_{i0}(t) \end{aligned}$$

(ii) Para $j = 1, 2, \dots$ se tiene para $t > 0$ que:

$$\begin{aligned} p'_{ij}(t) &= p_{i,j-1}(t)q_{j-1,j} + p_{i,j+1}(t)q_{j+1,j} - \nu_j p_{ij}(t) \\ &= p_{i,j-1}(t)\lambda_{j-1} + p_{i,j+1}(t)\mu_{j+1} - (\lambda_j + \mu_j)p_{ij}(t) \end{aligned}$$

(iii) Para $t = 0$ y para toda $i, j = 0, 1, 2, \dots$ se tiene $p_{ij}(0) = \delta_{ij}$, igual que el el caso (a)(iii).

De (i), (ii) y (iii) se sigue que las ecuaciones progresivas para un proceso de nacimiento y muerte están dadas por:

$$\begin{aligned} p'_{i0}(t) &= \mu_1 p_{i1}(t) - \lambda_0 p_{i0}(t) \\ p'_{ij}(t) &= \lambda_{j-1} p_{i,j-1}(t) + \mu_{j+1} p_{i,j+1}(t) - (\lambda_j + \mu_j) p_{ij}(t) \quad j = 1, 2, \dots \end{aligned}$$

sujetas a la condición inicial $p_{ij}(0) = \delta_{ij}$.

□

Ejemplo 1. Considérese una cadena de Markov en tiempo continuo $X = \{X_t : t \geq 0\}$ con espacio de estados $S = \{0, 1\}$. Supóngase que X permanece en el estado 0 un tiempo con distribución exponencial de parámetro λ , antes de cambiar al estado 1, donde a su vez permanece un tiempo con distribución exponencial de parámetro μ , antes de retornar al estado 0. Es decir, $\nu_0 = \lambda$, $\nu_1 = \mu$. Como por definición se tiene que $\nu_0 = \mu_0 + \lambda_0$ y $\mu_0 = 0$, entonces $\lambda = \nu_0 = \lambda_0$. También, como por definición $\nu_1 = \mu_1 + \lambda_1$ y $\lambda_1 = 0$, entonces $\mu = \nu_1 = \mu_1$.

De las ecuaciones progresivas de Kolmogorov (1.4.2), se obtiene para X las siguientes ecuaciones:

(a) Para $i = j = 0$:

$$\begin{aligned} p'_{00}(t) &= \mu_1 p_{01}(t) - \lambda_0 p_{00}(t) \\ &= \mu p_{01}(t) - \lambda p_{00}(t). \end{aligned}$$

Como $p_{01}(t) = 1 - p_{00}(t)$, entonces:

$$\begin{aligned} p'_{00}(t) &= \mu(1 - p_{00}(t)) - \lambda p_{00}(t) \\ &= -(\lambda + \mu)p_{00}(t) + \mu, \end{aligned}$$

o bien,

$$p'_{00}(t) + (\lambda + \mu)p_{00}(t) = \mu. \quad (1.4.3)$$

Multiplicando ambos lados de la igualdad en (1.4.3) por el factor integrante $\exp((\lambda + \mu)t)$ se tiene:

$$\exp((\lambda + \mu)t)[p'_{00}(t) + (\lambda + \mu)p_{00}(t)] = \mu \exp((\lambda + \mu)t),$$

o equivalentemente,

$$\frac{d}{dt}[\exp((\lambda + \mu)t)p_{00}(t)] = \mu \exp((\lambda + \mu)t).$$

Integrando esta última ecuación respecto a t en ambos lados de la igualdad, se obtiene:

$$\exp((\lambda + \mu)t)p_{00}(t) = \frac{\mu}{\lambda + \mu} \exp((\lambda + \mu)t) + c \quad (1.4.4)$$

Ahora, como $p_{00}(0) = 1$, sustituyendo en la ecuación (1.4.4) junto con $t = 0$, se tiene que $c = \lambda/(\lambda + \mu)$. Por tanto,

$$p_{00}(t) = \frac{\mu}{\lambda + \mu} + \frac{\lambda}{\lambda + \mu} \exp(-(\lambda + \mu)t).$$

(b) Para $i = j = 1$:

$$\begin{aligned} p'_{11}(t) &= \lambda_0 p_{10}(t) - \mu_1 p_{11}(t) \\ &= \lambda p_{10}(t) - \mu p_{11}(t) \end{aligned}$$

Como $p_{10}(t) = 1 - p_{11}(t)$, entonces:

$$\begin{aligned} p'_{11}(t) &= \lambda[1 - p_{11}(t)] - \mu p_{11}(t) \\ &= -(\lambda + \mu)p_{11}(t) + \lambda, \end{aligned}$$

o bien,

$$p'_{11}(t) + (\lambda + \mu)p_{11}(t) = \lambda.$$

El factor integrante en este caso es el mismo que el caso (a), es decir, $\exp((\lambda + \mu)t)$. Entonces, haciendo un razonamiento análogo al de dicho caso, la solución general de esta ecuación está dada por:

$$\exp((\lambda + \mu)t)p_{11}(t) = \frac{\lambda}{\lambda + \mu} \exp((\lambda + \mu)t) + c,$$

donde $c = \mu/(\lambda + \mu)$, pues $p_{11}(0) = 1$. Por tanto,

$$p_{11}(t) = \frac{\lambda}{\lambda + \mu} + \frac{\mu}{\lambda + \mu} \exp(-(\lambda + \mu)t).$$

Como $p_{01}(t) = 1 - p_{00}(t)$ y $p_{10}(t) = 1 - p_{11}(t)$, entonces:

$$p_{01}(t) = \frac{\lambda}{\lambda + \mu} [1 - \exp(-(\lambda + \mu)t)]$$

$$p_{10}(t) = \frac{\mu}{\lambda + \mu} [1 - \exp(-(\lambda + \mu)t)].$$

Ejemplo 2. Considérese un proceso de muerte pura $X = \{X_t : t \geq 0\}$ con espacio de estados $S = \{0, 1, \dots, N\}$, $N \in \mathbb{N}$, y tasas de muerte $\{\mu_i : i = 0, 1, \dots, N\}$ con $\mu_0 = 0$ (es decir, X es un proceso de nacimiento y muerte con $\lambda_i = 0$ para toda $i = 0, 1, \dots, N$). Entonces, las ecuaciones progresivas de Kolmogorov (1.4.2) para el proceso X se reducen a:

$$p'_{ii}(t) = -\mu_{i+1}p_{i+1}(t) - \mu_i p_{ii}(t)$$

$$= -\mu_i p_{ii}(t) \quad (1.4.5)$$

pues por ser X un proceso de muerte pura, $p_{i+1}(t) = 0$ para toda $t \geq 0$, y:

$$p'_{ij}(t) = \mu_{j+1}p_{j+1}(t) - \mu_j p_{ij}(t), \quad j < i. \quad (1.4.6)$$

La ecuación (1.4.5) se reescribe como:

$$p'_{ii}(t) + \mu_i p_{ii}(t) = 0 \quad (1.4.7)$$

Esta ecuación multiplicada de ambos lados de la igualdad por el factor integrante $\exp(\mu_i t)$, es igual a:

$$\exp(\mu_i t)[p'_{ii}(t) + \mu_i p_{ii}(t)] = 0,$$

o bien, a:

$$\frac{d}{dt}(\exp(\mu_i t) p_{ii}(t)) = 0.$$

Integrando respecto a t ambos lados la ecuación anterior, se obtiene:

$$\exp(\mu_i t) p_{ii}(t) = c, \quad c \text{ constante.}$$

Como $p_{ii}(0) = 1$, entonces $c = 1$. Por tanto,

$$p_{ii}(t) = \exp(-\mu_i t).$$

La solución de la ecuación (1.4.6) es recursiva y se obtiene de la siguiente forma: reescribiendo (1.4.6) como

$$\mu_{j+1} p_{i,j+1}(u) = p'_{ij}(u) + \mu_j p_{ij}(u), \quad j < i,$$

y multiplicando por el factor integrante $\exp(\mu_j u)$ ambos lados de la igualdad anterior, se tiene:

$$\begin{aligned} \exp(\mu_j u) \mu_{j+1} p_{i,j+1}(u) &= \exp(\mu_j u) [p'_{ij}(u) + \mu_j p_{ij}(u)] \\ &= \frac{d}{du} [\exp(\mu_j u) p_{ij}(u)], \quad j < i. \end{aligned}$$

Integrando respecto a u en el intervalo $[0, t]$ ambos lados de la igualdad anterior, se obtiene:

$$\exp(\mu_j t) p_{ij}(t) = \int_0^t \exp(\mu_j u) \mu_{j+1} p_{i,j+1}(u) du + c, \quad j < i.$$

Como $p_{ij}(0) = 0$ (pues $j < i$), entonces $c = 0$. Por tanto,

$$p_{ij}(t) = \mu_{j+1} \exp(-\mu_j t) \int_0^t \exp(\mu_j u) \mu_{j+1} p_{i,j+1}(u) du, \quad j < i.$$

1.4.3 Probabilidades límite

Sea $X = \{X_t : t \geq 0\}$ un proceso de nacimiento y muerte con espacio de estados $S = \{0, 1, 2, \dots\}$. Sea $X^s = \{X_n^s : n = 0, 1, \dots\}$ la cadena de saltos asociada a X , con probabilidades de transición p_{ij} , $i, j \in S$. X^s corresponde

a una caminata aleatoria en tiempo discreto en los enteros mayores o iguales a cero, con una barrera reflejante en el estado 0. Esta caminata aleatoria es una cadena de Markov irreducible y recurrente positiva (ver [14]). Entonces, por el Corolario del Teorema 1, para el proceso X existen probabilidades límite:

$$\lim_{t \rightarrow \infty} p_{ij}(t) = P(j) > 0, \quad \text{para toda } i, j \in S,$$

están dadas por (1.2.2) y satisfacen las ecuaciones de balance total:

$$\nu_j P(j) = \sum_{i \in S} P(i) q_{ij}, \quad \sum_{j \in S} P(j) = 1, \quad j \in S \quad (1.4.8)$$

Teorema 7. Sea $X = \{X_t : t \geq 0\}$ un proceso de nacimiento y muerte con espacio de estados $S = \{0, 1, \dots\}$ y tasas de nacimiento y muerte $\{\lambda_i, i = 0, 1, \dots\}$ y $\{\mu_i, i = 1, 2, \dots\}$, respectivamente. Entonces, si

$$\sum_{m=1}^{\infty} \frac{\lambda_{m-1} \lambda_{m-2} \cdots \lambda_1 \lambda_0}{\mu_m \mu_{m-1} \cdots \mu_2 \mu_1} < \infty$$

las probabilidades límite de X están dadas por:

$$P(0) = \left[1 + \sum_{m=1}^{\infty} \frac{\lambda_{m-1} \lambda_{m-2} \cdots \lambda_1 \lambda_0}{\mu_m \mu_{m-1} \cdots \mu_2 \mu_1} \right]^{-1}$$

$$P(m) = \frac{\lambda_{m-1} \lambda_{m-2} \cdots \lambda_1 \lambda_0}{\mu_m \mu_{m-1} \cdots \mu_2 \mu_1 \left(1 + \sum_{m=1}^{\infty} \frac{\lambda_{m-1} \lambda_{m-2} \cdots \lambda_1 \lambda_0}{\mu_m \mu_{m-1} \cdots \mu_2 \mu_1} \right)}, \quad \text{para } m = 1, 2, \dots$$

Demostración. Sea $P(i)$ la probabilidad límite de la cadena X , para cada $i \in S$. Entonces, $P(i)$, $i \in S$, debe satisfacer (1.4.8), con:

$$q_{ij} = \begin{cases} \lambda_i & \text{si } j = i + 1, i = 0, 1, \dots \\ \mu_i & \text{si } j = i - 1, i = 1, 2, \dots \\ 0 & \text{en otro caso,} \end{cases}$$

$$\nu_i = \lambda_i + \mu_i, \quad i = 0, 1, \dots,$$

es decir, se satisfacen las ecuaciones:

$$\lambda_0 P(0) = \mu_1 P(1) \quad (1.4.9)$$

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

$$(\mu_m + \lambda_m)P(m) = \mu_{m+1}P(m+1) + \lambda_{m-1}P(m-1). \quad (1.4.10)$$

Se demostrará por inducción que para $m = 1, 2, 3, \dots$ se tiene:

$$P(m) = \frac{\lambda_{m-1}\lambda_{m-2}\cdots\lambda_1\lambda_0}{\mu_m\mu_{m-1}\cdots\mu_2\mu_1}P(0).$$

De la ecuación (1.4.9) se tiene que para $k = 1$ es válido:

$$P(1) = \frac{\lambda_0}{\mu_1}P(0). \quad (1.4.11)$$

Para $k = 2$, utilizando (1.4.10) con $m = 1$, es válido:

$$(\mu_1 + \lambda_1)P(1) = \mu_2P(2) + \lambda_0P(0),$$

es decir,

$$\mu_2P(2) = (\lambda_1 + \mu_1)P(1) - \lambda_0P(0),$$

por lo que:

$$\begin{aligned} P(2) &= \frac{\lambda_1 + \mu_1}{\mu_2}P(1) - \frac{\lambda_0}{\mu_2}P(0) \\ &= \frac{\lambda_1 + \mu_1}{\mu_2} \frac{\lambda_0}{\mu_1}P(0) - \frac{\lambda_0}{\mu_2}P(0) \quad (\text{por (1.4.11)}) \\ &= \frac{(\lambda_1 + \mu_1)\lambda_0 - \lambda_0\mu_1}{\mu_2\mu_1}P(0) \\ &= \frac{\lambda_1\lambda_0 + \lambda_0\mu_1 - \lambda_0\mu_1}{\mu_2\mu_1}P(0) = \frac{\lambda_1\lambda_0}{\mu_2\mu_1}P(0). \end{aligned}$$

Supóngase que el resultado vale para toda $k \leq l$, se demostrará ahora que es válido para $l + 1$. Por (1.4.10), para $m = l$ se tiene:

$$(\mu_l + \lambda_l)P(l) = \mu_{l+1}P(l+1) + \lambda_{l-1}P(l-1).$$

Entonces,

$$\begin{aligned}
 P(l+1) &= \frac{\lambda_l + \mu_l}{\mu_{l+1}} P(l) - \frac{\lambda_{l-1}}{\mu_{l+1}} P(l-1) \\
 &= \frac{\lambda_l + \mu_l}{\mu_{l+1}} \frac{\lambda_{l-1} \lambda_{l-2} \cdots \lambda_1 \lambda_0}{\mu_l \mu_{l-1} \cdots \mu_1} P(0) - \frac{\lambda_{l-1}}{\mu_{l+1}} \frac{\lambda_{l-2} \lambda_{l-3} \cdots \lambda_1 \lambda_0}{\mu_{l-1} \mu_{l-2} \cdots \mu_1} P(0) \\
 &= \frac{\lambda_l \lambda_{l-1} \cdots \lambda_1 \lambda_0 + \mu_l \lambda_{l-1} \cdots \lambda_1 \lambda_0 - \mu_l \lambda_{l-1} \lambda_{l-2} \cdots \lambda_1 \lambda_0}{\mu_{l+1} \mu_l \cdots \mu_1} P(0) \\
 &= \frac{\lambda_l \lambda_{l-1} \cdots \lambda_1 \lambda_0}{\mu_{l+1} \mu_l \cdots \mu_1} P(0),
 \end{aligned}$$

donde la segunda igualdad se da por la hipótesis de inducción.

Dado que $\sum_{m=0}^{\infty} P(m) = 1$, se tiene que:

$$1 = P(0) + P(0) \sum_{m=1}^{\infty} \frac{\lambda_{m-1} \lambda_{m-2} \cdots \lambda_1 \lambda_0}{\mu_m \mu_{m-1} \cdots \mu_2 \mu_1}$$

y, por tanto,

$$P(0) = \left[1 + \sum_{m=1}^{\infty} \frac{\lambda_{m-1} \lambda_{m-2} \cdots \lambda_1 \lambda_0}{\mu_m \mu_{m-1} \cdots \mu_2 \mu_1} \right]^{-1}$$

Entonces:

$$P(m) = \frac{\lambda_{m-1} \lambda_{m-2} \cdots \lambda_1 \lambda_0}{\mu_m \mu_{m-1} \cdots \mu_2 \mu_1 \left(1 + \sum_{m=1}^{\infty} \frac{\lambda_{m-1} \lambda_{m-2} \cdots \lambda_1 \lambda_0}{\mu_m \mu_{m-1} \cdots \mu_2 \mu_1} \right)}, \quad m = 1, 2, \dots$$

Obsérvese que si $\sum_{m=1}^{\infty} \frac{\lambda_{m-1} \lambda_{m-2} \cdots \lambda_1 \lambda_0}{\mu_m \mu_{m-1} \cdots \mu_2 \mu_1} = \infty$, entonces necesariamente $P(0) = 0$ y $P(j) = 0$ para toda $j \in S$. De esta forma, para que $P(j) > 0$ para toda $j \in S$, se debe tener:

$$\sum_{m=1}^{\infty} \frac{\lambda_{m-1} \lambda_{m-2} \cdots \lambda_1 \lambda_0}{\mu_m \mu_{m-1} \cdots \mu_2 \mu_1} < \infty$$

y el resultado sigue. □

Capítulo 2

Proceso de difusión

En este capítulo se presentan resultados de procesos de difusión que se utilizarán posteriormente en el Capítulo 3.

En la Sección 2.1 se define un proceso de difusión y se presentan definiciones relacionadas con él. En la Sección 2.2 se presentan las ecuaciones diferenciales que satisfacen ciertas funciones de los procesos de difusión. En la Sección 2.3 se define la distribución estacionaria de un proceso de difusión y se obtiene su forma general. En la Sección 2.4 se da una clasificación de las fronteras del espacio de estados de un proceso de difusión y se calcula la distribución estacionaria del proceso para el caso de fronteras de entrada. En la Sección 2.5 se describe la convergencia de procesos estocásticos a un proceso de difusión.

2.1 Resultados preliminares

A continuación se define un proceso de difusión y se presentan definiciones relacionadas, que pueden encontrarse descritas más ampliamente en Ash(1972) y Karlin y Taylor(1981).

Definición 23. Sea (Ω, \mathcal{F}, P) un espacio de probabilidad. Una familia de σ -álgebras $\{\mathcal{F}_t : t \geq 0\}$ es llamada *filtración* en (Ω, \mathcal{F}) si:

- (i) para cada $t \geq 0$, $\mathcal{F}_t \subset \mathcal{F}$
- (ii) para $t_1, t_2 \geq 0$, si $t_1 \leq t_2$ entonces $\mathcal{F}_{t_1} \subset \mathcal{F}_{t_2}$.

Si $\{\mathcal{F}_t : t \geq 0\}$ es una filtración se define:

$$\mathcal{F}_\infty = \sigma \left(\bigcup_{t \geq 0} \mathcal{F}_t \right),$$

es decir, la σ -álgebra generada por $\bigcup_{t \geq 0} \mathcal{F}_t$. El espacio $(\Omega, \mathcal{F}, P, \{\mathcal{F}_t : t \geq 0\})$ es llamado *espacio de probabilidad filtrado*.

Definición 24. Sea $X = \{X_t : t \geq 0\}$ un proceso estocástico. La *filtración generada por X* , denotada por $\{\mathcal{F}_t^X : t \geq 0\}$ es definida por:

$$\mathcal{F}_t^X = \sigma(X_s : s \leq t).$$

Definición 25. Sea $X = \{X_t : t \geq 0\}$ un proceso estocástico definido en un espacio de probabilidad (Ω, \mathcal{F}, P) . La variable aleatoria τ definida en (Ω, \mathcal{F}, P) que toma valores en $\{0, 1, \dots\}$, es llamada *tiempo de Markov o tiempo de paro* con respecto a X si para cada $t \geq 0$, $\{\tau \leq t\} \subset \mathcal{F}_t^X$, es decir, el evento $\{\tau \leq t\}$ está completamente determinado por $\{X_s : s \leq t\}$.

Definición 26. Sea $X = \{X_t : t \geq 0\}$ un proceso de Markov con espacio de estados S y probabilidades de transición $p_{ij}(t)$, $i, j \in S$, $t \geq 0$. El proceso X es un *proceso de Markov fuerte* si para cada tiempo de paro τ con respecto a X , se tiene que para cada $t \geq 0$ y $A \subset S$, es válido:

$$P\{X_{\tau+t} \in A \mid X_s = x_s, s \leq \tau\} = P\{X_{\tau+t} \in A \mid X_\tau = x_\tau\}$$

para todo $x_s \in S$.

Definición 27 (Proceso de difusión). Un proceso de Markov fuerte $X = \{X_t : t \geq 0\}$ con espacio de estados $S = I \subseteq \mathbb{R}$ (donde I es un intervalo de la forma (l, r) , $(l, r]$, $[l, r)$ o $[l, r]$, $l < r$, y los valores $l = -\infty$ y/o $r = \infty$ son permitidos) es un *proceso de difusión* si:

1. Para cada $\omega \in \Omega$, $X_t(\omega)$ como función de t es continua por la derecha y tiene límite por la izquierda.
2. Para cada $\omega \in \Omega$, $X_t(\omega)$ como función de t es continua por la izquierda a través de tiempos de paro, es decir, si $\tau_1 \leq \tau_2 \leq \dots$ son tiempos de paro y $\lim_{n \rightarrow \infty} \tau_n = \tau \leq \infty$, entonces $\lim_{n \rightarrow \infty} X_{\tau_n}(\omega) = X_\tau(\omega)$ siempre que $\tau < \infty$. En otras palabras, la igualdad anterior se da casi seguramente en el sentido de que $P\{\tau < \infty \text{ y } \lim_{n \rightarrow \infty} X_{\tau_n}(\omega) \neq X_\tau(\omega)\} = 0$.

3. X_t satisface la *Condición de Dinkin*, es decir, dada $\varepsilon > 0$

$$\lim_{h \rightarrow 0} P \{ |X_{t+h} - X_t| > \varepsilon \mid X_t = x \} = 0 \quad (2.1.1)$$

uniformemente respecto a t para toda x en un subintervalo compacto de S y t en un intervalo $[0, N]$, $N < \infty$.

Definición 28. Sea $X = \{X_t : t \geq 0\}$ un proceso de difusión con espacio de estados $S = I$ un intervalo contenido en \mathbb{R} . Para $z \in I$, se define el tiempo T_z como el tiempo que tarda X en tomar el valor z por primera vez, es decir,

$$T_z = \begin{cases} \infty, & \text{si } X_t \neq z \text{ para todo } t \geq 0 \\ \inf\{t \geq 0 : X_t = z\}, & \text{en otro caso.} \end{cases}$$

Definición 29. Se dice que un proceso de difusión $X = \{X_t : t \geq 0\}$ con espacio de estados $S = (l, r)$, $l < r$, es *regular* si

$$P\{T_z < \infty \mid X_0 = x\} > 0$$

siempre que $l < x, z < r$. Es decir, partiendo de algún $x \in S$, cualquier otro punto $z \in S$ puede alcanzarse con probabilidad positiva.

Observación 11. En adelante sólo se considerarán procesos de difusión regulares.

Definición 30. Sea $X = \{X_t : t \geq 0\}$ un proceso estocástico. Se define el incremento de X en el intervalo $[t, t+h]$ como:

$$\Delta_h X_t = X_{t+h} - X_t.$$

Lema 1. Sea $X = \{X_t : t \geq 0\}$ un proceso estocástico con espacio de estados un intervalo S , que satisface las condiciones 1 y 2 de la Definición 27. Si X satisface la condición de momento infinitesimal:

$$\lim_{h \rightarrow 0} \frac{1}{h} E [|\Delta_h X_t|^p \mid X_t = x] = 0, \quad \text{para alguna } p > 2$$

uniformemente respecto a t para x en cualquier subintervalo compacto de S y t en un intervalo $[0, N]$, $N < \infty$, entonces el proceso satisface la condición de *Dinkin*.

Demostración. Ver [15]. □

Definición 31. Sea $X = \{X_t : t \geq 0\}$ un proceso de difusión con espacio de estados $S = I$. Para $x \in S$ y $t \geq 0$, los parámetros infinitesimales de X están definidos por:

$$\mu(x, t) = \lim_{h \rightarrow 0} \frac{1}{h} E[\Delta_h X_t | X_t = x] \quad (2.1.2)$$

$$\sigma^2(x, t) = \lim_{h \rightarrow 0} \frac{1}{h} E[\{\Delta_h X_t\}^2 | X_t = x] \quad (2.1.3)$$

Al parámetro $\mu(x, t)$ se le llama *media infinitesimal, desplazamiento infinitesimal esperado o parámetro de deriva*. Al parámetro $\sigma^2(x, t)$ se le llama *varianza infinitesimal o parámetro de difusión*.

Observación 12. 1. El proceso X también satisface:

$$\lim_{h \rightarrow 0} \frac{1}{h} E[|\Delta_h X_t|^r | X_t = x] = 0, \quad r = 3, 4, \dots \quad (2.1.4)$$

(ver [15]).

2. La motivación para denominar a $\mu(x, t)$ media infinitesimal proviene de que, por definición, se tiene:

$$E[\Delta_h X_t | X_t = x] = \mu(x, t)h + o_1(h),$$

donde $\lim_{h \rightarrow 0} \frac{o_1(h)}{h} = 0$.

Para la varianza infinitesimal $\sigma^2(x, t)$ se tiene que:

$$\begin{aligned} \text{Var}[\Delta_h X_t | X_t = x] &= E[(\Delta_h X_t)^2 | X_t = x] - (E[\Delta_h X_t | X_t = x])^2 \\ &= \sigma^2(x, t)h + o_2(h) - [\mu(x, t)h + o_1(h)]^2 \\ &= \sigma^2(x, t)h + o_3(h), \end{aligned}$$

donde $o_3(h) = o_2(h) - [\mu(x, t)h + o_1(h)]^2$ y es tal que $\lim_{h \rightarrow 0} \frac{o_3(h)}{h} = 0$.

3. Para un proceso de difusión regular típico se tiene que $\sigma^2(x, t) > 0$ para toda x tal que $l < x < r$ y $t > 0$.
4. Cuando el proceso es homogéneo en el tiempo, entonces se escribe $\mu(x, t) = \mu(x)$ y $\sigma^2(x, t) = \sigma^2(x)$. En adelante sólo se considerarán procesos de difusión homogéneos en el tiempo.

Teorema 8. Sea $X = \{X_t : t \geq 0\}$ un proceso de difusión regular y homogéneo con espacio de estados $S = (l, r)$, $l < r$, y parámetros infinitesimales $\mu(x)$ y $\sigma^2(x)$, $x \in S$. Sea g una función estrictamente monótona en S con segunda derivada $g''(x)$ continua para $l < x < r$. Entonces $Y_t = g(X_t)$ define un proceso de difusión regular en el intervalo con extremos $g(l)$ y $g(r)$, y $Y = \{Y_t : t \geq 0\}$ tiene parámetros infinitesimales:

$$\begin{aligned}\mu_Y(y) &= \frac{1}{2}\sigma^2(x)g''(x) + \mu(x)g'(x) \\ \sigma_Y^2(y) &= \sigma^2(x)[g'(x)]^2\end{aligned}$$

donde $y = g(x)$.

Demostración. Ver [15]. □

2.2 Ecuaciones diferenciales

En esta sección se presentan ecuaciones diferenciales que son satisfechas por ciertas funciones de los procesos de difusión. Las demostraciones serán omitidas, pero pueden consultarse, por ejemplo, en Karlin y Taylor(1981) y Tudor(1994).

Definición 32. Sea $X = \{X_t : t \geq 0\}$ un proceso de difusión regular y homogéneo con espacio de estados $S = (l, r)$, $l < r$, y parámetros infinitesimales $\mu(x)$ y $\sigma^2(x)$, $x \in S$. Se define a $P_{xy}(t)$, la función de distribución de transición de X_t , como:

$$P_{xy}(t) = P\{X_t \leq y \mid X_0 = x\},$$

sujeta a la distribución inicial:

$$P_{xy}(0) = \begin{cases} 1 & \text{si } x \leq y \\ 0 & \text{si } x > y, \end{cases}$$

(es decir, una distribución puntual concentrada en x).

Definición 33. Sea $X = \{X_t : t \geq 0\}$ un proceso de difusión regular y homogéneo con espacio de estados $S = (l, r)$, $l < r$, y parámetros infinitesimales $\mu(x)$ y $\sigma^2(x)$, $x \in S$. Se define a $p_{xy}(t)$, la densidad de transición de X_t correspondiente a la distribución $P_{xy}(t)$, como:

$$p_{xy}(t) = \frac{d}{dt} P_{xy}(t), \quad t \geq 0.$$

Teorema 9 (Ecuación retrógrada de Kolmogorov). Sea $X = \{X_t : t \geq 0\}$ un proceso de difusión regular y homogéneo con espacio de estados $S = (l, r)$, $l < r$, y parámetros infinitesimales $\mu(x)$ y $\sigma^2(x)$, $x \in S$. Se define para $t \geq 0$:

$$u(x, t) = E[g(X_t) | X_0 = x],$$

donde g es una función acotada y continua por partes sobre S . Si $u(x, t)$ es continua y diferenciable respecto a t , dos veces diferenciable respecto a x y continua por la derecha en cero, entonces $u(x, t)$ satisface la ecuación diferencial parcial

$$\frac{\partial u(x, t)}{\partial t} = \frac{1}{2} \sigma^2(x) \frac{\partial^2 u(x, t)}{\partial x^2} + \mu(x) \frac{\partial u(x, t)}{\partial x} \quad (2.2.1)$$

con la condición inicial $u(0+, x) = g(x)$, donde $u(0+, x) = \lim_{h \rightarrow 0+} u(h, x)$. En particular, si g es tal que

$$g(\eta) = \begin{cases} 1 & \text{si } \eta \leq y \\ 0 & \text{si } \eta > y, \end{cases}$$

entonces $u(x, t) = P_{xy}(t)$ (ver Definición 32). En este caso, la ecuación (2.2.1) se transforma en la siguiente ecuación:

$$\frac{\partial}{\partial t} P_{xy}(t) = \frac{1}{2} \sigma^2(x) \frac{\partial^2}{\partial x^2} P_{xy}(t) + \mu(x) \frac{\partial}{\partial x} P_{xy}(t), \quad t > 0, x, y \in S \quad (2.2.2)$$

y se denomina ecuación retrógrada de Kolmogorov, y está sujeta a la condición inicial

$$P_{xy}(0+) = \begin{cases} 1 & \text{si } x \leq y \\ 0 & \text{si } x > y. \end{cases}$$

La densidad de transición $p_{xy}(t)$ también satisface la ecuación retrógrada de Kolmogorov:

$$\frac{\partial}{\partial t} p_{xy}(t) = \frac{1}{2} \sigma^2(x) \frac{\partial^2}{\partial x^2} p_{xy}(t) + \mu(x) \frac{\partial}{\partial x} p_{xy}(t) \quad t > 0, x, y \in S. \quad (2.2.3)$$

Demostración. Ver [15]. □

Observación 13. No siempre las ecuaciones (2.2.2) y (2.2.3) admiten soluciones únicas aunque se requiera que $P_{xy}(t)$ sea una función de distribución en $y \in S$ para cada $t \in S$ y $x \in S$. Esto es debido a que (2.2.2) no considera el comportamiento del proceso en la frontera del espacio de estados S (ver [15]).

Existe una segunda ecuación diferencial parcial satisfecha por la densidad de transición $p_{xy}(t) = \frac{d}{dt} P_{xy}(t)$. Dicha ecuación puede verse como la *dual* de la ecuación (2.2.3) y las variables a considerar son $t \geq 0$ e $y \in S$ (es decir, se toma el estado al tiempo t en vez del estado inicial $x \in S$).

Teorema 10 (Ecuación progresiva de Kolmogorov o ecuación de evolución). Sea $X = \{X_t : t \geq 0\}$ un proceso de difusión regular y homogéneo con espacio de estados $S = (l, r)$, $l < r$, y parámetros infinitesimales $\mu(x)$ y $\sigma^2(x)$, $x \in S$. Y sea $p_{xy}(t)$, $x, y \in S$, $t \geq 0$, como en la Definición 33. Entonces, $p_{xy}(t)$ satisface la ecuación diferencial parcial:

$$\frac{\partial}{\partial t} p_{xy}(t) = \frac{1}{2} \frac{\partial^2}{\partial y^2} [\sigma^2(y) p_{xy}(t)] - \frac{\partial}{\partial y} [\mu(y) p_{xy}(t)], \quad t > 0, x, y, \in S.$$

Demostración. Ver [15]. □

2.3 Distribución estacionaria

A continuación se definirá la distribución estacionaria de un proceso de difusión y se obtendrá su forma general. Para mayores detalles, ver Karlin y Taylor (1981).

Definición 34. Sea $X = \{X_t : t \geq 0\}$ un proceso de difusión regular y homogéneo con espacio de estados $S = (l, r)$, $l < r$, y parámetros infinitesimales $\mu(x)$ y $\sigma^2(x)$, $x \in S$. Se dice que $\psi(\cdot)$ es una *distribución estacionaria*

de X si satisface:

$$\begin{aligned} \psi(y) &= \int_S \psi(x) p_{xy}(t) dx, & t > 0, y \in S \\ \psi(x) &\geq 0, & x \in S \\ \int_S \psi(\xi) d\xi &= 1. \end{aligned}$$

Proposición 3. Sea $X = \{X_t : t \geq 0\}$ un proceso de difusión regular y homogéneo con espacio de estados $S = (l, r)$, $l < r$, y parámetros infinitesimales $\mu(x)$ y $\sigma^2(x)$, $x \in S$. Si existe una distribución estacionaria $\psi(y)$ de X , entonces ésta satisface la ecuación diferencial parcial:

$$0 = \frac{1}{2} \frac{\partial^2}{\partial y^2} [\sigma^2(y)\psi(y)] - \frac{\partial}{\partial y} [\mu(y)\psi(y)], \quad y \in S. \quad (2.3.1)$$

Demostración. Un bosquejo de la demostración se encuentra en [15]. \square

2.3.1 Cálculo de la distribución estacionaria

A continuación se describirá cómo obtener la distribución estacionaria de un proceso de difusión.

Sea $X = \{X_t : t \geq 0\}$ un proceso de difusión regular y homogéneo con espacio de estados $S = (l, r)$, $l < r$, y parámetros infinitesimales $\mu(x)$ y $\sigma^2(x)$, $x \in S$. Sea $\psi(y)$, $y \in S$, su distribución estacionaria. Entonces, por la Proposición 3, $\psi(y)$, $y \in S$, debe satisfacer (2.3.1). De esta forma, integrando dicha ecuación respecto a y se tiene que:

$$\frac{\partial}{\partial y} \left(\frac{\sigma^2(y)}{2} \psi(y) \right) - \mu(y)\psi(y) = \frac{1}{2} c_1 \quad (2.3.2)$$

donde c_1 es una constante. Para resolver esta ecuación diferencial ordinaria, se multiplica ambos lados de (2.3.2) por el factor integrante

$$s(y) = \exp \left(- \int^y \frac{2\mu(\xi)}{\sigma^2(\xi)} d\xi \right) \quad (2.3.3)$$

con lo que (2.3.2) queda de la forma:

$$\begin{aligned} \exp\left(-\int^y \frac{2\mu(\xi)}{\sigma^2(\xi)} d\xi\right) \frac{d}{dy}[\sigma^2(y)\psi(y)] - 2\exp\left(-\int^y \frac{2\mu(\xi)}{\sigma^2(\xi)} d\xi\right) \mu(y)\psi(y) = \\ = c_1 \exp\left(-\int^y \frac{2\mu(\xi)}{\sigma^2(\xi)} d\xi\right) \end{aligned}$$

o bien,

$$\frac{d}{dy} \left[\exp\left(-\int^y \frac{2\mu(\xi)}{\sigma^2(\xi)} d\xi\right) \sigma^2(y)\psi(y) \right] = c_1 \exp\left(-\int^y \frac{2\mu(\xi)}{\sigma^2(\xi)} d\xi\right).$$

Por (2.3.3), esta ecuación se reescribe como:

$$\frac{d}{dy} [s(y)\sigma^2(y)\psi(y)] = c_1 s(y). \quad (2.3.4)$$

Integrando nuevamente con respecto a y , y haciendo

$$S(x) = \int^x s(y) dy,$$

para cada lado de la igualdad (2.3.4) se obtiene:

$$\begin{aligned} \int^x \frac{d}{dy} [s(y)\sigma^2(y)\psi(y)] dy &= s(x)\sigma^2(x)\psi(x) - c_2 \\ \int^x c_1 s(y) dy &= c_1 S(x), \end{aligned}$$

por lo que

$$s(x)\sigma^2(x)\psi(x) = c_1 S(x) + c_2,$$

de donde se sigue:

$$\psi(x) = c_1 \frac{S(x)}{s(x)\sigma^2(x)} + c_2 \frac{1}{s(x)\sigma^2(x)} = m(x)[c_1 S(x) + c_2],$$

con

$$m(x) = \frac{1}{s(x)\sigma^2(x)}.$$

Las constantes c_1 y c_2 se determinan de modo que se garantice que $\psi(x) \geq 0$ en $S = (l, r)$ y $\int_l^r \psi(\xi) d\xi = 1$. Si esto es posible, entonces existe una distribución estacionaria, de otro modo no se puede afirmar su existencia.

2.4 Comportamiento en la frontera del espacio de estados

A continuación se definirán sobre las fronteras del espacio de estados de un proceso de difusión, ciertas funciones que permitirán clasificar dichas fronteras. Posteriormente, para el caso de fronteras de entrada, se obtendrá una expresión única para la distribución estacionaria de un proceso de difusión. Para mayores detalles ver Karlin y Taylor (1981).

Definición 35. Sea $X = \{X_t : t \geq 0\}$ un proceso de difusión regular y homogéneo con espacio de estados $S = (l, r)$, $l < r$ y con parámetros infinitesimales $\mu(x)$ y $\sigma^2(x)$, $x \in S$.

(a) Se denomina *función escala* a la función:

$$S(x) = \int_{x_0}^x s(\xi) d\xi, \quad x \in S,$$

donde

$$s(\xi) = \exp \left(- \int_{\xi_0}^{\xi} \frac{2\mu(\eta)}{\sigma^2(\eta)} d\eta \right)$$

y $x_0, \xi_0 \in S$ son fijos y arbitrarios.

(b) Se define la *medida escala* de un intervalo $J = [c, d] \subset (l, r)$ como:

$$\tilde{S}[J] = \tilde{S}[c, d] = S(d) - S(c),$$

donde $S(x)$ es la función escala.

Propiedades 1

La medida escala satisface las siguientes propiedades:

- (i) $0 < \tilde{S}[c, d] < \infty$ para $l < c < d < r$, pues como $s(\xi) > 0$ para toda $\xi \in (l, r)$ y $[c, d] \subset (l, r)$, entonces $0 < \int_c^d s(\xi) d\xi < \infty$.
- (ii) $\tilde{S}[c, d] = \tilde{S}[c, x] + \tilde{S}[x, d]$ para x tal que $l < c < x < d < r$, por lo que para d fija, \tilde{S} es monótona en c . Esto es porque como $s(\xi)$ es una función monótona decreciente en ξ y $\tilde{S}[c, d] = \int_c^d s(\xi) d\xi$, entonces para d fija y $c_2 < c_1$ se tiene que $\tilde{S}[c_2, d] > \tilde{S}[c_1, d]$.

Observación 14. Sea $x \in [c, d] \subset (l, r)$, $l < r$. Se utilizará la notación $dS(x) = s(x) dx$.

Definición 36. Sea $X = \{X_t : t \geq 0\}$ un proceso de difusión regular y homogéneo con espacio de estados $S = (l, r)$, $l < r$, y parámetros infinitesimales $\mu(x)$ y $\sigma^2(x)$, $x \in S$. Sea $J = [c, d] \subset (l, r)$, $l < r$. Se define la *medida velocidad* inducida por la *densidad velocidad* $m(x) = 1/[s(x)\sigma^2(x)]$, como:

$$M[J] = M[c, d] = \int_c^d m(x) dx.$$

Observación 15. La medida velocidad de un intervalo $J = [c, d] \subset S$ satisface $0 < M[J] < \infty$ para $l < c < d < r$. Esto es porque como $m(x)$ es positiva y monótona creciente, y como $[c, d] \subset (l, r)$, entonces $0 < \int_c^d m(x) dx < \infty$.

Observación 16. Considérese $x \in (l, r)$, $l < r$. Se utilizará la notación $dM(x) = m(x) dx$.

Definición 37. Sea $X = \{X_t : t \geq 0\}$ un proceso de difusión regular con espacio de estados $S = (l, r)$, $l < r$. Sean $[a, b] \subset (l, r)$ un intervalo cerrado y $\tilde{S}[a, b]$ la medida escala de $[a, b]$. Entonces se define:

$$\tilde{S}(l, b) = \lim_{a \rightarrow l} \tilde{S}[a, b] \leq \infty, \quad l < a < b < r.$$

Observación 17. $\tilde{S}(l, b) = \infty$ para alguna $b \in S$ si y sólo si $\tilde{S}(l, b) = \infty$ para toda $b \in S$. Esto se debe a lo siguiente: supóngase que $\tilde{S}(l, b) = \infty$ para alguna $b \in S$. Tómesese $x \in S$ tal que $x \in [b, r)$. Entonces, por la Propiedad 1(ii) se tiene que:

$$\tilde{S}[a, x] = \tilde{S}[a, b] + \tilde{S}[b, x] \quad (2.4.1)$$

Cuando $a \rightarrow l$ el lado derecho de (2.4.1) tiende a infinito y, por tanto:

$$\tilde{S}(l, x) = \infty.$$

Ahora, tómesese $x \in [a, b]$. Entonces, al igual que antes, se tiene que:

$$\tilde{S}[a, b] = \tilde{S}[a, x] + \tilde{S}[x, b] \quad (2.4.2)$$

Cuando $a \rightarrow l$ el lado izquierdo de (2.4.2) tiende a infinito, por lo que:

$$\tilde{S}(l, x) = \infty.$$

De esta forma, si $\tilde{S}(l, b) = \infty$ para alguna $b \in S$, se tiene que $\tilde{S}(l, b) = \infty$ para toda $b \in S$. Y el recíproco de esta afirmación es inmediato.

Definición 38. Sea $X = \{X_t : t \geq 0\}$ un proceso de difusión regular con espacio de estados $S = (l, r)$, $l < r$. Se define:

$$\Sigma(l) = \lim_{a \rightarrow l} \int_a^x \tilde{S}[a, \xi] dM(\xi) = \lim_{a \rightarrow l} \int_a^x \tilde{S}[a, \xi] m(\xi) d\xi,$$

o equivalentemente,

$$\Sigma(l) = \int_l^x M[\eta, x] dS(\eta) = \int_l^x M[\eta, x] s(\eta) d\eta,$$

donde M es la medida velocidad y S la función escala.

Definición 39. Sea $X = \{X_t : t \geq 0\}$ un proceso de difusión regular con espacio de estados $S = (l, r)$, $l < r$. Se define:

$$\begin{aligned} M(l, x) &= \lim_{a \rightarrow l} M[a, x], & l < a < x < r \\ N(l) &= \int_l^x \tilde{S}[\eta, x] dM(\eta) = \int_l^x M(l, \xi) dS(\xi). \end{aligned}$$

Observación 18. 1. Si se considerara la frontera r en vez de l , los funcionales análogos para r serían, para $l < x < b < r$:

$$\begin{aligned} \tilde{S}[x, r] &= \lim_{b \rightarrow r} \tilde{S}[x, b] \\ \Sigma(r) &= \lim_{b \rightarrow r} \int_x^b \tilde{S}[\xi, b] m(\xi) d\xi = \int_x^r M[x, \eta] s(\eta) d\eta \\ M[x, r] &= \lim_{b \rightarrow r} M[x, b] \\ N(r) &= \int_x^r \tilde{S}[x, \eta] m(\eta) d\eta = \int_x^r M[\xi, r] s(\xi) d\xi. \end{aligned}$$

2. En adelante el desarrollo se hará para la frontera l , siendo análogo para la frontera r .

2.4.1 Clasificación de fronteras

Definición 40. Sea $X = \{X_t : t \geq 0\}$ un proceso de difusión regular con espacio de estados $S = (l, r)$, $l < r$. Se dice que la frontera l es *atrayernte* si $\tilde{S}(l, x) < \infty$, independientemente de la elección de $x \in S$.

Observación 19. Si la frontera l es atrayente, no necesariamente pertenece al espacio de estados del proceso ni el tiempo esperado para que el proceso llegue a ella tiene que ser finito. Un ejemplo es el caso del movimiento Browniano con $l = -\infty$, que es una frontera atrayente pero no pertenece al espacio de estados.

Definición 41. Sea $X = \{X_t : t \geq 0\}$ un proceso de difusión regular con espacio de estados $S = (l, r)$, $l < r$. Se dice que la frontera l es:

- *alcanzable* si $\Sigma(l) < \infty$
- *inalcanzable* si $\Sigma(l) = \infty$.

Lema 2. La relación entre las funciones $\tilde{S}(l, x)$, $\Sigma(l)$, $M(l, x)$ y $N(l)$ es la siguiente:

(i) $\tilde{S}(l, x) = \infty$ implica $\Sigma(l) = \infty$

(ii) $\Sigma(l) < \infty$ implica $\tilde{S}(l, x) < \infty$, es decir, si l es alcanzable entonces es atrayente.

(iii) $M(l, x) = \infty$ implica $N(l) = \infty$

(iv) $N(l) < \infty$ implica $M(l, x) < \infty$

(v) $\Sigma(l) + N(l) = \tilde{S}(l, x)M(l, x)$.

Demostración. Ver [15]. □

Observación 20. Para caracterizar totalmente un proceso de difusión, el comportamiento de éste en cualquiera de las fronteras de su espacio de estados $S = (l, r)$, $l < r$, debe ser especificado.

Definición 42. Sea $X = \{X_t : t \geq 0\}$ un proceso de difusión regular con espacio de estados $S = (l, r)$, $l < r$. Se dice que l es una *frontera regular* si $\tilde{S}(l, x) < \infty$ y $M(l, x) < \infty$, $x \in S$. Las condiciones son análogas si se considera r en vez de l .

Observación 21. Un proceso de difusión regular puede entrar y salir de una frontera regular.

Definición 43. Sea $X = \{X_t : t \geq 0\}$ un proceso de difusión regular con espacio de estados $S = (l, r)$, $l < r$. Se dice que l es una *frontera de salida* si $\Sigma(l) < \infty$ pero $M(l, x] = \infty$, $x \in S$. Las condiciones son análogas si se considera r en vez de l .

Observación 22. Si l es una frontera de salida, una vez que el proceso alcanzó l no puede producirse trayectoria alguna del proceso desde l .

Definición 44. Sea $X = \{X_t : t \geq 0\}$ un proceso de difusión regular con espacio de estados $S = (l, r)$, $l < r$. Se dice que l es una *frontera natural* (o de Feller) si $\Sigma(l) = \infty$ y $N(l) = \infty$. Las condiciones son análogas si se considera r en vez de l . Ver [15].

Observación 23. Si l es una frontera natural (o de Feller), no puede ser alcanzada en un tiempo esperado finito, o bien, no es posible que $X_0 = l$, por lo que toda barrera natural es excluida de S . Ver [15].

Definición 45. Sea $X = \{X_t : t \geq 0\}$ un proceso de difusión regular con espacio de estados $S = (l, r)$, $l < r$. Se dice que l es una *frontera de entrada* si $\bar{S}(l, x] = \infty$ a la vez que $N(l) < \infty$, $x \in S$. Las condiciones son análogas si se considera r en vez de l .

Observación 24. Si l es una frontera de entrada, no se puede llegar a ella desde el interior de S , pero es posible considerar que $X_0 = l$. Ver [15].

2.4.2. Fronteras de entrada y la distribución estacionaria

Teorema 11. Sea $X = \{X_t : t \geq 0\}$ un proceso de difusión regular con espacio de estados $S = (l, r)$, $l < r$, con l y r ambas fronteras de entrada. La distribución estacionaria del proceso X está dada por:

$$\psi(x) = \frac{m(x)}{\int_l^r m(\xi) d\xi} = \frac{1}{\sigma^2(x)s(x) \int_l^r \frac{d\xi}{\sigma^2(\xi)m(\xi)}}, \quad x \in S \quad (2.4.3)$$

Demostración. Como se obtuvo en la Sección 2.3.1, una distribución estacionaria del proceso X está dada por:

$$\psi(x) = m(x)[c_1 S(x) + c_2],$$

con

$$m(x) = \frac{1}{s(x)\sigma^2(x)}, \quad x \in S.$$

Nótese que, por definición, la función escala $S(x)$ es monótona sobre el espacio de estados S , $x \in S$, y como l y r son fronteras de entrada, entonces:

$$\lim_{x \rightarrow r} S(x) = \infty$$

y

$$\lim_{x \rightarrow l} S(x) = -\infty.$$

Esto se debe a que si $\tilde{S}(l, x) = \infty$. Entonces, $\tilde{S}(l, x) = \infty$ para toda $x \in (l, r)$. Además, como:

$$\tilde{S}(l, x) = \lim_{a \rightarrow l} \tilde{S}[a, x] = \lim_{a \rightarrow l} [S(x) - S(a)],$$

entonces para que $\tilde{S}(l, x) = \infty$ es necesario que $\lim_{a \rightarrow l} S(a) = -\infty$. Análogamente,

$$\tilde{S}(x, r) = \lim_{b \rightarrow r} \tilde{S}[x, b] = \lim_{b \rightarrow r} [S(b) - S(x)],$$

y para que $\tilde{S}(x, r) = \infty$ es necesario que $\lim_{b \rightarrow r} S(b) = \infty$. De esta forma, para preservar $\psi(x) \geq 0$ en S , es necesario que $c_1 = 0$. Y como por definición de distribución estacionaria se tiene que $\int_S \psi(x) dx = 1$, se debe elegir c_2 de modo que se satisfaga esta condición, por lo que $c_2 = \left(\int_l^r m(\xi) d\xi\right)^{-1}$. Así, la *única* distribución estacionaria de X en este caso es:

$$\psi(x) = \frac{m(x)}{\int_l^r m(\xi) d\xi} = \frac{1}{\sigma^2(x)s(x) \int_l^r \frac{d\xi}{\sigma^2(\xi)m(\xi)}}, \quad x \in S$$

y el denominador es finito ya que l y r son fronteras de entrada. □

2.5 Convergencia a una difusión

Considérese la secuencia de procesos estocásticos en tiempo discreto $X^{(N)} = \{X_k^{(N)} : k = 0, 1, 2, \dots\}$, no necesariamente cadenas de Markov, indexada por $N = 1, 2, \dots$ y con espacio de estados S un intervalo cerrado acotado

contenido en \mathbb{R} y adaptado ¹ a una secuencia creciente $\{\mathcal{F}_k^{(N)} : k \geq 0\}$ de σ -álgebras. Es usual tomar $\mathcal{F}_k^{(N)} = \sigma(X_i^{(N)} : i \leq k)$.

Sea $\Delta X_n^{(N)} = X_{n+1}^{(N)} - X_n^{(N)}$. Supóngase que $\Delta X_n^{(N)}$ satisface:

$$E[\Delta X_n^{(N)} | \mathcal{F}_n^{(N)}] = h_N \mu(X_n^{(N)}) + \varepsilon_{1,n}^{(N)} \quad (2.5.1)$$

$$E[(\Delta X_n^{(N)})^2 | \mathcal{F}_n^{(N)}] = h_N \sigma^2(X_n^{(N)}) + \varepsilon_{2,n}^{(N)} \quad (2.5.2)$$

$$E[(\Delta X_n^{(N)})^4 | \mathcal{F}_n^{(N)}] = \varepsilon_{3,n}^{(N)}, \quad (2.5.3)$$

donde $h_N > 0$ y $h_N \rightarrow 0$ conforme $N \rightarrow \infty$, y los términos de error son suficientemente pequeños de modo que satisfacen para toda $t \geq 0$:

$$\sum_{n < [t/h_N]} E[|\varepsilon_{i,n}^{(N)}|] \xrightarrow{N \rightarrow \infty} 0 \quad \text{para } i = 1, 2, 3,$$

donde $[t/h_N]$ denota el mayor entero menor o igual a t/h_N .

Observación 25. Cuando los procesos $\{X_k^{(N)} : k = 0, 1, 2, \dots\}$ son cadenas de Markov, condicionar en (2.5.1), (2.5.2) y (2.5.3) respecto a $\mathcal{F}_n^{(N)}$ equivale simplemente a condicionar sobre el estado de $X_n^{(N)}$.

Defínase el proceso en tiempo continuo

$$X_t^{(N)} = X_{[t/h_N]}^{(N)}, \quad t \geq 0.$$

Ahora, si $\{X_t^{(N)} : t \geq 0\}$ como funciones de t satisfacen:

- (1) para cada $\omega \in \Omega$ fijo son funciones continuas por la derecha con límite por la izquierda,
- (2) para todo conjunto de puntos $0 = t_0 < t_1 < t_2 < \dots < t_r$, las distribuciones finito-dimensionales de $\{X_{t_0}^{(N)}, X_{t_1}^{(N)}, X_{t_2}^{(N)}, \dots, X_{t_r}^{(N)}\}$ convergen a las de $\{X_{t_0}, X_{t_1}, X_{t_2}, \dots, X_{t_r}\}$ conforme $N \rightarrow \infty$ (donde X_{t_i} , para $i = 0, 1, \dots, r$, es un proceso de difusión con coeficientes de deriva y varianza $\mu_i(x)$ y $\sigma_i^2(x)$, $x \in S$, dados por (2.1.2) y (2.1.3), respectivamente),

¹i.e. $X_k^{(N)}$ es medible respecto a $\mathcal{F}_k^{(N)}$, $k \geq 0$

entonces para N grande la secuencia de procesos $X^{(N)} = \{X_k^{(N)} : k = 0, 1, 2, \dots\}$ puede aproximarse por la difusión $X = \{X_t : t \geq 0\}$ (ver Apéndice A).

Sea $Y^{(N)} = \{Y_t^{(N)} : t \geq 0\}$ una sucesión de cadenas de Markov en tiempo continuo, indexada por $N = 1, 2, \dots$, y defínase el proceso $Z^{(N)} = \{Z_t^{(N)} : t \geq 0\}$ por :

$$Z_t^{(N)} = a(N)[Y_{b(N)t}^{(N)} - c(N)],$$

donde $c(N)$ es la constante que centra el proceso, $a(N)$ es la constante que cambia la escala de los estados y $b(N)$ la constante que cambia la escala de tiempo. De esta manera, una unidad de tiempo del proceso $Z_t^{(N)}$ corresponde a un periodo de $b(N)$ unidades de tiempo del proceso original $Y_t^{(N)}$. Se se eligen adecuadamente las constantes $a(N)$, $b(N)$ y $c(N)$, y para toda $t \geq 0$, $Z_t^{(N)}$ como función de t satisface las condiciones (1) y (2) anteriores, se tiene que $Z^{(N)}$ converge en distribución, conforme $N \rightarrow \infty$, a un proceso de difusión $Y = \{Y_t : t \geq 0\}$. De este modo, propiedades de ciertos funcionales de las cadenas $Y^{(N)}$ pueden obtenerse mediante el cálculo de los correspondientes funcionales del proceso Y .

Capítulo 3

El modelo de Wright-Fisher y el modelo de Moran

En este capítulo se describirán dos modelos matemáticos para la evolución de poblaciones genéticas. En ambos modelos se supondrá que la población es haploide, es decir, cada gene (que en adelante se llamará "individuo") proviene de un único progenitor. Se observará un único locus ¹ fijo que, por hipótesis, tendrá dos alelos posibles, tipo A_1 y tipo A_2 . Se permitirá mutación entre ambos tipos de alelos a ciertas tasas de mutación y se considerará que la población es neutral, es decir, el tipo alélico de un individuo de la población no influye en su capacidad de reproducción. Para ambos modelos se supondrá que la población mantendrá un tamaño constante, que se denotará por N ; y se tomará X_t (con $t = 0, 1, 2, \dots$ ó $t \geq 0$) igual al número de individuos de alelo tipo A_1 en la población al tiempo t . A medida que el tiempo transcurre la evolución de la población será descrita por el proceso $X = \{X_t : t \in T\}$, con $T = \{0, 1, \dots\}$ o $T = [0, \infty)$, según se trate del modelo de Wright-Fisher o del modelo de Moran, respectivamente. En ambos casos, X será una cadena de Markov con espacio de estados $S = \{0, 1, 2, \dots, N\}$.

Se describirá primeramente en tiempo discreto el modelo de Wright-Fisher, posteriormente en tiempo continuo el modelo de Moran, y finalmente resultados de procesos de difusión para ambos modelos.

Este capítulo se basa en [5], [6], [8], [12], [13], [15], [18] y [20].

¹i.e. la posición o localización de un gene en un cromosoma.

3.1 Modelo de Wright-Fisher

3.1.1 Descripción del modelo

El modelo de Wright-Fisher es un modelo de evolución en tiempo discreto. En él, cada evento de reproducción da origen a una nueva generación, y de este modo no hay intersección de generaciones.

Demografía de la población

Dada la generación presente, la siguiente generación se forma de la siguiente manera: se seleccionan, con reemplazo, N individuos de la generación presente para reproducirse, y dicha selección se hace de forma independiente y uniforme. Así, la siguiente generación está formada por los descendientes de los individuos seleccionados.

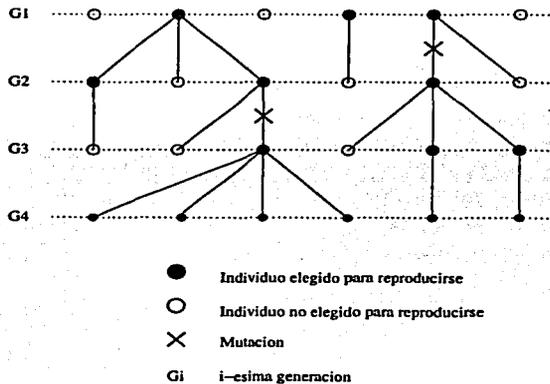


Figura 3.1: Dinámica poblacional del modelo de Wright-Fisher

Dinámica de mutación en la población

- Un individuo con alelo tipo A_1 produce un individuo con alelo tipo A_2 con probabilidad $\gamma_1 \geq 0$. En otro caso, produce un individuo de su mismo tipo con probabilidad $1 - \gamma_1 \geq 0$.

- Un individuo con alelo tipo A_2 produce un individuo con alelo tipo A_1 con probabilidad $\gamma_2 \geq 0$. En otro caso, produce un individuo de su mismo tipo con probabilidad $1 - \gamma_2 \geq 0$.

Observación 26. La dinámica de mutación supone que la mutación del alelo tipo A_1 al alelo tipo A_2 ocurre con probabilidad $\gamma_1 \geq 0$; y del alelo tipo A_2 al alelo tipo A_1 , con probabilidad $\gamma_2 \geq 0$.

Supóngase que el estado actual de la cadena X es $j \in S$, es decir, existen j individuos tipo A_1 en la población actual. Entonces, la probabilidad de que se produzca para la siguiente generación un descendiente tipo A_1 es:

$$p_j = \frac{j}{N}(1 - \gamma_1) + \frac{N-j}{N}\gamma_2. \quad (3.1.1)$$

Esto se debe a que existen dos posibilidades, mutuamente excluyentes, de que el descendiente producido sea de tipo A_1 . La primera consiste en seleccionar de la generación presente un individuo tipo A_1 que produzca un descendiente no mutante, lo cual ocurre con probabilidad $(j/N)(1 - \gamma_1)$; y la segunda consiste en elegir de la generación presente un individuo tipo A_2 que produzca un descendiente mutante, y esto ocurre con probabilidad $[(N-j)/N]\gamma_2$. Haciendo un razonamiento análogo, la probabilidad de que se produzca para la siguiente generación un descendiente tipo A_2 es:

$$q_j = \frac{j}{N}\gamma_1 + \frac{N-j}{N}(1 - \gamma_2). \quad (3.1.2)$$

Si se considera como un evento exitoso a la producción de un individuo tipo A_1 , entonces, por la evolución del modelo, dada la generación presente, la asignación de tipos alélicos en la siguiente generación puede verse como la realización de N ensayos Bernoulli con probabilidad de éxito igual a p_j y de fracaso igual a q_j . Entonces, la probabilidad de que haya k individuos tipo A_1 en la generación $n+1$, dado que en la generación n hay j , es:

$$P\{X_{n+1} = k \mid X_n = j\} = \binom{N}{k} p_j^k q_j^{N-k} \quad (3.1.3)$$

Observación 27. Aunque se está considerando una población haploide, podría considerarse una población diploide², en cuyo caso se tendrían $2N$ individuos y se tomaría $M = 2N$ como el tamaño total de la población.

²i.e. cada individuo tiene dos progenitores.

Por lo anterior, si X_n es el número de individuos tipo A_1 en la n -ésima generación, $n = 0, 1, 2, \dots$, entonces $\{X_n : n = 0, 1, \dots\}$ es una cadena de Markov discreta con espacio de estados $S = \{0, 1, \dots, N\}$ y probabilidades de transición dadas por la distribución binomial (3.1.3).

3.2 Modelo de Moran

3.2.1 Descripción del modelo

Demografía de la población

A cada unidad de tiempo un individuo es seleccionado, de manera uniforme e independiente, para producir un descendiente; y a la vez, un individuo es seleccionado, de manera uniforme e independiente, para morir.

Observación 28. Nótese que por la descripción de la evolución demográfica es posible que el individuo elegido para morir sea el mismo que se elige para reproducirse, en cuyo caso dicho individuo sería reemplazado por su propio descendiente.

Dinámica de mutación en la población

La dinámica de mutación es la misma que la presentada en el modelo de Wright-Fisher.

Observación 29. 1. En este modelo, una generación equivale a N eventos de reproducción-muerte (aunque no necesariamente los N individuos de la población son reemplazados en este lapso. Ver, por ejemplo, [7]).

2. Como la versión del modelo de Moran que se utiliza es en tiempo continuo, las dinámicas de demografía y mutación ocurren tal como se describieron, pero los eventos de nacimiento y muerte de un individuo suceden en intervalos de tiempo distribuidos exponencialmente con parámetro λ , independientemente de los eventos pasados. Entonces, para $t \geq 0$:

(a) Para cada individuo, independientemente de los otros individuos y de eventos pasados, existe una probabilidad $\lambda h + o(h)$ de morir en el intervalo $(t, t + h)$, $h > 0$.

(b) La probabilidad de más de una muerte en $(t, t + h)$ es $o(h)$, $h > 0$, donde $\lim_{h \rightarrow 0} \frac{o(h)}{h} = 0$.

Supóngase que en la población actual existen j individuos tipo A_1 , es decir, $X_t = j$, $j \in S$. Entonces:

$$\begin{aligned} j/N &= P\{\text{elegir un individuo tipo } A_1 \text{ para morir}\} \\ &= P\{\text{elegir un individuo tipo } A_1 \text{ para reproducirse}\}, \end{aligned}$$

$$\begin{aligned} (N - j)/N &= P\{\text{elegir un individuo tipo } A_2 \text{ para morir}\} \\ &= P\{\text{elegir un individuo tipo } A_2 \text{ para reproducirse}\}. \end{aligned}$$

De este modo, la probabilidad de reemplazar al individuo seleccionado para morir por un individuo tipo A_1 , está dada por (3.1.1); y la probabilidad de reemplazarlo por un individuo tipo A_2 , está dada por (3.1.2).

Observación 30. Por la demografía poblacional del modelo, y dado $X_t = j$, $j \in S$, los únicos estados distintos de j que es posible visitar en la siguiente transición son $j + 1$, $j - 1 \in S$. Entonces, por la Definición 21, $X = \{X_t : t \geq 0\}$ es un proceso de nacimiento y muerte.

Sea $h > 0$. Las tasas de nacimiento y muerte de X están dadas de la siguiente manera: si al tiempo presente la cadena X está en el estado $j \in S$,

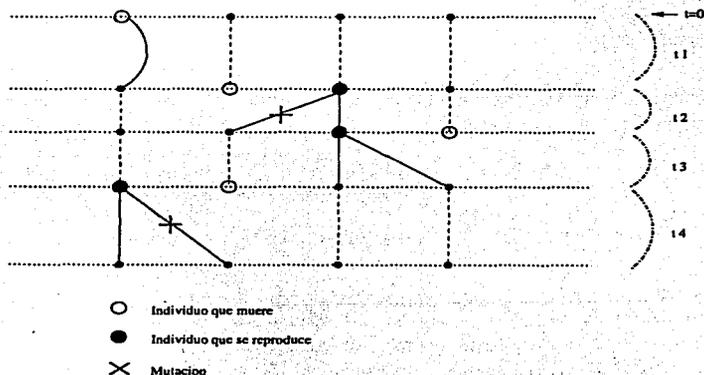


Figura 3.2: Dinámica poblacional del modelo de Moran

entonces la probabilidad de que el próximo salto sea al estado $j + 1$ es:

$$p_{jj+1} = \left(\frac{N-j}{N} \right) p_j,$$

dado que el número de individuos tipo A_1 sólo puede aumentar una unidad si muere un individuo tipo A_2 (el cual se elige con probabilidad $(N-j)/N$) y nace un individuo tipo A_1 (lo cual ocurre con probabilidad p_j). Por un razonamiento análogo se tiene que:

$$p_{jj-1} = \frac{j}{N} q_j.$$

Entonces, dado que X permanece en el estado $j \in S$ una cantidad de tiempo distribuida exponencialmente con parámetro λ , se tiene que ν_j , $j \in S$, de la Definición 12 es tal que

$$\nu_j = \lambda, \quad j \in S.$$

Por tanto, por las definiciones 17 y 22, las tasas de nacimiento y muerte de la cadena X están dadas, respectivamente, por:

$$\lambda_j = \nu_j p_{jj+1} = \lambda \left(\frac{N-j}{N} \right) p_j, \quad j = 0, 1, \dots, N$$

$$\mu_j = \nu_j p_{jj-1} = \lambda \left(\frac{j}{N} \right) q_j, \quad j = 0, 1, \dots, N.$$

De esta forma, por la Sección 1.4.1, las probabilidades de transición de X están dadas por:

$$P\{X_{t+h} = i | X_t = j\} = \begin{cases} \lambda \left(\frac{N-j}{N} \right) p_j h + o(h) & \text{si } i=j+1, j=0,1,\dots,N-1 \\ \lambda \left(\frac{j}{N} \right) q_j h + o(h) & \text{si } i=j-1, j=1,2,\dots,N \\ \lambda \left(\frac{j}{N} p_j + \frac{N-j}{N} q_j \right) h + o(h) & \text{si } i=j, j=0,1,\dots,N \\ 0 & \text{en otro caso.} \end{cases}$$

Observación 31. Si $\gamma_1 = 0 = \gamma_2$ y la cadena X se encuentra en el estado $j \in S$ al tiempo presente, entonces $p_j = j/N$, $q_j = (N-j)/N$, y los estados 0 y N son estados absorbentes, es decir, si $X_s = 0$ ($X_s = N$) para alguna $s \geq 0$, entonces $X_t = 0$ ($X_t = N$) para toda $t \geq s$. De esta forma, la cadena X no es irreducible y, por tanto, no posee una única distribución estacionaria. Pero es posible preguntarse, por ejemplo, por el tiempo de fijación

de la cadena a un tipo de alelo u otro; es decir, dado el estado inicial de la población, se pregunta sobre la cantidad de tiempo que transcurre hasta que todos los individuos son de tipo A_1 o todos son de tipo A_2 . Sin embargo, esta pregunta no es el interés de la presente tesis.

En adelante se considerará solamente el caso $\gamma_1 > 0$ y $\gamma_2 > 0$.

3.2.2 Distribución límite

Dado que para $\gamma_1 > 0$ y $\gamma_2 > 0$ la cadena X es un proceso de nacimiento y muerte irreducible con espacio de estados finito, y la cadena de saltos $X^s = \{X_n^s : n = 0, 1, \dots\}$ asociada a X es recurrente positiva, entonces X tiene una única distribución estacionaria, que coincide con su distribución límite (por el Teorema 2). Entonces, la distribución límite de X existe y está dada por el siguiente teorema.

Teorema 12. *Si $X = \{X_t : t \geq 0\}$ es la cadena de Markov que describe la evolución de una población genética de acuerdo al modelo de Moran, con dos tipos de alelos, A_1 y A_2 , probabilidad de mutación de A_1 a A_2 igual a $\gamma_1 > 0$ y de A_2 a A_1 igual a $\gamma_2 > 0$, entonces su distribución límite $P(k)$, $k \in S$, $S = \{0, 1, \dots, N\}$, está dada por:*

$$P(k) = \frac{\Gamma(N+1)}{\Gamma(k+1)\Gamma(N-k+1)} \frac{\Gamma\left(\frac{N(\gamma_1+\gamma_2)}{1-\gamma_1-\gamma_2}\right) \Gamma\left(\frac{N(1-\gamma_2)}{1-\gamma_1-\gamma_2}\right)}{\Gamma\left(\frac{N}{1-\gamma_1-\gamma_2}\right) \Gamma\left(\frac{N\gamma_1}{1-\gamma_1-\gamma_2}\right)} \frac{\Gamma\left(\frac{N\gamma_2}{1-\gamma_1-\gamma_2} + k\right) \Gamma\left(\frac{N(1-\gamma_2)}{1-\gamma_1-\gamma_2} - k\right)}{\Gamma\left(\frac{N\gamma_2}{1-\gamma_1-\gamma_2}\right) \Gamma\left(\frac{N(1-\gamma_2)}{1-\gamma_1-\gamma_2}\right)}$$

para todo $k \in S$, y donde

$$\Gamma(\alpha) = \begin{cases} (\alpha-1)! & \text{si } \alpha \text{ es un entero positivo} \\ \int_0^\infty y^{\alpha-1} e^{-y} dy & \text{en otro caso.} \end{cases}$$

Demostración. Por el Teorema 7 se tiene que la distribución límite de X está dada por:

$$P(k) = P(0) \frac{\lambda_{k-1} \cdots \lambda_1 \lambda_0}{\mu_k \cdots \mu_2 \mu_1}, \quad k = 1, 2, \dots, N,$$

donde

$$P(0) = \left[1 + \sum_{k=1}^N \left(\prod_{i=0}^{k-1} \frac{\lambda_i}{\mu_{i+1}} \right) \right]^{-1}$$

dado que el espacio de estados es finito y, por tanto:

$$\sum_{k=1}^N \left(\prod_{i=0}^{k-1} \frac{\lambda_i}{\mu_{i+1}} \right) < \infty.$$

De la Sección 3.2.1 se tiene para $i = 0, 1, \dots, N$ que:

$$\begin{aligned} \lambda_i &= \lambda \left(\frac{N-i}{N} \right) p_i \\ \mu_{i+1} &= \lambda \left(\frac{i+1}{N} \right) q_{i+1}. \end{aligned}$$

Entonces:

$$\begin{aligned} P(k) &= P(0) \left(\prod_{i=0}^{k-1} \frac{\lambda_i}{\mu_{i+1}} \right) \\ &= P(0) \left(\prod_{i=0}^{k-1} \frac{N-i}{i+1} \frac{p_i}{q_{i+1}} \right) \\ &= P(0) \frac{N!}{k! (N-k)!} \frac{p_0 p_1 \cdots p_{k-1}}{q_1 q_2 \cdots q_k} \\ &= P(0) \frac{\Gamma(N+1)}{\Gamma(k+1) \Gamma(N-k+1)} \frac{p_0 p_1 \cdots p_{k-1}}{q_1 q_2 \cdots q_k} \end{aligned}$$

y

$$\begin{aligned} P(0) &= \left[1 + \sum_{k=1}^N \left(\prod_{i=0}^{k-1} \frac{\lambda_i}{\mu_{i+1}} \right) \right]^{-1} \\ &= \left[1 + \sum_{k=1}^N \left(\frac{N!}{k! (N-k)!} \frac{p_0 p_1 \cdots p_{k-1}}{q_1 q_2 \cdots q_k} \right) \right]^{-1} \end{aligned}$$

de la cadena a un tipo de alelo u otro; es decir, dado el estado inicial de la población, se pregunta sobre la cantidad de tiempo que transcurre hasta que todos los individuos son de tipo A_1 o todos son de tipo A_2 . Sin embargo, esta pregunta no es el interés de la presente tesis.

En adelante se considerará solamente el caso $\gamma_1 > 0$ y $\gamma_2 > 0$.

3.2.2 Distribución límite

Dado que para $\gamma_1 > 0$ y $\gamma_2 > 0$ la cadena X es un proceso de nacimiento y muerte irreducible con espacio de estados finito, y la cadena de saltos $X^s = \{X_n^s : n = 0, 1, \dots\}$ asociada a X es recurrente positiva, entonces X tiene una única distribución estacionaria, que coincide con su distribución límite (por el Teorema 2). Entonces, la distribución límite de X existe y está dada por el siguiente teorema.

Teorema 12. *Si $X = \{X_t : t \geq 0\}$ es la cadena de Markov que describe la evolución de una población genética de acuerdo al modelo de Moran, con dos tipos de alelos, A_1 y A_2 , probabilidad de mutación de A_1 a A_2 igual a $\gamma_1 > 0$ y de A_2 a A_1 igual a $\gamma_2 > 0$, entonces su distribución límite $P(k)$, $k \in S$, $S = \{0, 1, \dots, N\}$, está dada por:*

$$P(k) = \frac{\Gamma(N+1)}{\Gamma(k+1)\Gamma(N-k+1)} \frac{\Gamma\left(\frac{N(\gamma_1+\gamma_2)}{1-\gamma_1-\gamma_2}\right) \Gamma\left(\frac{N(1-\gamma_2)}{1-\gamma_1-\gamma_2}\right)}{\Gamma\left(\frac{N}{1-\gamma_1-\gamma_2}\right) \Gamma\left(\frac{N\gamma_1}{1-\gamma_1-\gamma_2}\right)} \frac{\Gamma\left(\frac{N\gamma_2}{1-\gamma_1-\gamma_2} + k\right) \Gamma\left(\frac{N(1-\gamma_2)}{1-\gamma_1-\gamma_2} - k\right)}{\Gamma\left(\frac{N\gamma_2}{1-\gamma_1-\gamma_2}\right) \Gamma\left(\frac{N(1-\gamma_2)}{1-\gamma_1-\gamma_2}\right)}$$

para todo $k \in S$, y donde

$$\Gamma(\alpha) = \begin{cases} (\alpha-1)! & \text{si } \alpha \text{ es un entero positivo} \\ \int_0^\infty y^{\alpha-1} e^{-y} dy & \text{en otro caso.} \end{cases}$$

Demostración. Por el Teorema 7 se tiene que la distribución límite de X está dada por:

$$P(k) = P(0) \frac{\lambda_{k-1} \cdots \lambda_1 \lambda_0}{\mu_k \cdots \mu_2 \mu_1}, \quad k = 1, 2, \dots, N,$$

A continuación se demostrará por inducción que:

$$p_0 p_1 \cdots p_{k-1} = \frac{\Gamma\left(\frac{N\gamma_2}{1-\gamma_1-\gamma_2} + k\right)}{\Gamma\left(\frac{N\gamma_2}{1-\gamma_1-\gamma_2}\right)} \left(\frac{1-\gamma_1-\gamma_2}{N}\right)^k \quad (3.2.1)$$

$$q_1 q_2 \cdots q_k = \frac{\Gamma\left(\frac{N(1-\gamma_2)}{1-\gamma_1-\gamma_2}\right)}{\Gamma\left(\frac{N(1-\gamma_2)}{1-\gamma_1-\gamma_2} - k\right)} \left(\frac{1-\gamma_1-\gamma_2}{N}\right)^k \quad (3.2.2)$$

para toda $k = 1, 2, \dots, N$.

• Prueba de (3.2.1)

Por la Proposición 7 del Apéndice B se tiene que:

$$\frac{\Gamma\left(\frac{N\gamma_2}{1-\gamma_1-\gamma_2} + k\right)}{\Gamma\left(\frac{N\gamma_2}{1-\gamma_1-\gamma_2}\right)} = \prod_{s=0}^{k-1} \left(\frac{N\gamma_2 + s(1-\gamma_1-\gamma_2)}{1-\gamma_1-\gamma_2}\right) \quad (3.2.3)$$

para toda $k = 1, 2, \dots$. Entonces, para $k=1$ se tiene lo siguiente:

Por la ecuación (3.1.1), $p_0 = \gamma_2$. Por la ecuación (3.2.1),

$$\begin{aligned} p_0 &= \frac{\Gamma\left(\frac{N\gamma_2}{1-\gamma_1-\gamma_2} + 1\right)}{\Gamma\left(\frac{N\gamma_2}{1-\gamma_1-\gamma_2}\right)} \left(\frac{1-\gamma_1-\gamma_2}{N}\right) \\ &= \frac{N\gamma_2}{1-\gamma_1-\gamma_2} \left(\frac{1-\gamma_1-\gamma_2}{N}\right) = \gamma_2 \quad (\text{por (3.2.3)}) \end{aligned}$$

por lo que el resultado es válido para $k = 1$. Supóngase que lo es para $k \leq r$, y se demostrará que vale para $r + 1$.

Por hipótesis de inducción:

$$\begin{aligned} p_0 p_1 \cdots p_{r-1} &= \frac{\Gamma\left(\frac{N\gamma_2}{1-\gamma_1-\gamma_2} + r\right)}{\Gamma\left(\frac{N\gamma_2}{1-\gamma_1-\gamma_2}\right)} \left(\frac{1-\gamma_1-\gamma_2}{N}\right)^r \\ &= \prod_{s=0}^{r-1} \left(\frac{N\gamma_2 + s(1-\gamma_1-\gamma_2)}{1-\gamma_1-\gamma_2}\right) \left(\frac{1-\gamma_1-\gamma_2}{N}\right)^r \quad (\text{por (3.2.3)}) \\ &= \prod_{s=0}^r \left(\frac{N\gamma_2 + s(1-\gamma_1-\gamma_2)}{N}\right). \end{aligned}$$

Y por la ecuación (3.1.1),

$$p_r = \frac{r}{N}(1 - \gamma_1) + \left(\frac{N-r}{N}\right) \gamma_2 = \frac{N\gamma_2 + r(1 - \gamma_1 - \gamma_2)}{N}.$$

Entonces:

$$\begin{aligned} p_0 \cdots p_{r-1} p_r &= \prod_{s=0}^{r-1} \left(\frac{N\gamma_2 + s(1 - \gamma_1 - \gamma_2)}{N} \right) \left(\frac{N\gamma_2 + r(1 - \gamma_1 - \gamma_2)}{N} \right) \\ &= \prod_{s=0}^r \left(\frac{N\gamma_2 + s(1 - \gamma_1 - \gamma_2)}{N} \right) \\ &= \prod_{s=0}^r \left(\frac{N\gamma_2 + s(1 - \gamma_1 - \gamma_2)}{1 - \gamma_1 - \gamma_2} \right) \left(\frac{1 - \gamma_1 - \gamma_2}{N} \right)^{r+1} \\ &= \frac{\Gamma \left(\frac{N\gamma_2}{1 - \gamma_1 - \gamma_2} + (r+1) \right)}{\Gamma \left(\frac{N\gamma_2}{1 - \gamma_1 - \gamma_2} \right)} \left(\frac{1 - \gamma_1 - \gamma_2}{N} \right)^{r+1} \quad (\text{por (3.2.3)}) \end{aligned}$$

• Prueba de (3.2.2)

Por la Proposición 7 del Apéndice B se tiene que:

$$\frac{\Gamma \left(\frac{N(1-\gamma_2)}{1-\gamma_1-\gamma_2} \right)}{\Gamma \left(\frac{N(1-\gamma_2)}{1-\gamma_1-\gamma_2} - k \right)} = \prod_{s=1}^k \left(\frac{N(1-\gamma_2) - s(1-\gamma_1-\gamma_2)}{1-\gamma_1-\gamma_2} \right) \quad (3.2.4)$$

para toda $k = 1, 2, \dots$. Entonces, para $k = 1$ se tiene lo siguiente:

Por la ecuación (3.1.2),

$$q_1 = \frac{1}{N} \gamma_1 + \left(\frac{N-1}{N} \right) (1 - \gamma_2) = \frac{N(1-\gamma_2) - (1-\gamma_1-\gamma_2)}{N}.$$

Por la ecuación (3.2.2),

$$\begin{aligned} \frac{\Gamma \left(\frac{N(1-\gamma_2)}{1-\gamma_1-\gamma_2} \right)}{\Gamma \left(\frac{N(1-\gamma_2)}{1-\gamma_1-\gamma_2} - 1 \right)} \left(\frac{1 - \gamma_1 - \gamma_2}{N} \right) &\stackrel{(1)}{=} \frac{N(1-\gamma_2) - (1-\gamma_1-\gamma_2)}{1-\gamma_1-\gamma_2} \left(\frac{1 - \gamma_1 - \gamma_2}{N} \right) \\ &= \frac{N(1-\gamma_2) - (1-\gamma_1-\gamma_2)}{N} = q_1, \end{aligned}$$

donde (1) es por la igualdad (3.2.4). Por tanto, el resultado es válido para $k = 1$. Supóngase que lo es para $k \leq r$, y se demostrará que vale para $r + 1$. Por hipótesis de inducción:

$$\begin{aligned} q_1 q_2 \cdots q_r &= \frac{\Gamma\left(\frac{N(1-\gamma_2)}{1-\gamma_1-\gamma_2}\right)}{\Gamma\left(\frac{N(1-\gamma_2)}{1-\gamma_1-\gamma_2} - r\right)} \left(\frac{1-\gamma_1-\gamma_2}{N}\right)^r \\ &= \prod_{s=1}^r \left(\frac{N(1-\gamma_2) - s(1-\gamma_1-\gamma_2)}{1-\gamma_1-\gamma_2}\right) \left(\frac{1-\gamma_1-\gamma_2}{N}\right)^r \quad (\text{por (3.2.4)}) \\ &= \prod_{s=1}^r \left(\frac{N(1-\gamma_2) - s(1-\gamma_1-\gamma_2)}{N}\right). \end{aligned}$$

Y por la ecuación (3.1.2),

$$\begin{aligned} q_{r+1} &= \left(\frac{r+1}{N}\right) \gamma_1 + \left(\frac{N-(r+1)}{N}\right) (1-\gamma_2) \\ &= \frac{N(1-\gamma_2) - (r+1)(1-\gamma_1-\gamma_2)}{N}. \end{aligned}$$

Entonces,

$$\begin{aligned} q_1 q_2 \cdots q_{r+1} &= \prod_{s=1}^r \left(\frac{N(1-\gamma_2) - s(1-\gamma_1-\gamma_2)}{N}\right) \left(\frac{N(1-\gamma_2) - (r+1)(1-\gamma_1-\gamma_2)}{N}\right) \\ &= \prod_{s=1}^{r+1} \left(\frac{N(1-\gamma_2) - s(1-\gamma_1-\gamma_2)}{N}\right) \\ &= \prod_{s=1}^{r+1} \left(\frac{N(1-\gamma_2) - s(1-\gamma_1-\gamma_2)}{1-\gamma_1-\gamma_2}\right) \left(\frac{1-\gamma_1-\gamma_2}{N}\right)^{r+1} \\ &= \frac{\Gamma\left(\frac{N(1-\gamma_2)}{1-\gamma_1-\gamma_2}\right)}{\Gamma\left(\frac{N(1-\gamma_2)}{1-\gamma_1-\gamma_2} - (r+1)\right)} \left(\frac{1-\gamma_1-\gamma_2}{N}\right)^{r+1} \quad (\text{por (3.2.4)}) \end{aligned}$$

Quedan así demostradas (3.2.1) y (3.2.2). Y de este modo se sigue que:

$$P(k) = P(0) \frac{\Gamma(N+1)}{\Gamma(k+1)\Gamma(N-k+1)} \frac{\Gamma\left(\frac{N\gamma_2}{1-\gamma_1-\gamma_2} + k\right) \Gamma\left(\frac{N(1-\gamma_2)}{1-\gamma_1-\gamma_2} - k\right)}{\Gamma\left(\frac{N\gamma_2}{1-\gamma_1-\gamma_2}\right) \Gamma\left(\frac{N(1-\gamma_2)}{1-\gamma_1-\gamma_2}\right)}.$$

Para hallar una expresión para $P(0)$ considérese lo siguiente:

Lema 3. Sean $x, y \in \mathbb{R}$ y $M \in \mathbb{N}$. Sean $u_j = x + jy$, $j = 0, \dots, N-1$, y $v_j = 1 - x - jy$, $j = 1, \dots, N$. Se define:

$$S_N(x, y) = v_1 v_2 \cdots v_N + \binom{N}{1} u_0 v_2 \cdots v_N + \binom{N}{2} u_0 u_1 v_3 \cdots v_N + \cdots + u_0 u_1 \cdots u_{N-1}.$$

Entonces:

$$S_N(x, y) = (1-y)(1-2y) \cdots (1-Ny) = \prod_{k=1}^N (1-ky).$$

Demostración. Ver [18]. □

Si se toma $x = \gamma_2$ y $y = (1 - \gamma_1 - \gamma_2)/N$, entonces:

$$\begin{aligned} v_i &= 1 - x - iy = (1 - \gamma_2) - i \frac{(1 - \gamma_1 - \gamma_2)}{N}, & i = 1, \dots, N \\ u_i &= x + iy = \gamma_2 + i \frac{(1 - \gamma_1 - \gamma_2)}{N}, & i = 0, \dots, N-1. \end{aligned}$$

Y por las ecuaciones (3.1.1) y (3.1.2) se tiene que:

$$\begin{aligned} v_i &= q_i, & i = 1, \dots, N \\ u_i &= p_i, & i = 0, \dots, N-1 \end{aligned}$$

Observando que:

$$\begin{aligned} &\left(1 + \sum_{r=1}^N \frac{N!}{r!(N-r)!} \frac{p_0 p_1 \cdots p_{r-1}}{q_1 q_2 \cdots q_r} \right) = \\ &= \frac{1}{q_1 \cdots q_N} \left(q_1 \cdots q_N + \binom{N}{1} p_0 q_2 \cdots q_N + \binom{N}{2} p_0 p_1 q_3 \cdots q_N \right. \\ &\quad \left. + \binom{N}{3} p_0 p_1 p_2 q_4 \cdots q_N + \cdots + p_0 p_1 \cdots p_{N-1} \right), \end{aligned} \quad (3.2.5)$$

por el Lema 3, la ecuación (3.2.5) es igual a:

$$\prod_{k=1}^N \left[\frac{(1 - k \frac{1 - \gamma_1 - \gamma_2}{N})}{((1 - \gamma_2) - k \frac{(1 - \gamma_1 - \gamma_2)}{N})} \right].$$

Ahora, por la Proposición 8 del Apéndice B se tiene:

$$\left(\frac{N}{1-\gamma_1-\gamma_2}\right)^N \prod_{k=1}^N \left(1 - k \frac{(1-\gamma_1-\gamma_2)}{N}\right) = \frac{\Gamma\left(\frac{N}{1-\gamma_1-\gamma_2}\right)}{\Gamma\left(\frac{N(\gamma_1+\gamma_2)}{1-\gamma_1-\gamma_2}\right)}$$

$$\left(\frac{N}{1-\gamma_1-\gamma_2}\right)^N \prod_{k=1}^N \left((1-\gamma_2) - k \frac{(1-\gamma_1-\gamma_2)}{N}\right) = \frac{\Gamma\left(\frac{N(1-\gamma_2)}{1-\gamma_1-\gamma_2}\right)}{\Gamma\left(\frac{N\gamma_1}{1-\gamma_1-\gamma_2}\right)},$$

y entonces (3.2.5) es igual a:

$$\frac{\Gamma\left(\frac{N}{1-\gamma_1-\gamma_2}\right) \Gamma\left(\frac{N\gamma_1}{1-\gamma_1-\gamma_2}\right)}{\Gamma\left(\frac{N(\gamma_1+\gamma_2)}{1-\gamma_1-\gamma_2}\right) \Gamma\left(\frac{N(1-\gamma_2)}{1-\gamma_1-\gamma_2}\right)}.$$

Por tanto,

$$P(0) = \frac{\Gamma\left(\frac{N(\gamma_1+\gamma_2)}{1-\gamma_1-\gamma_2}\right) \Gamma\left(\frac{N(1-\gamma_2)}{1-\gamma_1-\gamma_2}\right)}{\Gamma\left(\frac{N}{1-\gamma_1-\gamma_2}\right) \Gamma\left(\frac{N\gamma_1}{1-\gamma_1-\gamma_2}\right)}.$$

y el resultado sigue. □

3.3 Aproximación por difusión

Otra forma de estudiar la evolución de una población de acuerdo a cualquiera de los dos modelos presentados anteriormente, es utilizando procesos de difusión en lugar de cadenas de Markov.

Como pudo observarse en la Sección 3.2.2, obtener la distribución límite del proceso que describe la composición de la población que evoluciona de acuerdo al modelo de Moran, resulta muy elaborado y se complica aún más si la población evoluciona de acuerdo al modelo de Wright-Fisher. Si se utiliza un proceso de difusión para aproximar al proceso que describe la composición de la población conforme ésta evoluciona, la expresión de la distribución límite tiene una forma más sencilla. De esta manera, si se desea muestrear valores a partir de la distribución límite, es más conveniente usar la versión del modelo que utiliza un proceso de difusión. Sin embargo, la forma de las

probabilidades de transición del proceso de Markov que describe la evolución de la población, no facilita el cálculo de los parámetros infinitesimales del proceso de difusión.

Por otro lado, obsérvese que si en el modelo de Moran se normaliza el tiempo en N^2 eventos de nacimiento y muerte, y si en el modelo de Wright-Fisher se acelera el tiempo en N unidades, se obtiene el mismo proceso de difusión (con la normalización apropiada del espacio de estados; ver [7]).

En esta sección se obtendrán un proceso de difusión que aproxima el proceso que describe la composición de la población cuando ésta evoluciona de acuerdo al modelo de Wright-Fisher, así como su correspondiente distribución estacionaria. Para el modelo de Moran se obtendrá igualmente un proceso de difusión, que resultará el mismo obtenido para el modelo de Wright-Fisher bajo la elección adecuada de ciertos parámetros.

Los resultados de esta sección pueden hallarse en Feller(1951), Karlin y Taylor(1981) y Rodrigues(2001).

3.3.1 Difusión para el modelo de Wright-Fisher

Considérese el modelo de Wright-Fisher y sea X_n el número de individuos tipo A_1 en la n -ésima generación, $n = 0, 1, 2, \dots$. Sean $\gamma_1, \gamma_2 > 0$ tales que un individuo tipo A_1 muta a uno tipo A_2 con probabilidad γ_1 , y un individuo tipo A_2 muta a uno tipo A_1 con probabilidad γ_2 . Sean $\alpha, \beta > 0$ constantes fijas tales que $\gamma_1 = \alpha/N$ y $\gamma_2 = \beta/N$ (por lo que la intensidad de mutación del tipo A_1 al tipo A_2 en la población es $N\gamma_1 = \alpha$, y del tipo A_2 al tipo A_1 es $N\gamma_2 = \beta$).

Considérese el proceso $Y^{(N)} = \{Y_t^{(N)} : t \geq 0\}$ asociado a $X = \{X_n : n = 0, 1, \dots\}$, dado por:

$$Y_t^{(N)} = \frac{X_{[Nt]}}{N}, \quad t \geq 0,$$

donde $[Nt]$ denota el mayor entero menor o igual a Nt .

El proceso $Y^{(N)}$ es un proceso en tiempo continuo que representa la proporción o frecuencia de individuos tipo A_1 en la población al tiempo $[Nt]$ para el proceso X .

Observación 32. (a) Una unidad $t = 1$ del proceso $Y^{(N)}$ corresponde a un lapso de N generaciones del proceso X .

(b) $\gamma_1 = \alpha/N$ y $\gamma_2 = \beta/N$ son las probabilidades de mutación por individuo para el proceso X .

Teorema 13. *Los parámetros infinitesimales de la difusión $Y = \{Y_t : t \geq 0\}$, que es el límite de $Y^{(N)}$ cuando $N \rightarrow \infty$, están dados por:*

$$\begin{aligned}\mu(x) &= -\alpha x + (1-x)\beta \\ \sigma^2(x) &= x(1-x),\end{aligned}$$

donde $x \in [0, 1]$.

Demostración. Tómese $h = 1/N$. Entonces el incremento del proceso $Y^{(N)} = \{Y_t^{(N)} : t \geq 0\}$ está dado por:

$$\Delta Y_{(t,h)}^{(N)} = Y_{t+\frac{1}{N}}^{(N)} - Y_t^{(N)} = \frac{X_{[Nt]+1} - X_{[Nt]}}{N}.$$

Supóngase que $Y_t^{(N)} = \frac{i}{N}$, $i \in S$. Entonces, por la dinámica del proceso X y por la definición de $Y^{(N)}$, dado que $Y_t^{(N)} = \frac{i}{N}$, $NY_{t+\frac{1}{N}}^{(N)}$ tiene distribución binomial con parámetros N y p_i . De este modo,

$$\begin{aligned}E \left[\Delta Y_{(t,h)}^{(N)} \mid Y_t^{(N)} = \frac{i}{N} \right] &= E \left[Y_{t+\frac{1}{N}}^{(N)} - Y_t^{(N)} \mid Y_t^{(N)} = \frac{i}{N} \right] \\ &\stackrel{(1)}{=} E \left[Y_{t+\frac{1}{N}}^{(N)} \mid Y_t^{(N)} = \frac{i}{N} \right] - E \left[Y_t^{(N)} \mid Y_t^{(N)} = \frac{i}{N} \right] \\ &\stackrel{(1)}{=} \frac{1}{N} E \left[NY_{t+\frac{1}{N}}^{(N)} \mid Y_t^{(N)} = \frac{i}{N} \right] - \frac{i}{N} \\ &= \frac{1}{N} N p_i - \frac{i}{N} \stackrel{(2)}{=} \left(\frac{i}{N} (1 - \gamma_1) + \frac{(N-i)}{N} \gamma_2 \right) - \frac{i}{N} \\ &= -\gamma_1 \frac{i}{N} + \left(1 - \frac{i}{N} \right) \gamma_2 \\ &\stackrel{(3)}{=} \frac{1}{N} \left(-\alpha \frac{i}{N} + \left(1 - \frac{i}{N} \right) \beta \right)\end{aligned}$$

- (1) por propiedades de la esperanza de una variable aleatoria.
- (2) por definición de p_i , dada por (3.1.1).
- (3) pues $\gamma_1 = \alpha/N$ y $\gamma_2 = \beta/N$.

Supóngase que $i/N \rightarrow x$ conforme $N \rightarrow \infty$. Entonces se tiene que:

$$\lim_{h \rightarrow 0} \frac{1}{h} E[\Delta Y_{(t,h)}^{(N)} | Y_t^{(N)} = x] = \lim_{N \rightarrow \infty} N E[\Delta Y_{(t,h)}^{(N)} | Y_t^{(N)} = x] \\ = -\alpha x + (1-x)\beta,$$

y la convergencia es uniforme para $0 \leq x \leq 1$ (ver [15]).

Usando nuevamente fórmulas para la distribución binomial, se obtiene:

$$\begin{aligned} E \left[(\Delta Y_{(t,h)}^{(N)})^2 | Y_t^{(N)} = \frac{i}{N} \right] &= E \left[(Y_{t+\frac{1}{N}}^{(N)} - Y_t^{(N)})^2 | Y_t^{(N)} = \frac{i}{N} \right] \\ &= E \left[(Y_{t+\frac{1}{N}}^{(N)})^2 - 2Y_{t+\frac{1}{N}}^{(N)} Y_t^{(N)} + (Y_t^{(N)})^2 | Y_t^{(N)} = \frac{i}{N} \right] \\ &\stackrel{(1)}{=} E \left[(Y_{t+\frac{1}{N}}^{(N)})^2 | Y_t^{(N)} = \frac{i}{N} \right] - 2E \left[Y_{t+\frac{1}{N}}^{(N)} Y_t^{(N)} | Y_t^{(N)} = \frac{i}{N} \right] \\ &\quad + E \left[(Y_t^{(N)})^2 | Y_t^{(N)} = \frac{i}{N} \right] \\ &\stackrel{(1)}{=} \frac{1}{N^2} E \left[(N Y_{t+\frac{1}{N}}^{(N)})^2 | Y_t^{(N)} = \frac{i}{N} \right] - 2 \frac{i}{N} E \left[Y_{t+\frac{1}{N}}^{(N)} | Y_t^{(N)} = \frac{i}{N} \right] + \left(\frac{i}{N} \right)^2 \\ &\stackrel{(2)}{=} \frac{1}{N^2} (N p_i (1 - p_i) + N^2 p_i^2) - \frac{2i}{N^2} (N p_i) + \frac{i^2}{N^2} \\ &= \frac{1}{N^2} (N p_i (1 - p_i) + N^2 p_i^2 - 2N i p_i + i^2) \\ &= \frac{1}{N} p_i (1 - p_i) + p_i^2 - 2 \frac{i}{N} p_i + \left(\frac{i}{N} \right)^2 \\ &= \frac{1}{N} p_i (1 - p_i) + \left(p_i - \frac{i}{N} \right)^2 \end{aligned}$$

$$\begin{aligned}
& \stackrel{(3)}{=} \frac{1}{N} \left(\frac{i}{N}(1-\gamma_1) + \left(\frac{N-i}{N}\right) \gamma_2 \right) \left(\frac{i}{N} \gamma_1 + \left(\frac{N-i}{N}\right) (1-\gamma_2) \right) \\
& + \left(\left(\frac{N-i}{N}\right) \gamma_2 - \frac{i}{N} \gamma_1 \right)^2 \\
& = \frac{1}{N} \left[\frac{i}{N} \left(\frac{N-i}{N}\right) [(1-\gamma_1)(1-\gamma_2) + \gamma_1 \gamma_2] + \left(\frac{i}{N}\right)^2 \gamma_1(1-\gamma_1) \right. \\
& \quad \left. + \left(\frac{N-i}{N}\right)^2 \gamma_2(1-\gamma_2) \right] + \left(\frac{N-i}{N}\right)^2 \gamma_2^2 - \frac{2i}{N} \left(\frac{N-i}{N}\right) \gamma_1 \gamma_2 + \left(\frac{i}{N}\right)^2 \gamma_1^2 \\
& = \frac{1}{N} \left[\frac{i}{N} \left(\frac{N-i}{N}\right) (1-\gamma_1-\gamma_2+2\gamma_1\gamma_2-2N\gamma_1\gamma_2) + \left(\frac{i}{N}\right)^2 (\gamma_1(1-\gamma_1) + N\gamma_1^2) \right. \\
& \quad \left. + \left(\frac{N-i}{N}\right)^2 (\gamma_2(1-\gamma_2) + N\gamma_2^2) \right] \\
& \stackrel{(4)}{=} \frac{1}{N} \left[\frac{i}{N} \left(\frac{N-i}{N}\right) + \frac{i}{N} \left(\frac{N-i}{N}\right) (-\gamma_1-\gamma_2+2\gamma_1\gamma_2+2\alpha\gamma_2) \right. \\
& \quad \left. + \left(\frac{i}{N}\right)^2 (\gamma_1(1-\gamma_1) + \alpha\gamma_1) + \left(\frac{N-i}{N}\right)^2 (\gamma_2(1-\gamma_2) + \beta\gamma_2) \right] \\
& = \frac{1}{N} \left[\frac{i}{N} \left(\frac{N-i}{N}\right) + O\left(\frac{1}{N}\right) \right] \tag{3.3.1}
\end{aligned}$$

donde $\lim_{N \rightarrow \infty} \frac{O(\frac{1}{N})}{\frac{1}{N}} = 0$.

- (1) propiedades de la esperanza de una variable aleatoria.
- (2) para toda variable aleatoria Z se satisface $\text{Var}[Z] = E[Z^2] - E^2[Z]$, o bien, $E[Z^2] = \text{Var}[Z] + E^2[Z]$. Como $NY_{t+\frac{1}{N}}^{(N)}$ tiene distribución binomial con parámetros N y p_i , entonces $E[NY_{t+\frac{1}{N}}^{(N)}] = Np_i$ y $\text{Var}[NY_{t+\frac{1}{N}}^{(N)}] = Np_i(1-p_i)$, por lo que $E[(NY_{t+\frac{1}{N}}^{(N)})^2] = Np_i(1-p_i) + (Np_i)^2$.
- (3) p_i está dada por la ecuación (3.1.1), y $1-p_i = q_i$ está dada por la ecuación (3.1.2).
- (4) pues $\gamma_1 = \alpha/N$ y $\gamma_2 = \beta/N$.

Entonces,

$$\lim_{h \rightarrow 0} \frac{1}{h} E[(\Delta Y_{(t,h)}^{(N)})^2 | Y_t^{(N)} = x] = \lim_{N \rightarrow \infty} N E[(\Delta Y_{(t,h)}^{(N)})^2 | Y_t^{(N)} = x] = x(1-x),$$

ya que $i/N \rightarrow x$ conforme $N \rightarrow \infty$. Y la convergencia es uniforme para $0 \leq x \leq 1$ (ver [15]).

Por otro lado, también se satisface:

$$\lim_{h \rightarrow 0} \frac{1}{h} E[(\Delta Y_{(t,h)}^{(N)})^4 | Y_t^{(N)} = x] = 0$$

(ver [15]). Lo anterior sugiere fuertemente que el proceso $Y_t^{(N)}$ converge conforme $N \rightarrow \infty$ a un proceso de difusión $Y = \{Y_t : t \geq 0\}$ cuyo espacio de estados es $S = [0, 1]$, y cuyos parámetros infinitesimales son:

$$\begin{aligned} \mu(x) &= -\alpha x + (1-x)\beta \\ \sigma^2(x) &= x(1-x), \end{aligned} \quad x \in [0, 1]$$

(ver [15] y [20]). □

Distribución estacionaria

A continuación se obtendrá la distribución estacionaria del proceso de difusión $Y = \{Y_t : t \geq 0\}$, con base en el análisis de los puntos frontera 0 y 1 del espacio de estados de Y .

Dado que para $Y = \{Y_t : t \geq 0\}$ se tiene que $\mu(x) = -\alpha x + (1-x)\beta$ y $\sigma^2(x) = x(1-x)$, $x \in [0, 1]$, entonces por definición (Sección 2.4) se obtiene lo siguiente:

$$\begin{aligned} s(x) &= \exp\left(-\int^x \frac{2\mu(\xi)}{\sigma^2(\xi)} d\xi\right) \\ &= \exp\left(-\int^x \frac{2\beta(1-\xi) - 2\alpha\xi}{\xi(1-\xi)} d\xi\right) \\ &= \exp\left(-2\beta \int^x \frac{d\xi}{\xi} + 2\alpha \int^x \frac{d\xi}{1-\xi}\right) \\ &= \exp(-2\beta \log x - 2\alpha \log(1-x)) \\ &= \exp(\log[x^{-2\beta}(1-x)^{-2\alpha}]) \\ &= x^{-2\beta}(1-x)^{-2\alpha} \end{aligned} \quad (3.3.2)$$

y

$$m(x) = \frac{1}{\sigma^2(x)s(x)} = \frac{x^{2\beta}(1-x)^{2\alpha}}{x(1-x)} = x^{2\beta-1}(1-x)^{2\alpha-1} \quad (3.3.3)$$

De esta forma,

$$S(x) = \int_a^x s(\xi) d\xi = \int_a^x \frac{d\xi}{\xi^{2\beta}(1-\xi)^{2\alpha}} \quad a \in [0, 1], a < x \quad (3.3.4)$$

$$dS(\xi) = s(\xi) d\xi = \frac{d\xi}{\xi^{2\beta}(1-\xi)^{2\alpha}} \quad (3.3.5)$$

$$M[a, x] = \int_a^x m(u) du = \int_a^x u^{2\beta-1}(1-u)^{2\alpha-1} du. \quad (3.3.6)$$

A continuación se describirá el comportamiento para la frontera 0, siendo análogo para la frontera 1.

Tómense $x \in (0, 1)$ y $\xi \in (0, x]$. Por la ecuación (3.3.4) se tiene:

$$\begin{aligned} \bar{S}(0, x) &= \lim_{a \rightarrow 0} \bar{S}[a, x] = \lim_{a \rightarrow 0} [S(x) - S(a)] \\ &= \lim_{a \rightarrow 0} \int_a^x \frac{d\xi}{\xi^{2\beta}(1-\xi)^{2\alpha}} \\ &= \int_0^x \frac{d\xi}{\xi^{2\beta}(1-\xi)^{2\alpha}} \end{aligned} \quad (3.3.7)$$

Si $2\beta \geq 1$ y $2\alpha \geq 1$, la ecuación (3.3.7) es infinita, es decir, $\bar{S}(0, x) = \infty$.

Por la ecuación (3.3.6) se tiene:

$$\begin{aligned} M(0, \xi) &= \lim_{a \rightarrow 0} M[a, \xi] \\ &= \lim_{a \rightarrow 0} \int_a^\xi u^{2\beta-1}(1-u)^{2\alpha-1} du \\ &= \int_0^\xi u^{2\beta-1}(1-u)^{2\alpha-1} du \end{aligned} \quad (3.3.8)$$

Si $2\beta - 1 \geq 0$ y $2\alpha - 1 \geq 0$, o bien, si $2\beta \geq 1$ y $2\alpha \geq 1$, la ecuación (3.3.8) es finita.

Por (3.3.8) y (3.3.5) se tiene:

$$\begin{aligned} N(0) &= \int_0^x M(0, \xi] dS(\xi) \\ &= \int_0^x \left(\int_0^\xi u^{2\beta-1} (1-u)^{2\alpha-1} du \right) \frac{d\xi}{\xi^{2\beta} (1-\xi)^{2\alpha}} \end{aligned} \quad (3.3.9)$$

Si $\beta > 0$, la ecuación (3.3.9) es finita, es decir, $N(0) < \infty$, y es infinita si $\beta = 0$.

Por lo anterior, si $2\beta \geq 1$ y $2\alpha \geq 1$ se tiene que $\tilde{S}(0, x] = \infty$ y $N(0) < \infty$, $x \in (0, 1)$. Y de manera análoga se obtiene que $\tilde{S}[x, 1) = \infty$ y $N(1) < \infty$, $x \in (0, 1)$. Por tanto, si $2\beta \geq 1$ y $2\alpha \geq 1$, las fronteras 0 y 1 del espacio de estados de Y son fronteras de entrada; y por la Sección 2.4.2 se tiene que la única distribución estacionaria de Y está dada por:

$$\begin{aligned} \psi(y) &= \frac{m(y)}{\int_0^1 m(u) du} \\ &= \frac{(1-y)^{2\alpha-1} y^{2\beta-1}}{\int_0^1 u^{2\beta-1} (1-u)^{2\alpha-1} du} \\ &= \frac{(1-y)^{2\alpha-1} y^{2\beta-1}}{\Gamma(2\alpha) \Gamma(2\beta) / \Gamma(2\alpha + 2\beta)} \\ &= \frac{\Gamma(2\alpha + 2\beta)}{\Gamma(2\alpha) \Gamma(2\beta)} (1-y)^{2\alpha-1} y^{2\beta-1}, \quad y \in [0, 1], \end{aligned} \quad (3.3.10)$$

que es la densidad Beta($2\alpha, 2\beta$), a partir de la cual es sencillo muestrear valores.

Obsérvese que si $2\alpha < 1$ y $2\beta < 1$, la distribución estacionaria del proceso $Y = \{Y_t : t \geq 0\}$, dada por:

$$\psi(y) = c_1 \left(\int_a^y \frac{dx}{(1-x)^{2\alpha} x^{2\beta}} \right) (1-y)^{2\alpha-1} y^{2\beta-1} + c_2 (1-y)^{2\alpha-1} y^{2\beta-1} \quad (3.3.11)$$

(ver Sección 2.3.1), constituye una familia biparamétrica de distribuciones estacionarias, y para obtener una en particular (si existe) se requiere mayor información sobre el proceso real (ver, por ejemplo, [15]).

3.3.2 Difusión para el modelo de Moran

Considérese el modelo de Moran y sea X_t el número de individuos de tipo A_1 al tiempo t , $t \geq 0$. Sean $\gamma_1, \gamma_2 \geq 0$ las probabilidades de mutación de un individuo tipo A_1 a uno tipo A_2 , y de un individuo tipo A_2 a uno tipo A_1 , respectivamente. Y sean $\alpha, \beta > 0$ constantes fijas tales que $\gamma_1 = \alpha/N$, $\gamma_2 = \beta/N$. Considérese el proceso $Z^{(N)} = \{Z_t^{(N)} : t \geq 0\}$, asociado al proceso de nacimiento y muerte $X = \{X_t : t \geq 0\}$, dado por:

$$Z_t^{(N)} = \frac{X_t}{N}, \quad t \geq 0.$$

El proceso $Z^{(N)}$ representa la proporción de individuos tipo A_1 en la población al tiempo t , $t \geq 0$, para el proceso X .

Teorema 14. *Los parámetros infinitesimales de la difusión $Z = \{Z_t : t \geq 0\}$, que es el límite de $Z^{(N)}$ cuando $N \rightarrow \infty$, están dados por:*

$$\begin{aligned} \mu(x) &= -\alpha x + (1-x)\beta \\ \sigma^2(x) &= x(1-x), \end{aligned}$$

donde $x \in [0, 1]$.

Demostración. Tómesese $h = 1/N$ y $\lambda = N^2$. Se define el incremento del proceso $Z^{(N)} = \{Z_t^{(N)} : t \geq 0\}$ como:

$$\Delta Z_{(t,h)}^{(N)} = Z_{t+h}^{(N)} - Z_t^{(N)}.$$

Supóngase que $Z_t^{(N)} = i/N$, $i \in S$. Entonces:

$$\begin{aligned} E \left[\Delta Z_{(t,h)}^{(N)} \mid Z_t^{(N)} = \frac{i}{N} \right] &= E \left[Z_{t+h}^{(N)} - Z_t^{(N)} \mid Z_t^{(N)} = \frac{i}{N} \right] \\ &= E \left[\frac{X_{t+h}}{N} - \frac{X_t}{N} \mid \frac{X_t}{N} = \frac{i}{N} \right] \stackrel{(1)}{=} \frac{1}{N} E [X_{t+h} - X_t \mid X_t = i] \\ &= \frac{1}{N} [P\{X_{t+h} - X_t = 1 \mid X_t = i\} - P\{X_{t+h} - X_t = -1 \mid X_t = i\}] \\ &= \frac{1}{N} [P\{X_{t+h} = i+1 \mid X_t = i\} - P\{X_{t+h} = i-1 \mid X_t = i\}] \end{aligned}$$

$$\begin{aligned}
& \stackrel{(2)}{=} \frac{1}{N} \left[\lambda \left(\frac{N-i}{N} \right) p_i h + o_1(h) - \lambda \left(\frac{i}{N} \right) q_i h + o_2(h) \right] \\
&= \frac{1}{N} \left[\lambda h \left(\left(\frac{N-i}{N} \right) p_i - \frac{i}{N} q_i \right) + o(h) \right] \\
&= \frac{1}{N} \left[\frac{N^2}{N} \left(\left(\frac{N-i}{N} \right) \left(\frac{i}{N} (1-\gamma_1) + \frac{(N-i)}{N} \gamma_2 \right) \right. \right. \\
&\quad \left. \left. - \frac{i}{N} \left(\frac{i}{N} \gamma_1 + \frac{(N-i)}{N} (1-\gamma_2) \right) \right) + o(h) \right] \\
&= \left(\frac{N-i}{N} \right) \left(\frac{i}{N} \right) (1-\gamma_1 - 1 + \gamma_2) + \left(\frac{N-i}{N} \right)^2 \gamma_2 - \left(\frac{i}{N} \right)^2 \gamma_1 + \frac{o(1/N)}{N} \\
&= \frac{1}{N^2} [(N-i)i(\gamma_2 - \gamma_1) + (N-i)^2 \gamma_2 - i^2 \gamma_1] + \frac{o(1/N)}{N} \\
&= \frac{1}{N^2} [Ni\gamma_2 - Ni\gamma_1 - i^2 \gamma_2 + i^2 \gamma_1 + N^2 \gamma_2 - 2Ni\gamma_2 + i^2 \gamma_2 - i^2 \gamma_1] + \frac{o(1/N)}{N} \\
&= \frac{1}{N^2} [-Ni\gamma_1 - Ni\gamma_2 + N^2 \gamma_2] + \frac{o(1/N)}{N} \\
& \stackrel{(3)}{=} \frac{1}{N^2} [-\alpha i - \beta i + N\beta] + \frac{o(1/N)}{N} \\
&= \frac{1}{N} \left[-\alpha \frac{i}{N} - \beta \frac{i}{N} + \beta + o(1/N) \right] \\
&= \frac{1}{N} \left[-\alpha \frac{i}{N} + \left(1 - \frac{i}{N} \right) \beta + o(1/N) \right],
\end{aligned}$$

(1) por propiedades de la esperanza de una variable aleatoria.

(2) por la Sección 3.2.1.

(3) pues $\gamma_1 = \alpha/N$ y $\gamma_2 = \beta/N$.

Por tanto, si $i/N \rightarrow x$ conforme $N \rightarrow \infty$, $x \in [0, 1]$, se tiene:

$$\lim_{h \rightarrow 0} \frac{1}{h} E[\Delta Z_{(t,h)}^{(N)} \mid Z_t^{(N)} = x] = \lim_{N \rightarrow \infty} N E[\Delta Z_{(t,h)}^{(N)} \mid Z_t^{(N)} = x] = -\alpha x + (1-x)\beta.$$

Ahora,

$$E \left[(\Delta Z_{(t,h)}^{(N)})^2 \mid Z_t^{(N)} = \frac{i}{N} \right] = E \left[(Z_{t+h}^{(N)} - Z_t^{(N)})^2 \mid Z_t^{(N)} = \frac{i}{N} \right]$$

$$\begin{aligned}
&= E \left[\frac{1}{N^2} (X_{t+h} - X_t)^2 \mid X_t = i \right] \\
&\stackrel{(1)}{=} \frac{1}{N^2} E [(X_{t+h} - X_t)^2 \mid X_t = i] \\
&= \frac{1}{N^2} [(1)^2 P\{X_{t+h} - X_t = 1 \mid X_t = i\} + (-1)^2 P\{X_{t+h} - X_t = -1 \mid X_t = i\}] \\
&= \frac{1}{N^2} [P\{X_{t+h} = i + 1 \mid X_t = i\} + P\{X_{t+h} = i - 1 \mid X_t = i\}] \\
&\stackrel{(2)}{=} \frac{1}{N^2} \left[\lambda h \left(\left(\frac{N-i}{N} \right) p_i + \frac{i}{N} q_i \right) + o(h) \right] \\
&= \frac{1}{N^2} \left[\lambda h \left(\left(\frac{N-i}{N} \right) \left(\frac{i}{N} (1 - \gamma_1) + \frac{(N-i)}{N} \gamma_2 \right) + \right. \right. \\
&\quad \left. \left. \frac{i}{N} \left(\frac{i}{N} \gamma_1 + \frac{(N-i)}{N} (1 - \gamma_2) \right) \right) + o(h) \right] \\
&= \frac{1}{N^2} \left[\frac{\lambda h}{N^2} (i(N-i)(1 - \gamma_1) + (N-i)^2 \gamma_2 + i^2 \gamma_1 + i(N-i)(1 - \gamma_2)) + o(h) \right] \\
&= \frac{1}{N^2} \left[\frac{\lambda h}{N^2} (i(N-i) - i(N-i)\gamma_1 + N^2 \gamma_2 - 2Ni\gamma_2 + i^2 \gamma_2 + i^2 \gamma_1 + i(N-i) \right. \\
&\quad \left. - i(N-i)\gamma_2) + o(h) \right] \\
&= \frac{1}{N^2} \left[\lambda h \left(2 \frac{i}{N} \left(\frac{N-i}{N} \right) - \frac{i}{N} \left(\frac{N-i}{N} \right) (\gamma_1 + \gamma_2) + \gamma_2 - 2 \frac{i}{N} \gamma_2 \right. \right. \\
&\quad \left. \left. + \left(\frac{i}{N} \right)^2 (\gamma_1 + \gamma_2) \right) + o(h) \right] \\
&\stackrel{(3)}{=} \frac{N^2}{N^2(2N)} \left[2 \frac{i}{N} \left(\frac{N-i}{N} \right) - \frac{i}{N} \left(\frac{N-i}{N} \right) \left(\frac{\alpha}{N} + \frac{\beta}{N} \right) + \frac{\beta}{N} - 2 \frac{i}{N} \frac{\beta}{N} \right. \\
&\quad \left. + \left(\frac{i}{N} \right)^2 \left(\frac{\alpha}{N} + \frac{\beta}{N} \right) \right] + \frac{1}{N^2} o(h) \\
&= \frac{1}{N} \left[\frac{i}{N} \left(1 - \frac{i}{N} \right) \right] + \frac{1}{2N} \left[-\frac{i}{N} \left(\frac{N-i}{N} \right) \left(\frac{\alpha}{N} + \frac{\beta}{N} \right) + \frac{\beta}{N} \right. \\
&\quad \left. - 2 \frac{i}{N} \frac{\beta}{N} + \left(\frac{i}{N} \right)^2 \left(\frac{\alpha}{N} + \frac{\beta}{N} \right) \right] + \frac{1}{N^2} o(h)
\end{aligned}$$

(1) por propiedades de la esperanza de una variable aleatoria.

Capítulo 4

El proceso n -coalescente de Kingman y el proceso de líneas de descendencia

En este capítulo se describirá primeramente el *proceso n -coalescente de Kingman* para una población haploide en la que no se considera dinámica de mutación. Dicho proceso describe para una muestra de tamaño n extraída de la población, las relaciones familiares entre sus individuos buscando el ancestro común a ellos.

Posteriormente se describirá el *proceso de líneas de descendencia* para una población haploide con dinámica de mutación tal que cada mutación que ocurre es a un tipo alélico nuevo en la población. Tomando una muestra de la población, este proceso agrupa a los individuos de la muestra en dos tipos de clases de equivalencia, según provengan de ancestros mutantes o de ancestros no mutantes. Se observará que el proceso n -coalescente de Kingman es un caso particular del proceso de líneas de descendencia cuando no se considera dinámica de mutación en la población.

Y dada la simplicidad del modelo de Moran y la correspondencia que existe entre éste y el modelo de Wright-Fisher (según los resultados del Capítulo 3), se describirán ambos procesos exclusivamente para el modelo de Moran, siendo los resultados igualmente válidos para el modelo de Wright-Fisher considerando el cambio adecuado en la escala de tiempo. Este capítulo se basa en los artículos [7], [16], [17] y [25], así como en [20].

4.1 El proceso n -coalescente de Kingman

Considérese una población genética haploide de tamaño N grande, en la que existe una relación de parentesco entre los individuos. Se tomará de ella una muestra sin reemplazo de tamaño n , y se rastreará hacia atrás en el tiempo la relación de parentesco (es decir, los ancestros) de los individuos de esta muestra. En ese rastreo existirá un tiempo en que dos de los individuos en la muestra tendrán un ancestro común. Se dice entonces que sus líneas ancestrales han *coalescido* en este ancestro. Avanzando más atrás en el tiempo ocurrirá una segunda coalescencia, y así sucesivamente hasta que todos los ancestros coalezcan en un solo individuo, que será el ancestro común más reciente de todos los individuos en la muestra.

En adelante se considerará que la población evoluciona de acuerdo al modelo de Moran.

Definición 46. El n -coalescente de Kingman es un proceso que describe las relaciones familiares entre individuos de una muestra sin reemplazo de tamaño n , retirada de una población haploide de tamaño N .

Definición 47. Sea $t_0 > 0$ un tiempo fijo no aleatorio hasta el cual se deja evolucionar a la población. Tómese en t_0 una muestra sin reemplazo de tamaño n de la población. Sea $t > 0$. Considérese la relación de equivalencia entre individuos de la muestra, definida de la siguiente manera: dos individuos están en la misma *clase de equivalencia al tiempo t* si comparten el mismo ancestro al tiempo $t_0 - t$.

Sea \mathcal{E}_n el conjunto finito de las relaciones de equivalencia en $\{1, 2, \dots, n\}$. De este modo, si $R \in \mathcal{E}_n$, entonces $|R|$ corresponde al número de clases de equivalencia en R . Obsérvese que \mathcal{E}_n contiene dos relaciones de equivalencia particulares, la más refinada y la menos refinada, a saber:

$$\Delta = \{(1), (2), \dots, (n)\}$$

$$\Theta = \{(1, 2, \dots, n)\},$$

es decir, Δ es la clase de equivalencia donde cada individuo está relacionado únicamente consigo mismo, y Θ es la clase de equivalencia donde todos los individuos están relacionados entre sí.

Observación 34. (a) Sean $\xi, \eta \in \mathcal{E}_n$. Se dice que $\xi \prec \eta$ si η puede obtenerse combinando dos de las clases de equivalencia de ξ .

(b) Si $\xi \prec \eta$, entonces $|\xi| = |\eta| + 1$.

Definición 48. Una cadena de Markov en tiempo continuo $R = \{R_t : t \geq 0\}$ con espacio de estados \mathcal{E}_n es llamada un *n-coalescente* si:

(i) $R_0 = \Delta = \{(i) : i = 1, 2, \dots, n\}$

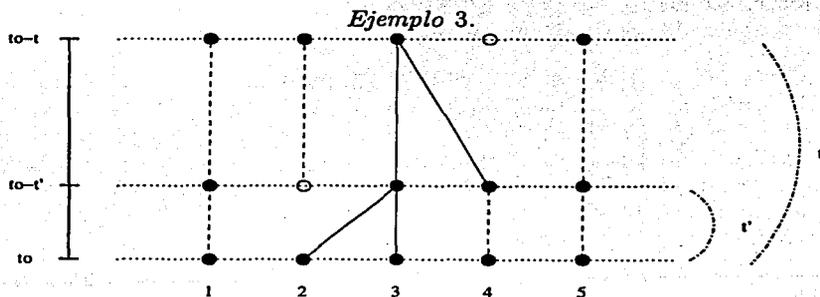
(ii) tiene tasa de transición

$$q_{\xi\eta} = \lim_{h \rightarrow 0} \frac{1}{h} P\{R_{t+h} = \eta \mid R_t = \xi\}$$

dada para $\xi, \eta \in \mathcal{E}_n$, $\xi \neq \eta$, de la siguiente forma:

$$q_{\xi\eta} = \begin{cases} 1 & \text{si } \xi \prec \eta \\ 0 & \text{en otro caso.} \end{cases}$$

Definición 49. Se define a R_t como el conjunto de las clases de equivalencia al tiempo t , es decir, las clases de equivalencia presentes al tiempo $t_0 - t$, $0 \leq t \leq t_0$, donde $R_0 = \Delta$, es decir, en R_0 cada individuo es su propio ancestro al tiempo t_0 .



$$R_0 = \{(1), (2), (3), (4), (5)\}$$

$$R_{t'} = \{(1), (2, 3), (4), (5)\}$$

$$R_t = \{(1), (2, 3, 4), (5)\}$$

Entonces $R = \{R_t : t \geq 0\}$ es una cadena de Markov en tiempo continuo con estado inicial Δ , estado absorbente Θ y tasa de transición $q_{\xi\eta}$ dada por (ii) de la Definición 48.

Observación 35. (a) R pasa de Δ a Θ a través de una secuencia de estados de \mathcal{E}_n que guardan entre ellos la relación " \prec ".

(b) De la Definición 12 se tiene que R permanece en el estado ξ por un tiempo exponencialmente distribuido con parámetro

$$\nu_\xi = \sum_{\eta \neq \xi} q_{\xi\eta} = \binom{|\xi|}{2} = \frac{|\xi|(|\xi| - 1)}{2}.$$

(c) Si el proceso R está en el estado $\xi \in \mathcal{E}_n$ al tiempo t , entonces las probabilidades de transición de R están dadas por:

$$(a) P\{R_{t+h} = \eta \mid R_t = \xi\} = q_{\xi\eta}h + O(h^2), \quad \eta \in \mathcal{E}_n \text{ tal que } \xi \prec \eta.$$

$$(b) P\{R_{t+h} = \xi \mid R_t = \xi\} = 1 - \nu_\xi h + O(h^2)$$

para $h > 0$ y la probabilidad de cualquier otra transición es de orden $O(h^2)$, donde $\lim_{h \rightarrow 0} \frac{O(h^2)}{h^2} = 0$.

Definición 50. Sea $D = \{D_t : t \geq 0\}$ donde $D_t = |R_t|$, es decir, D_t es el número de clases de equivalencia al tiempo t .

Observación 36. (a) Sea $h > 0$. Si $R_t = \xi$, para alguna $\xi \in \mathcal{E}_n$, entonces el estado $\eta \in \mathcal{E}_n$ tal que $R_{t+h} = \eta$ debe ser tal que $\xi \prec \eta$, por lo que $|\eta| = |\xi| - 1$. De esta forma, si $|R_t| = k$ y ocurre una transición, entonces $|R_{t+h}| = k - 1$. Por tanto, $D = \{D_t : t \geq 0\}$ es un proceso de muerte pura con espacio de estados $S = \{1, 2, \dots, n\}$ tal que si $D_t = k$, entonces D permanece una cantidad de tiempo τ_k en k con distribución exponencial de parámetro

$$\nu_k = \binom{k}{2} = \frac{k(k-1)}{2}. \quad (4.1.1)$$

y después cambia al estado $k-1$, $k = 2, 3, \dots, n$. De este modo, para el proceso D la tasa de muerte es $\mu_k = k(k-1)/2$, y la tasa de nacimiento

es cero. Así, las probabilidades de transición del proceso D son:

$$P\{D_{t+h} = l \mid D_t = k\} = \begin{cases} \mu_k h + o(h) & \text{si } l = k - 1 \\ 1 - \mu_k h + o(h) & \text{si } l = k \\ 0 & \text{en otro caso.} \end{cases}$$

- (b) Los estados n y 1 de D corresponden al estado inicial $\Delta = \{(1), (2), \dots, (n)\}$ y al estado absorbente $\Theta = \{(1, 2, \dots, n)\}$ de R , respectivamente.

Definición 51. Se define a $T = \inf\{t \geq 0 : D_t = 1\} = \inf\{t \geq 0 : R_t = \Theta\}$ como el mínimo tiempo de absorción de D en el estado 1 (o bien, de R_t en el estado Θ).

Observación 37. Nótese que $T = \sum_{k=2}^n \tau_k$ y los τ_k , $k = 2, 3, \dots, n$, son independientes y cada uno tiene distribución exponencial con parámetro μ_k dado por (4.1.1). Se sigue entonces que el tiempo esperado de absorción de D en el estado 1 es:

$$\begin{aligned} E[T] &= \sum_{k=2}^n E[\tau_k] = \sum_{k=2}^n \frac{2}{k(k-1)} = 2 \left(\sum_{k=2}^n \frac{1}{k-1} - \sum_{k=2}^n \frac{1}{k} \right) \\ &= 2 \left(\sum_{k=1}^{n-1} \frac{1}{k} - \sum_{k=2}^n \frac{1}{k} \right) = 2 \left(1 - \frac{1}{n} \right). \end{aligned}$$

Entonces, para n grande, $E[T]$ se aproxima a 2 . De esta forma, cuando n es grande, se necesitan 2 generaciones para llegar al ancestro común de la muestra de tamaño n . Como en el modelo de Moran una generación corresponde a N eventos de nacimiento y muerte, se tiene que para una muestra de tamaño N , $E[T]$ es aproximadamente $2N$.

Definición 52. Se define a $\mathcal{R} = \{\Delta = \mathcal{R}_n, \mathcal{R}_{n-1}, \dots, \mathcal{R}_2, \mathcal{R}_1 = \Theta\}$ como la sucesión en \mathcal{E}_n de los estados visitados por el proceso R hasta su absorción en Θ .

Observación 38. Nótese que R permanece en \mathcal{R}_k un periodo de tiempo τ_k , y que $|\mathcal{R}_k| = k$, pues si para alguna $t > 0$, $R_t = \mathcal{R}_k$ y $D_t = k$, entonces $|\mathcal{R}_k| = |R_t| = D_t = k$. De esta forma, \mathcal{R} es una cadena de Markov discreta con espacio de estados \mathcal{E}_n , estado inicial Δ y estado absorbente Θ , y corresponde a la cadena de saltos del proceso n-coalescente R .

Teorema 15. Para el n-coalescente, el proceso de muerte pura $D = \{D_t : t \geq 0\}$ y la cadena de saltos $\mathcal{R} = \{\mathcal{R}_k : k = n, n-1, \dots, 1\}$ son independientes y

$$R_t = \mathcal{R}_{D_t}, \quad t \geq 0.$$

Además, la cadena de Markov \mathcal{R} tiene probabilidades de transición:

$$P\{\mathcal{R}_{k-1} = \eta \mid \mathcal{R}_k = \xi\} = \begin{cases} \frac{2}{k(k-1)} & \text{si } \xi \prec \eta \\ 0 & \text{en otro caso,} \end{cases} \quad (4.1.2)$$

siempre que $\xi \in \mathcal{E}_n$, $|\xi| = k$ y $2 \leq k \leq n$. Y la distribución de \mathcal{R}_k , $k = n, n-1, \dots, 1$ está dada por:

$$P\{\mathcal{R}_k = \xi\} = \frac{(n-k)! k! (k-1)!}{n!(n-1)!} \eta_1! \eta_2! \dots \eta_k! \quad (4.1.3)$$

donde $\eta_1, \eta_2, \dots, \eta_k$ son los tamaños de las clases de equivalencia de ξ .

Demostración. Por el Capítulo 1 se tiene para el n-coalescente que si $\mathcal{R}_k = \xi$ y $|\xi| = k$, entonces $D_t = k$ y, por tanto, $\nu_\xi = \nu_k = \mu_k$. Y además los τ_k , $k = n, n-1, \dots, 2$, son independientes con distribución exponencial de parámetro μ_k , por la Observación 37. Se sigue entonces que, dado \mathcal{R} , la distribución de D_t para cada t , es la misma que su distribución no condicionada y, por tanto, dado \mathcal{R} , las distribuciones conjuntas de D son las mismas que sus distribuciones no condicionadas. Luego, D y \mathcal{R} son independientes.

Por definición, se tiene que $D_t = |R_t|$, y si $|R_t| = k$ también por definición $R_t = \mathcal{R}_k$. De esta forma, para todo $t \geq 0$:

$$R_t = \mathcal{R}_{D_t}.$$

Por la definición de $q_{\xi\eta}$ y por el Capítulo 1 se tiene que:

$$P\{\mathcal{R}_{k-1} = \eta \mid \mathcal{R}_k = \xi\} = \begin{cases} \frac{q_{\xi\eta}}{\nu_\xi} & \text{para } \xi \prec \eta \\ 0 & \text{en otro caso.} \end{cases}$$

Entonces, si $|\xi| = k$ se tiene que $\nu_\xi = k(k-1)/2$ y el resultado se sigue por la definición de $q_{\xi\eta}$, el hecho de que $\nu_\xi = \sum_{\xi \prec \eta} q_{\xi\eta}$ y dado que existen $\binom{k}{2}$ posibles formas en que coalescen dos clases de un total de k .

Se probará (4.1.3) por inducción sobre k con el tiempo en reversa. Considérense:

$$P\{\mathcal{R}_k = \xi\}, \xi \in \mathcal{E}_n \text{ y } |\xi| = k.$$

Tómese $k = n$. Entonces, por definición,

$$P\{\mathcal{R}_n = \xi\} = \begin{cases} 1 & \text{si } \xi = \Delta \\ 0 & \text{en otro caso.} \end{cases}$$

Lo anterior concuerda con (4.1.3) sustituyendo $k = n$ y tomando $\eta_i = 1$, $i = 1, 2, \dots, n$, por lo que el resultado es válido para $k = n$. Supóngase que lo es para $k = l < n$ y se demostrará que vale para $k = l - 1$. Entonces:

$$\begin{aligned} P\{\mathcal{R}_{l-1} = \eta\} &\stackrel{(1)}{=} \sum_{\xi \prec \eta} P\{\mathcal{R}_{l-1} = \eta \mid \mathcal{R}_l = \xi\} P\{\mathcal{R}_l = \xi\} \\ &\stackrel{(2)}{=} \sum_{\xi \prec \eta} \frac{2}{l(l-1)} P\{\mathcal{R}_l = \xi\} \end{aligned} \quad (4.1.4)$$

(1) por la ley de la probabilidad total.

(2) por (4.1.2).

Sean $\eta_1, \eta_2, \dots, \eta_{l-1}$ los tamaños de las clases de equivalencia en η . Entonces, para algún m , $m = 1, 2, \dots, l - 1$, se tiene que η_m es la suma de los tamaños de dos clases de equivalencia de ξ . Sea M tal que $1 \leq M \leq \eta_m - 1$ y sean $\eta_1, \dots, \eta_{m-1}, M, \eta_m - M, \eta_{m+1}, \dots, \eta_{l-1}$ los tamaños de las clases de equivalencia en ξ . Entonces (4.1.4) equivale a:

$$\begin{aligned} P\{\mathcal{R}_{l-1} = \eta\} &\stackrel{(1)}{=} \sum_{m=1}^{l-1} \sum_{M=1}^{\eta_m-1} \frac{2}{l(l-1)} \frac{(n-l)! l! (l-1)!}{n! (n-1)!} \eta_1! \cdots \eta_{m-1}! M! (\eta_m - M)! \\ &\quad \eta_{m+1}! \cdots \eta_{l-1}! \frac{1}{2} \binom{\eta_m}{M} \\ &= \frac{(n-l)! l! (l-1)!}{n! (n-1)! l(l-1)} \sum_{m=1}^{l-1} \sum_{M=1}^{\eta_m-1} \eta_1! \cdots \eta_{m-1}! M! (\eta_m - M)! \eta_{m+1}! \\ &\quad \cdots \eta_{l-1}! \frac{\eta_m!}{(\eta_m - M)! M!} \end{aligned}$$

$$\begin{aligned}
&= \frac{(n-l)!(l-1)!(l-2)!}{n!(n-1)!} \eta_1! \cdots \eta_{m-1}! \eta_m! \eta_{m+1}! \cdots \eta_{l-1}! \sum_{m=1}^{l-1} \sum_{M=1}^{\eta_m-1} 1 \\
&= \frac{(n-l)!(l-1)!(l-2)!}{n!(n-1)!} \eta_1! \cdots \eta_{l-1}! \sum_{m=1}^{l-1} (\eta_m - 1) \\
&= \frac{(n-l)!(l-1)!(l-2)!}{n!(n-1)!} \eta_1! \cdots \eta_{l-1}! \left(\sum_{m=1}^{l-1} \eta_m - \sum_{m=1}^{l-1} 1 \right) \\
\text{(2)} \quad &= \frac{(n-l)!(l-1)!(l-2)!}{n!(n-1)!} \eta_1! \cdots \eta_{l-1}! [n - (l-1)] \\
&= \frac{(n-l+1)(n-l)!(l-1)!(l-2)!}{n!(n-1)!} \eta_1! \cdots \eta_{l-1}! \\
&= \frac{(n-l+1)!(l-1)!(l-2)!}{n!(n-1)!} \eta_1! \cdots \eta_{l-1}!
\end{aligned}$$

y el resultado se sigue.

(1) por lo siguiente:

- $|\eta| = |\xi| - 1 = l - 1$ y $1 \leq M \leq \eta_m - 1$.
- por hipótesis de inducción, $P\{\mathcal{R}_l = k\} = [(n-l)! / (l-1)! / n! (n-1)!] \eta_1! \cdots \eta_{m-1}! M! (\eta_m - M)! \eta_{m+1}! \cdots \eta_{l-1}!$
- porque $\frac{1}{2} \binom{\eta_m}{M}$ es el número de posibles maneras, sin importar el orden, en que pueden formarse un conjunto de tamaño M y otro de tamaño $\eta_m - M$, a partir de un conjunto de tamaño η_m .

(2) porque la suma de los tamaños de las clases de equivalencia en η es igual al tamaño de la muestra, es decir, $\sum_{m=1}^{l-1} \eta_m = n$.

□

Proposición 4. Para $l < k$, $|\xi| = k$, $|\eta| = l$, $\xi \subset \eta$,

$$P\{\mathcal{R}_l = \eta \mid \mathcal{R}_k = \xi\} = \frac{(k-l)! l! (l-1)!}{k! (k-1)!} \eta_1! \eta_2! \cdots \eta_l!,$$

donde $\eta_1, \eta_2, \dots, \eta_k$ son los tamaños de las clases de equivalencia de η inducidas por ξ en \mathcal{E}_k .

Demostración. Por inducción sobre l . Para $l = k - 1$, se obtiene (4.1.2), pues $\xi \prec \eta$ y, por tanto, $\eta_i = 2$ para alguna $i = 1, 2, \dots, l$. Se supone cierto el resultado para $l < k - 1$ y se prueba para $l - 1$ de manera análoga a como se hizo en el Teorema 15. \square

Proposición 5. Sean $R = \{R_t : t \geq 0\}$ un proceso n -coalescente, $\mathcal{R} = \{\mathcal{R}_k : k = n, n - 1, \dots, 1\}$ la cadena de saltos asociada a R y $D = \{D_t : t \geq 0\}$ la cadena que cuenta el número de ancestros de la muestra de tamaño n , de una población evolucionando de acuerdo al modelo de Moran. Entonces, para $\xi \in \mathcal{E}_n$ tal que $|\xi| = k$, se tiene que:

$$P\{R_t = \xi\} = P\{D_t = k\}P\{\mathcal{R}_k = \xi\}.$$

Demostración. Sea $\xi \in \mathcal{E}_n$ tal que $|\xi| = k$. Entonces:

$$\begin{aligned} P\{R_t = \xi\} &\stackrel{(1)}{=} P\{\mathcal{R}_{D_t} = \xi\} \\ &= \sum_{l=1}^n P\{\mathcal{R}_l = \xi \mid D_t = l\}P\{D_t = l\} \\ &\stackrel{(2)}{=} P\{\mathcal{R}_k = \xi \mid D_t = k\}P\{D_t = k\} \\ &\stackrel{(3)}{=} P\{\mathcal{R}_k = \xi\}P\{D_t = k\} \end{aligned}$$

(1) pues $R_t = \mathcal{R}_{D_t}$, $t \geq 0$, por el Teorema 15.

(2) pues $|\xi| = k$.

(3) pues \mathcal{R} y D son independientes, por el Teorema 15. \square

Proposición 6.

$$P\{D_t = k\} = P\left\{\sum_{i=k+1}^n \tau_i \leq t\right\} - P\left\{\sum_{i=k}^n \tau_i \leq t\right\}.$$

Demostración. El evento $\{D_t = k\}$ equivale al evento

$$\left\{t_0 - \sum_{i=k}^n \tau_i \leq t_0 - t \leq t_0 - \sum_{i=k+1}^n \tau_i\right\},$$

o bien, al evento

$$\left\{ \sum_{i=k+1}^n \tau_i \leq t \leq \sum_{i=k}^n \tau_i \right\}.$$

Entonces:

$$\begin{aligned} P\{D_t = k\} &= P\left\{ \sum_{i=k+1}^n \tau_i \leq t, \sum_{i=k}^n \tau_i \geq t \right\} \\ &= P\left\{ \sum_{i=k+1}^n \tau_i \leq t, \tau_k + \sum_{i=k+1}^n \tau_i \geq t \right\} \\ &= P\left\{ \tau_k + \sum_{i=k+1}^n \tau_i \geq t \mid \sum_{i=k+1}^n \tau_i \leq t \right\} P\left\{ \sum_{i=k+1}^n \tau_i \leq t \right\} \\ &= \left[1 - P\left\{ \tau_k + \sum_{i=k+1}^n \tau_i \leq t \mid \sum_{i=k+1}^n \tau_i \leq t \right\} \right] P\left\{ \sum_{i=k+1}^n \tau_i \leq t \right\} \\ &= P\left\{ \sum_{i=k+1}^n \tau_i \leq t \right\} - P\left\{ \tau_k + \sum_{i=k+1}^n \tau_i \leq t, \sum_{i=k+1}^n \tau_i \leq t \right\} \\ &= P\left\{ \sum_{i=k+1}^n \tau_i \leq t \right\} - P\left\{ \sum_{i=k}^n \tau_i \leq t \right\} \end{aligned}$$

□

4.2 El proceso de líneas de descendencia

4.2.1 Generalidades

Considérese para cada tiempo discreto n en $\{\dots, -2, -1, 0, 1, 2, \dots\}$ una población de tamaño N que evoluciona de acuerdo al modelo de Moran, con un número infinito de alelos; es decir, la demografía de la población es la misma que la descrita en la Sección 3.2.1, pero la dinámica de mutación está dada de la siguiente manera: el descendiente producido será un mutante con un nuevo tipo alélico (es decir, un tipo alélico que no existió previamente ni existe actualmente en la población) con probabilidad γ , y tendrá el mismo

tipo que su progenitor con probabilidad $1 - \gamma$.

Tómese una muestra sin reemplazo de tamaño m de la población al tiempo $n_0 > 0$, y considérese su composición respecto a una población al tiempo $n_0 - n$, $n \in \{0, 1, \dots\} \subset [0, n_0]$, es decir, al tiempo n si se regresa en el tiempo n unidades contando a partir del tiempo n_0 . Los individuos de dicha muestra al tiempo n estarán divididos en dos tipos de clases que se definen como sigue.

Definición 53. Los individuos en la muestra que descienden de algún ancestro al tiempo $n_0 - n$, sin que haya ocurrido una mutación, se dice que pertenecen a una *clase de equivalencia antigua al tiempo n* . Se denotará por D_n al número de clases antiguas al tiempo n , $n \in \{0, 1, \dots\}$ (ver Figura 4.1).

Definición 54. Los individuos en la muestra que descienden de ancestros mutantes al tiempo más reciente que $n_0 - n$, se dice que pertenecen a una *clase de equivalencia nueva al tiempo n* . Se denotará por F_n al número de clases nuevas al tiempo n , $n \in \{0, 1, \dots\}$ (ver Figura 4.2).

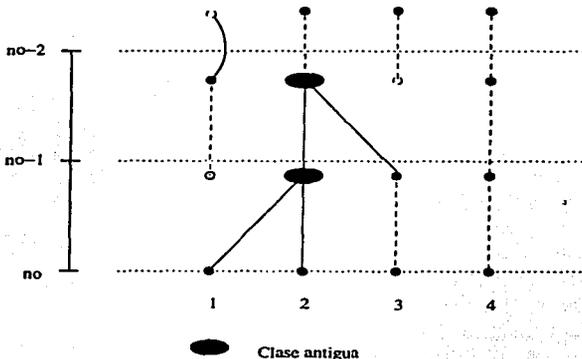


Figura 4.1: Los individuos 1 y 2 están en una misma clase antigua al tiempo 1, y los individuos 1, 2 y 3 están en una misma clase antigua al tiempo 2.

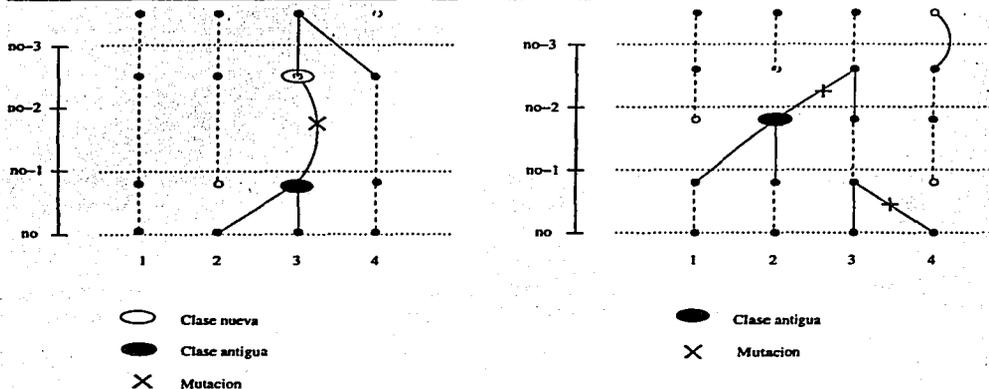


Figura 4.2: En la figura a la izquierda, los individuos 2 y 3 están en una misma clase antigua al tiempo 1, pero por una mutación ocurrida entre $n_0 - 2$ y $n_0 - 3$, dichos individuos están en una misma clase nueva al tiempo 3. En la figura a la derecha, el individuo 4 está en una clase antigua al tiempo 0, pero por una mutación ocurrida entre n_0 y $n_0 - 1$, está en una clase nueva al tiempo 1. Los individuos 1 y 2 están en una misma clase antigua al tiempo 2, pero dicha clase se transforma, por una mutación ocurrida entre $n_0 - 2$ y $n_0 - 3$, en una clase nueva al tiempo 3.

- Observación 39. (a) Los individuos en una misma clase antigua al tiempo n comparten el mismo ancestro al tiempo $n_0 - n$ sin que intervenga mutación (ver Figura 4.1).
- (b) Los individuos en una misma clase nueva al tiempo n son descendientes de un individuo que al tiempo $n_0 - k$, con $0 < k < n$, es un mutante (ver Figura 4.2).
- (c) Nótese que el proceso D dado por $D = \{D_n : n = 0, 1, \dots\}$ es una cadena de Markov con espacio de estados $\{0, 1, \dots, m\}$ y tal que $D_0 = m$, ya que un individuo puede ser considerado como su propio ancestro.
- (d) Se denotará por ξ_i , $i = 1, 2, \dots, D_n$, a las clases de equivalencia antiguas

al tiempo n , y por λ_i al número de individuos en ξ_i , es decir, $\lambda_i = |\xi_i|$, $i = 1, 2, \dots, D_n$.

- (e) Nótese que el proceso F dado por $F = \{F_n : n = 0, 1, \dots\}$ es una cadena de Markov con espacio de estados $\{0, 1, \dots, m\}$, tal que $F_0 = 0$.
- (f) Se denotará por η_j , $j = 1, 2, \dots, F_n$, a las clases de equivalencia nuevas al tiempo n , y por μ_j al número de individuos en η_j , es decir, $\mu_j = |\eta_j|$, $j = 1, 2, \dots, F_n$.
- (g) El orden en las clases de equivalencia antiguas y en las nuevas es irrelevante, pero es importante hacer distinción entre las clases que son nuevas y las que son antiguas.
- (h) Para cada $n = 0, 1, \dots$ se tiene que:

$$m = \sum_{i=1}^{D_n} \lambda_i + \sum_{i=1}^{F_n} \mu_i \quad (4.2.1)$$

4.2.2 Descripción del proceso

Definición 55. El proceso $R = \{R_n : n = 0, 1, \dots\}$ que a cada tiempo n representa las clases de equivalencia nuevas y antiguas de la muestra de la población, se denomina *proceso de líneas de descendencia*. Si al tiempo n se tienen las clases antiguas ξ_1, \dots, ξ_{D_n} y las clases nuevas $\eta_1, \dots, \eta_{F_n}$, entonces:

$$R_n = \{\xi_1, \xi_2, \dots, \xi_{D_n}; \eta_1, \eta_2, \dots, \eta_{F_n}\} = \{\bar{\xi}; \bar{\eta}\}.$$

• Transiciones del proceso R

A) *Unión de dos clases antiguas*

Supóngase que $R_n = \{\xi_1, \xi_2, \dots, \xi_{D_n}; \eta_1, \eta_2, \dots, \eta_{F_n}\} = \{\bar{\xi}; \bar{\eta}\}$, y sean $\xi_i, \xi_j \in \bar{\xi}$ dos clases antiguas al tiempo n . Supóngase que los individuos en cada una de estas clases son descendientes sin mutación de un mismo ancestro al tiempo $n_0 - (n+1)$. Entonces, el conjunto de clases antiguas:

$$\bar{\xi}^{(i,j)} = \{\xi_1, \xi_2, \dots, \xi_i \cup \xi_j, \dots, \xi_{D_n}\}$$

es obtenido al tiempo $n_0 - (n + 1)$ por la unión de las clases de equivalencia ξ_i y ξ_j de $\bar{\xi}$. Luego,

$$R_{n+1} = R_n^{(i,j)} = \{\bar{\xi}^{(i,j)}; \bar{\eta}\},$$

y además $D_{n+1} = D_n - 1$ y $F_{n+1} = F_n$.

(B) *Conversión de una clase antigua en una clase nueva*

Supóngase que $R_n = \{\xi_1, \xi_2, \dots, \xi_{D_n}; \eta_1, \eta_2, \dots, \eta_{F_n}\} = \{\bar{\xi}; \bar{\eta}\}$. Sea $\xi_i \in \bar{\xi}$ una clase antigua al tiempo n . Supóngase que los individuos de esta clase (presentes al tiempo $n_0 - n$) son descendientes mutantes de un individuo al tiempo $n_0 - (n + 1)$. Entonces, la clase antigua ξ_i es reclasificada como clase nueva al tiempo $n + 1$. Es decir, si:

$$\bar{\xi}^{(i)} = \{\xi_1, \xi_2, \dots, \xi_{i-1}, \xi_{i+1}, \dots, \xi_{D_n}\}$$

$$\bar{\eta}^{(i)} = \{\eta_1, \dots, \eta_{F_n}, \xi_i\},$$

entonces $\{\bar{\xi}^{(i)}; \bar{\eta}^{(i)}\}$ es obtenida al tiempo $n_0 - (n + 1)$ a partir de $\{\bar{\xi}; \bar{\eta}\}$, eliminando la clase ξ_i de $\bar{\xi}$ y añadiéndola a $\bar{\eta}$. De este modo,

$$R_{n+1} = R_n^{(i)} = \{\bar{\xi}^{(i)}; \bar{\eta}^{(i)}\}$$

y además $D_{n+1} = D_n - 1$ y $F_{n+1} = F_n + 1$.

(C) *No hay cambio*

Supóngase que $R_n = \{\xi_1, \xi_2, \dots, \xi_{D_n}; \eta_1, \eta_2, \dots, \eta_{F_n}\} = \{\bar{\xi}; \bar{\eta}\}$. Al no haber cambio, $R_{n+1} = R_n$ y además $D_{n+1} = D_n$, $F_{n+1} = F_n$.

Y dado que al tiempo n_0 todos los individuos son descendientes de ellos mismos y sin mutación, se tiene que en n_0 todas las clases son antiguas y, por tanto, $R_0 = \{\xi_1, \xi_2, \dots, \xi_m; \emptyset\}$ con $\xi_i = (i)$, $i = 1, 2, \dots, m$.

Por lo anterior, el proceso R es una cadena de Markov con espacio de estados el conjunto de todas las particiones de la muestra que son del tipo $\{\bar{\xi}; \bar{\eta}\}$ (donde $\bar{\xi}$ o $\bar{\eta}$ puede ser vacía).

Definición 56. Se define la *cadena de saltos* de R como $\mathcal{R} = \{\mathcal{R}_k : k = m, m - 1, \dots, 0\}$ tal que $\mathcal{R}_k = \{\xi_1, \dots, \xi_k; \eta_1, \dots, \eta_l\}$, donde: ξ_i es una clase antigua, $i = 1, \dots, k$, η_j es una clase nueva, $j = 1, \dots, l$ y $k = 0, 1, \dots, m$, $l = 0, 1, \dots, m - k$.

- Probabilidades de transición del proceso R

$$P\{R_{n+1} = R_n^{(i,j)} \mid R_n = \{\bar{\xi}; \bar{\eta}\}\} = \frac{2}{N^2}(1 - \gamma), \quad 1 \leq i < j \leq D_n \quad (4.2.2)$$

pues $2/N^2$ es la probabilidad de que los individuos i y j sean el padre y su descendiente (en cualquier orden) que sobreviven en la siguiente transición, y $1 - \gamma$ es la probabilidad de que el descendiente no sea un mutante.

$$P\{R_{n+1} = R_n^{(i)} \mid R_n = \{\bar{\xi}; \bar{\eta}\}\} = \frac{1}{N}\gamma, \quad i = 1, 2, \dots, D_n \quad (4.2.3)$$

pues $1/N$ es la probabilidad de elegir un individuo en la población y γ es la probabilidad de que dicho individuo mute.

$$P\{R_{n+1} = \{\bar{\xi}; \bar{\eta}\} \mid R_n = \{\bar{\xi}; \bar{\eta}\}\} = 1 - \binom{D_n}{2} \frac{2}{N^2}(1 - \gamma) - \binom{D_n}{1} \frac{1}{N}\gamma \quad (4.2.4)$$

pues la probabilidad de que haya un cambio de estado es $\binom{D_n}{2}(2(1 - \gamma)/N^2) + \binom{D_n}{1}(\gamma/N)$, que corresponde a todas las posibles maneras de elegir 2 clases antiguas para unirse, por la probabilidad de que se unan, más todas las posibles maneras de elegir una clase antigua que se transforme en nueva, por la probabilidad de que la transformación ocurra.

4.2.3 Teorema principal

El teorema presentado a continuación permite obtener la probabilidad de que exista al tiempo n una cierta partición de la muestra con un determinado número de clases de equivalencia nuevas, dado que a ese tiempo se tiene cierto número de clases de equivalencia antiguas en la muestra, para $n \in \{0, 1, \dots\}$.

En el resto de esta sección se utilizará la siguiente notación:

$$\begin{aligned} x_{(k)} &= x(x+1) \cdots (x+k-1) && \text{para } k \geq 0 \\ x_{[k]} &= x(x-1) \cdots (x-k+1) && \text{para } 0 \leq k \leq x+1 \\ x_{(0)} &= 1 = x_{[0]} \end{aligned}$$

Teorema 16. Sean $\theta = N\gamma/(1-\gamma)$, $k = 0, 1, 2, \dots, m$ y $l = 0, 1, \dots, m-k$. Entonces, dado R_0 ,

$$P\{F_n = l, R_n = \{\bar{\xi}; \bar{\eta}\} \mid D_n = k\} = \frac{(m-k)! k! \theta^l}{m! (k+\theta)_{(m-k)}} \lambda_1! \lambda_2! \dots \lambda_k! (\mu_1 - 1)! (\mu_2 - 1)! \dots (\mu_l - 1)! \quad (4.2.5)$$

para $n \in \{0, 1, \dots\}$ y donde m satisface (4.2.1).

Antes de demostrar el Teorema 16 se harán algunas observaciones:

- Caso sin mutación

Cuando no existe mutación, se tiene de los resultados del proceso n-coalescente de Kingman, que las probabilidades de transición de R_n a R_{n+1} , dadas por (4.2.2), (4.2.3) y (4.2.4), sólo dependen del número de clases antiguas presentes D_n . La cadena de saltos $\mathcal{R} = \{\mathcal{R}_k : k = m, m-1, \dots, 1, 0\}$ del proceso R es tal que si $D_n = k$ y $\bar{R} = \{\bar{\xi}; \bar{\eta}\}$, se tiene que:

$$P\{R_n = \bar{R}\} = P\{\mathcal{R}_k = \bar{R}\} P\{D_n = k\},$$

(ver [25] y [16]).

- Caso con mutación

En este caso, la cadena de saltos del proceso R alcanza un estado absorbente $\mathcal{R}_0 = \{\emptyset; \eta_1, \eta_2, \dots, \eta_l\}$ sin clases antiguas y con l clases nuevas, para alguna $l = 1, 2, \dots, m$. Dicho estado es absorbente pues las únicas transiciones del proceso R donde hay cambio son las tipo (A) y (B), que dependen de la existencia de clases antiguas en la muestra.

Demostración del Teorema 16. Sea $\bar{R} = \{\xi_1, \xi_2, \dots, \xi_k; \eta_1, \eta_2, \dots, \eta_l\} = \{\bar{\xi}; \bar{\eta}\}$ una partición de la muestra de m individuos en k clases antiguas y l clases nuevas. La prueba se hará por inducción:

Tómese $k = m$, por lo que $\mathcal{R}_m = R_0$ y \mathcal{R}_m consta sólo de clases antiguas con un único individuo y ninguna clase nueva; es decir, $\lambda_1 = \lambda_2 = \dots = \lambda_k = 1$, $l = 0$ y $\mathcal{R}_m = \{(1), (2), \dots, (m); \emptyset\}$ con probabilidad 1. Y obsérvese que:

$$P\{F_0 = 0, \mathcal{R}_m = \{(1), (2), \dots, (m); \emptyset\} \mid D_0 = m\} = \frac{(m-m)! m!}{m! (m+\theta)_{(m-m)}} = 1,$$

por lo que el resultado vale para $k = m$.

Supóngase que el resultado es cierto para $k + 1$, y se demostrará que es válido para k . Nótese que:

$$P\{\mathcal{R}_k = \bar{R}\} = \sum_{R'} P\{\mathcal{R}_{k+1} = R'\} P\{\mathcal{R}_k = \bar{R} \mid \mathcal{R}_{k+1} = R'\}.$$

La probabilidad de transición $P\{\mathcal{R}_k = \bar{R} \mid \mathcal{R}_{k+1} = R'\}$ es distinta de cero cuando uno de los cambios (A) o (B), dados en la Sección 4.2.2, ocurren en el proceso R (el caso (C) no se considera dado que se está tratando con la cadena de saltos asociada a R).

• *Caso 1*

Este caso corresponde al cambio tipo (A). Sea:

$$R' = \{\xi_1, \xi_2, \dots, \xi_{i1}, \xi_{i2}, \dots, \xi_k; \eta_1, \eta_2, \dots, \eta_l\},$$

donde ξ_{i1} y ξ_{i2} son dos clases antiguas tales que $\xi_{i1}, \xi_{i2} \neq \emptyset$ y $\xi_{i1} \cup \xi_{i2} = \xi_i \in \bar{\xi}$, $|\xi_{i1}| = \lambda_{i1}$, $|\xi_{i2}| = \lambda_{i2}$, $\lambda_{i1} + \lambda_{i2} = \lambda_i = |\xi_i|$.

Obsérvese que existen $\binom{\lambda_i}{\lambda_{i1}}$ maneras en las que la clase ξ_i puede partirse en dos subconjuntos no vacíos de tamaño λ_{i1} y $\lambda_i - \lambda_{i1}$, y λ_{i1} puede variar entre 1 y $\lambda_i - 1$. Además, i puede tomar cualquier valor en $\{1, 2, \dots, k\}$.

Ahora,

$$P\{\xi_{i1} \text{ y } \xi_{i2} \text{ se unan para formar una sola clase antigua}\} = \frac{2}{N^2}(1-\gamma) \quad (4.2.6)$$

$$\begin{aligned} P\{\text{haya cambio}\} &= \binom{k+1}{2} \frac{2}{N^2}(1-\gamma) + \binom{k+1}{1} \frac{1}{N}\gamma \\ &= \frac{k(k+1)(1-\gamma)}{N^2} + \frac{(k+1)\gamma}{N} \\ &= \frac{(k+1)(1-\gamma)}{N^2} \left(k + \frac{N\gamma}{1-\gamma} \right) \\ &= \frac{(k+1)(1-\gamma)(k+\theta)}{N^2} \end{aligned} \quad (4.2.7)$$

Entonces, como $P\{\mathcal{R}_k = \bar{R} \mid \mathcal{R}_{k+1} = R'\}$ es igual a (4.2.6) entre (4.2.7), se tiene:

$$P\{\mathcal{R}_k = \bar{R} \mid \mathcal{R}_{k+1} = R'\} = \frac{2(1-\gamma)/N^2}{(k+1)(1-\gamma)(k+\theta)/N^2} = \frac{2}{(k+1)(k+\theta)}.$$

Por otro lado, $P\{\mathcal{R}_{k+1} = R'\}$ está dada por la hipótesis de inducción. Entonces, la contribución del Caso 1 a la $P\{\mathcal{R}_k = \bar{R}\}$ es:

$$\begin{aligned} & \sum_{i=1}^k \sum_{\lambda_{i1}=1}^{\lambda_i-1} \frac{1}{2} \binom{\lambda_i}{\lambda_{i1}} \frac{(m-k-1)!(k+1)!}{m!(k+1+\theta)_{(m-k-1)}} \theta^l \lambda_1! \lambda_2! \cdots \lambda_{i1}! \lambda_{i2}! \cdots \lambda_k! \\ & \qquad \qquad \qquad \left(\prod_{j=1}^l (\mu_j - 1)! \right) \frac{2}{(k+1)(k+\theta)} = \\ = & \sum_{i=1}^k \sum_{\lambda_{i1}=1}^{\lambda_i-1} \frac{\lambda_i!}{(\lambda_i - \lambda_{i1})! \lambda_{i1}!} \frac{(m-k-1)!(k+1)k!}{m!(k+1+\theta)_{(m-k-1)}} \frac{1}{(k+1)(k+\theta)} \theta^l \\ & \qquad \qquad \qquad \lambda_1! \cdots \lambda_{i-1}! \lambda_{i1}! \lambda_{i2}! \lambda_{i+1}! \cdots \lambda_k! \left(\prod_{j=1}^l (\mu_j - 1)! \right) \quad (4.2.8) \end{aligned}$$

Utilizando la notación para $x_{(k)}$ se tiene que (4.2.8) es igual a:

$$\begin{aligned} & \sum_{i=1}^k \sum_{\lambda_{i1}=1}^{\lambda_i-1} \frac{(m-k-1)!k!}{m!(k+\theta)_{(m-k)}} \theta^l \lambda_1! \lambda_2! \cdots \lambda_k! \left(\prod_{j=1}^l (\mu_j - 1)! \right) = \\ & = \frac{(m-k-1)!k!}{m!(k+\theta)_{(m-k)}} \theta^l \left(\prod_{j=1}^k \lambda_j! \right) \left(\prod_{j=1}^l (\mu_j - 1)! \right) \sum_{i=1}^k (\lambda_i - 1) \quad (4.2.9) \end{aligned}$$

• *Caso 2*

Este caso corresponde al cambio tipo (B). Sea:

$$R' = \{\xi_1, \xi_2, \dots, \xi_k, \eta_j; \eta_1, \eta_2, \dots, \eta_{j-1}, \eta_{j+1}, \dots, \eta_l\},$$

donde η_j es la clase antigua que se transformará en una clase nueva. Entonces, \bar{R} es obtenido a partir de R' cuando la clase η_j se reclasifica como clase nueva. Por la ecuación (4.2.3),

$$P\{\text{una clase antigua } \eta_j \text{ se transforme en nueva}\} = \frac{1}{N} \gamma, \quad (4.2.10)$$

Como $P\{\mathcal{R}_k = \bar{R} \mid \mathcal{R}_{k+1} = R'\}$ es igual a la probabilidad (4.2.10) entre la

probabilidad de que haya cambio, dada por (4.2.7), se sigue que:

$$\begin{aligned} P\{\mathcal{R}_k = \bar{R} \mid \mathcal{R}_{k+1} = R'\} &= \frac{\gamma}{N} \frac{N^2}{(k+1)(1-\gamma)(k+\theta)} \\ &= \frac{\gamma N}{1-\gamma} \frac{1}{(k+1)(k+\theta)} \\ &= \frac{\theta}{(k+1)(k+\theta)} \end{aligned}$$

Entonces, la contribución del Caso 2 a $P\{\mathcal{R}_k = \bar{R}\}$ es:

$$\begin{aligned} &\sum_{j=1}^l \frac{(m-k-1)!(k+1)k!}{m!(k+1+\theta)_{(m-k-1)}} \theta^{l-1} \left(\prod_{j=1}^k \lambda_j! \right) \mu_j! (\mu_1 - 1)! \cdots (\mu_{j-1} - 1)! (\mu_{j+1} - 1)! \\ &\quad \cdots (\mu_l - 1)! \frac{\theta}{(k+1)(k+\theta)} = \\ &= \sum_{j=1}^l \frac{(m-k-1)k!}{m!(k+\theta)_{(m-k)}} \theta^l \left(\prod_{j=1}^k \lambda_j! \right) \left(\prod_{j=1}^l (\mu_j - 1)! \right) \mu_j \\ &= \frac{(m-k-1)k!}{m!(k+\theta)_{(m-k)}} \theta^l \left(\prod_{j=1}^k \lambda_j! \right) \left(\prod_{j=1}^l (\mu_j - 1)! \right) \sum_{j=1}^l \mu_j \quad (4.2.11) \end{aligned}$$

Sumando las contribuciones (4.2.9) y (4.2.11), de los casos 1 y 2, respectivamente, se tiene que:

$$\begin{aligned} P\{\mathcal{R}_k = \bar{R}\} &= \frac{(m-k-1)k!}{m!(k+\theta)_{(m-k)}} \theta^l \left(\prod_{j=1}^k \lambda_j! \right) \left(\prod_{j=1}^l (\mu_j - 1)! \right) \left(\sum_{i=1}^k (\lambda_i - 1) + \sum_{j=1}^l \mu_j \right) \\ &= \frac{(m-k-1)k!}{m!(k+\theta)_{(m-k)}} \theta^l \left(\prod_{j=1}^k \lambda_j! \right) \left(\prod_{j=1}^l (\mu_j - 1)! \right) \left(\sum_{i=1}^k \lambda_i + \sum_{j=1}^l \mu_j - k \right) \\ &= \frac{(m-k)(m-k-1)k!}{m!(k+\theta)_{(m-k)}} \theta^l \left(\prod_{j=1}^k \lambda_j! \right) \left(\prod_{j=1}^l (\mu_j - 1)! \right) \quad (\text{por (4.2.1)}) \end{aligned}$$

y el resultado se sigue. \square

Observación 40. (a) Nótese que $P\{D_n = k\} = 0$ si $k < m - n$ ($m \geq n$), porque a lo más n clases antiguas pueden perderse hasta el tiempo n ,

por lo que a este tiempo quedan exactamente $m - n$ clases antiguas, no menos.

- (b) Caso particular del Teorema 16: distribución del m -coalescente de Kingman. Cuando no hay mutación, la muestra sólo puede tener clases antiguas en cualquier tiempo n . Luego, $F_n = 0$ para toda $n \in \{0, 1, \dots\}$, $l = 0$ y $\theta = 0$ (pues $\gamma = 0$). Entonces (4.2.5) se reduce a:

$$\begin{aligned} P\{R_n = \{\bar{\xi}; \bar{\eta}\} \mid D_n = k\} &= P\{F_n = 0, R_n = \{\bar{\xi}; \bar{\eta}\} \mid D_n = k\} \\ &= \frac{(m-k)! k!}{m! k_{(m-k)}} \lambda_1! \lambda_2! \cdots \lambda_k! \end{aligned}$$

Y como $k_{(m-k)} = k(k+1) \cdots (k+m-k-1) = (m-1)!/(k-1)!$, entonces,

$$P\{R_n = \{\bar{\xi}; \bar{\eta}\} \mid D_n = k\} = \frac{(m-k)! k! (k-1)!}{m! (m-1)!} \lambda_1! \lambda_2! \cdots \lambda_k!$$

que corresponde a la distribución $P\{\mathcal{R}_k = \bar{\xi}\}$ del m -coalescente de Kingman.

Apéndice A

Para mayores detalles de los resultados presentados en este Apéndice, consultar [2], [3], [10], [22] y [24].

Definición 57. Un subconjunto I de \mathbb{R} de la forma $[l, r]$, $l < r$, se denomina un *intervalo compacto* en \mathbb{R} .

Definición 58. Sea $\{f_n : n = 1, 2, \dots\}$ una secuencia de funciones definidas en un conjunto E . Se dice que $\{f_n : n = 1, 2, \dots\}$ converge *uniformemente* en E a una función f si para toda $\varepsilon > 0$ existe un entero N tal que $n \geq N$ implica

$$|f_n(x) - f(x)| \leq \varepsilon$$

para toda $x \in E$.

Definición 59. Sea M un conjunto y sean $p, q, \in M$. La función $d : M \times M \rightarrow [0, \infty)$ que satisface:

- (a) $d(p, q) > 0$ si $p \neq q$
- (b) $d(p, q) = 0$ si y sólo si $p = q$
- (c) $d(p, q) = d(q, p)$
- (d) $d(p, q) \leq d(p, r) + d(r, q)$ para toda $r \in M$,

es una *métrica* sobre M .

Definición 60. Sea M un conjunto en el que es posible definir una métrica d . Se dice entonces que M es un *espacio métrico*, y usualmente se denota por M o (M, d) .

Definición 61. Sean M un espacio métrico con métrica d y $E \subset M$.

Una *vecindad* de radio $r > 0$ de un punto $p \in M$ es un conjunto $B_r(p)$ de puntos $q \in M$ tales que $d(q, p) < r$.

Se dice que E es un *conjunto abierto* si para todo $p \in E$ existe $r > 0$ tal que $B_r(p) \subset E$.

Una *cubierta abierta* de E es una colección $\{G_\alpha\}$ de conjuntos abiertos en M tal que $E \subset \bigcup_\alpha G_\alpha$.

Definición 62. Sea M un espacio métrico y K un subconjunto de M . Sea $\{G_\alpha\}$ una cubierta abierta de K . Se dice que K es *compacto* si existe un número finito de subíndices $\alpha_1, \dots, \alpha_n$ tales que:

$$K \subset G_{\alpha_1} \cup \dots \cup G_{\alpha_n}.$$

Definición 63. Sean M un espacio métrico y $E \subset M$.

Se dice que $p \in E$ es un *punto límite* de E si toda vecindad de p contiene un punto $q \neq p$ tal que $q \in E$.

Se dice que E es *denso en M* si todo punto de M es un punto límite de E , o un punto de E (o ambos).

Definición 64. Sea M un espacio métrico. Se dice que M es *separable* si contiene un subconjunto denso numerable.

Definición 65. Sean $\{X_n : n = 1, 2, \dots\}$ un conjunto de variables aleatorias y X una variable aleatoria. Se dice que X_n *converge en distribución (o débilmente)* a X si

$$P\{X_n \leq x\} \rightarrow P\{X \leq x\}$$

conforme $n \rightarrow \infty$, para toda x tal que $F_X(x) = P\{X \leq x\}$ es continua.

Se denotará por $\mathcal{C}(T, \mathbb{R}^d)$ al espacio de las funciones continuas definidas en T con valores en \mathbb{R}^d , y por $Pr(\mathcal{C}(T, \mathbb{R}^d))$ al conjunto de todas las funciones de probabilidad definidas en $\mathcal{C}(T, \mathbb{R}^d)$.

Definición 66. Una familia $\mathcal{A} \subset Pr(\mathcal{C}(T, \mathbb{R}^d))$ es *uniformemente tensa* si para toda $\varepsilon > 0$ existe un conjunto compacto $K_\varepsilon \subset \mathcal{C}(T, \mathbb{R}^d)$ tal que:

$$P\{K_\varepsilon\} \leq \varepsilon$$

para toda $P \in \mathcal{A}$.

La familia \mathcal{A} es *relativamente compacta* si para cada sucesión $\{\theta_n : n =$

$1, 2, \dots\} \subset \mathcal{A}$ existen una subsucesión $\{\theta_{n_k} : k = 1, 2, \dots\}$ y una probabilidad $\theta \in Pr(\mathcal{C}(T, \mathbb{R}^d))$ (no necesariamente en \mathcal{A}), tales que para $f \in \mathcal{C}(T, \mathbb{R}^d)$ acotada se tiene:

$$\lim_{n \rightarrow \infty} \int f d\theta_n = \int f d\theta$$

que se denotará por $\theta_{n_k} \implies \theta$ cuando $k \rightarrow \infty$ (para mayores detalles, ver [3]).

Teorema 17 (Teorema de Prohorov). *Una familia $\mathcal{A} \subset Pr(\mathcal{C}(T, \mathbb{R}^d))$ es relativamente compacta si y sólo si \mathcal{A} es uniformemente tensa.*

Demostración. Ver [3]. □

Definición 67. Una familia $\{X^{(n)} : n = 1, 2, \dots\}$ de procesos estocásticos con distribuciones $\{P_{(n)} : n = 1, 2, \dots\}$, respectivamente, es relativamente compacta si $\{P_{(n)} : n = 1, 2, \dots\}$ es relativamente compacta.

Teorema 18. *Sea (E, d) un espacio métrico separable. Sean $X, X^{(n)}, n = 1, 2, \dots$ procesos estocásticos definidos en un espacio de probabilidad (Ω, \mathcal{F}, P) , que toman valores en (E, d) y tales que para cada $\omega \in \Omega, X_t(\omega)$ y $X_t^{(n)}(\omega), t \in [0, \infty)$, son funciones continuas por la derecha con límite por la izquierda. Entonces, si $\{X^{(n)} : n = 1, 2, \dots\}$ es relativamente compacta y existe un conjunto denso $D \subset [0, \infty)$ tal que para todo conjunto finito $\{t_1, t_2, \dots, t_k\} \subset D$ es válido*

$$(X_{t_1}^{(n)}, X_{t_2}^{(n)}, \dots, X_{t_k}^{(n)}) \implies (X_{t_1}, X_{t_2}, \dots, X_{t_k})$$

cuando $n \rightarrow \infty$, entonces $X^{(n)} \implies X$ cuando $n \rightarrow \infty$.

1. **Objetivo:** El propósito de este capítulo es proporcionar una visión general de los conceptos básicos de la programación en C++.

2. **Contenido:** Este capítulo cubre los fundamentos de la programación en C++, incluyendo la sintaxis, los tipos de datos y las estructuras de control.

3. **Objetivo:** El propósito de este capítulo es introducir al lector en el uso de las funciones y el manejo de archivos en C++.

4. **Contenido:** Este capítulo detalla cómo declarar y utilizar funciones, así como las técnicas para leer y escribir datos en archivos.

5. **Objetivo:** El propósito de este capítulo es enseñar al lector sobre las estructuras de datos y cómo utilizarlas en C++.

6. **Contenido:** Este capítulo explora los arreglos, los vectores, las listas enlazadas y otros tipos de estructuras de datos.

7. **Objetivo:** El propósito de este capítulo es introducir al lector en el uso de las clases y la programación orientada a objetos en C++.

8. **Contenido:** Este capítulo cubre los conceptos de clases, objetos, herencia y polimorfismo.

9. **Objetivo:** El propósito de este capítulo es enseñar al lector sobre el uso de las bibliotecas estándar y cómo utilizarlas en C++.

10. **Contenido:** Este capítulo detalla el uso de las bibliotecas estándar de C++, incluyendo las funciones de entrada y salida.

11. **Objetivo:** El propósito de este capítulo es introducir al lector en el uso de las bibliotecas de terceros y cómo utilizarlas en C++.

12. **Contenido:** Este capítulo cubre el uso de bibliotecas de terceros, como las bibliotecas de matemáticas y de gráficos.

13. **Objetivo:** El propósito de este capítulo es enseñar al lector sobre el uso de las bibliotecas de programación y cómo utilizarlas en C++.

14. **Contenido:** Este capítulo cubre el uso de bibliotecas de programación, como las bibliotecas de programación de gráficos y de programación de juegos.

15. **Objetivo:** El propósito de este capítulo es introducir al lector en el uso de las bibliotecas de programación de gráficos y de programación de juegos en C++.

16. **Contenido:** Este capítulo cubre el uso de bibliotecas de programación de gráficos y de programación de juegos, como las bibliotecas de OpenGL y de SDL.

17. **Objetivo:** El propósito de este capítulo es enseñar al lector sobre el uso de las bibliotecas de programación de gráficos y de programación de juegos en C++.

18. **Contenido:** Este capítulo cubre el uso de bibliotecas de programación de gráficos y de programación de juegos, como las bibliotecas de OpenGL y de SDL.

19. **Objetivo:** El propósito de este capítulo es introducir al lector en el uso de las bibliotecas de programación de gráficos y de programación de juegos en C++.

20. **Contenido:** Este capítulo cubre el uso de bibliotecas de programación de gráficos y de programación de juegos, como las bibliotecas de OpenGL y de SDL.

Apéndice B

Proposición 7. Sean $\gamma_1, \gamma_2 \geq 0$ tales que $\gamma_1 + \gamma_2 < 1$, N un entero mayor que cero y $k = 1, 2, \dots$ Entonces se satisface:

$$\frac{\Gamma\left(\frac{N\gamma_2}{1-\gamma_1-\gamma_2} + k\right)}{\Gamma\left(\frac{N\gamma_2}{1-\gamma_1-\gamma_2}\right)} = \prod_{s=0}^{k-1} \left(\frac{N\gamma_2 + s(1-\gamma_1-\gamma_2)}{1-\gamma_1-\gamma_2}\right) \quad (\text{B.0.1})$$

$$\frac{\Gamma\left(\frac{N(1-\gamma_2)}{1-\gamma_1-\gamma_2}\right)}{\Gamma\left(\frac{N(1-\gamma_2)}{1-\gamma_1-\gamma_2} - k\right)} = \prod_{s=1}^k \left(\frac{N(1-\gamma_2) - s(1-\gamma_1-\gamma_2)}{1-\gamma_1-\gamma_2}\right) \quad (\text{B.0.2})$$

Demostración. Como la función Gamma satisface $\Gamma(\alpha + 1) = \alpha\Gamma(\alpha)$ para toda $\alpha > 0$ (ver [9]), entonces:

$$\begin{aligned} \Gamma\left(\frac{N\gamma_2}{1-\gamma_1-\gamma_2} + k\right) &= \left(\frac{N\gamma_2}{1-\gamma_1-\gamma_2} + k - 1\right) \Gamma\left(\frac{N\gamma_2}{1-\gamma_1-\gamma_2} + k - 1\right) \\ &= \left(\frac{N\gamma_2}{1-\gamma_1-\gamma_2} + k - 1\right) \Gamma\left(\left(\frac{N\gamma_2}{1-\gamma_1-\gamma_2} + k - 2\right) + 1\right) \\ &= \left(\frac{N\gamma_2}{1-\gamma_1-\gamma_2} + k - 1\right) \left(\frac{N\gamma_2}{1-\gamma_1-\gamma_2} + k - 2\right) \Gamma\left(\frac{N\gamma_2}{1-\gamma_1-\gamma_2} + k - 2\right) \\ &= \left(\frac{N\gamma_2}{1-\gamma_1-\gamma_2} + k - 1\right) \left(\frac{N\gamma_2}{1-\gamma_1-\gamma_2} + k - 2\right) \cdots \left(\frac{N\gamma_2}{1-\gamma_1-\gamma_2} + k - (k - 1)\right) \\ &\quad \left(\frac{N\gamma_2}{1-\gamma_1-\gamma_2} + k - k\right) \Gamma\left(\frac{N\gamma_2}{1-\gamma_1-\gamma_2} + k - k\right) \\ &= \left(\frac{N\gamma_2}{1-\gamma_1-\gamma_2} + k - 1\right) \left(\frac{N\gamma_2}{1-\gamma_1-\gamma_2} + k - 2\right) \cdots \left(\frac{N\gamma_2}{1-\gamma_1-\gamma_2} + 1\right) \\ &\quad \left(\frac{N\gamma_2}{1-\gamma_1-\gamma_2}\right) \Gamma\left(\frac{N\gamma_2}{1-\gamma_1-\gamma_2}\right) \end{aligned}$$

Por tanto,

$$\begin{aligned} \frac{\Gamma\left(\frac{N\gamma_2}{1-\gamma_1-\gamma_2} + k\right)}{\Gamma\left(\frac{N\gamma_2}{1-\gamma_1-\gamma_2}\right)} &= \left(\frac{N\gamma_2}{1-\gamma_1-\gamma_2} + k - 1\right) \left(\frac{N\gamma_2}{1-\gamma_1-\gamma_2} + k - 2\right) \cdots \\ &\quad \left(\frac{N\gamma_2}{1-\gamma_1-\gamma_2} + 1\right) \left(\frac{N\gamma_2}{1-\gamma_1-\gamma_2}\right) \\ &= \left(\frac{N\gamma_2 + (k-1)(1-\gamma_1-\gamma_2)}{1-\gamma_1-\gamma_2}\right) \left(\frac{N\gamma_2 + (k-2)(1-\gamma_1-\gamma_2)}{1-\gamma_1-\gamma_2}\right) \\ &\quad \cdots \left(\frac{N\gamma_2 + (1-\gamma_1-\gamma_2)}{1-\gamma_1-\gamma_2}\right) \left(\frac{N\gamma_2}{1-\gamma_1-\gamma_2}\right) \\ &= \prod_{s=0}^{k-1} \left(\frac{N\gamma_2 + s(1-\gamma_1-\gamma_2)}{1-\gamma_1-\gamma_2}\right), \end{aligned}$$

y queda así demostrada la igualdad (B.0.1).

Para probar la igualdad (B.0.2) obsérvese lo siguiente:

$$\Gamma\left(\frac{N(1-\gamma_2)}{1-\gamma_1-\gamma_2} - k + 1\right) = \left(\frac{N(1-\gamma_2)}{1-\gamma_1-\gamma_2} - k\right) \Gamma\left(\frac{N(1-\gamma_2)}{1-\gamma_1-\gamma_2} - k\right).$$

Y como

$$\Gamma\left(\frac{N(1-\gamma_2)}{1-\gamma_1-\gamma_2} - k + 2\right) = \left(\frac{N(1-\gamma_2)}{1-\gamma_1-\gamma_2} - k + 1\right) \Gamma\left(\frac{N(1-\gamma_2)}{1-\gamma_1-\gamma_2} - k + 1\right),$$

entonces:

$$\begin{aligned} \Gamma\left(\frac{N(1-\gamma_2)}{1-\gamma_1-\gamma_2} - k + 2\right) &= \left(\frac{N(1-\gamma_2)}{1-\gamma_1-\gamma_2} - k + 1\right) \left(\frac{N(1-\gamma_2)}{1-\gamma_1-\gamma_2} - k\right) \\ &\quad \Gamma\left(\frac{N(1-\gamma_2)}{1-\gamma_1-\gamma_2} - k\right). \end{aligned}$$

Ahora, como

$$\Gamma\left(\frac{N(1-\gamma_2)}{1-\gamma_1-\gamma_2} - k + 3\right) = \left(\frac{N(1-\gamma_2)}{1-\gamma_1-\gamma_2} - k + 2\right) \Gamma\left(\frac{N(1-\gamma_2)}{1-\gamma_1-\gamma_2} - k + 2\right),$$

entonces:

$$\begin{aligned} \Gamma\left(\frac{N(1-\gamma_2)}{1-\gamma_1-\gamma_2} - k + 3\right) &= \left(\frac{N(1-\gamma_2)}{1-\gamma_1-\gamma_2} - k + 2\right) \left(\frac{N(1-\gamma_2)}{1-\gamma_1-\gamma_2} - k + 1\right) \\ &\quad \left(\frac{N(1-\gamma_2)}{1-\gamma_1-\gamma_2} - k\right) \Gamma\left(\frac{N(1-\gamma_2)}{1-\gamma_1-\gamma_2} - k\right). \end{aligned}$$

Así sucesivamente hasta el paso $k - 1$ se tiene:

$$\begin{aligned} \Gamma\left(\frac{N(1-\gamma_2)}{1-\gamma_1-\gamma_2} - 1\right) &= \Gamma\left(\frac{N(1-\gamma_2)}{1-\gamma_1-\gamma_2} - k + (k-1)\right) \\ &= \left(\frac{N(1-\gamma_2)}{1-\gamma_1-\gamma_2} - k + (k+2)\right) \cdots \left(\frac{N(1-\gamma_2)}{1-\gamma_1-\gamma_2} - k + 1\right) \\ &\quad \left(\frac{N(1-\gamma_2)}{1-\gamma_1-\gamma_2} - k\right) \Gamma\left(\frac{N(1-\gamma_2)}{1-\gamma_1-\gamma_2} - k\right) \\ &= \prod_{s=2}^k \left(\frac{N(1-\gamma_2)}{1-\gamma_1-\gamma_2} - s\right) \Gamma\left(\frac{N(1-\gamma_2)}{1-\gamma_1-\gamma_2} - k\right). \end{aligned}$$

Entonces:

$$\begin{aligned} \left(\frac{N(1-\gamma_2)}{1-\gamma_1-\gamma_2} - 1\right) \Gamma\left(\frac{N(1-\gamma_2)}{1-\gamma_1-\gamma_2} - 1\right) &= \\ &= \left(\frac{N(1-\gamma_2)}{1-\gamma_1-\gamma_2} - 1\right) \prod_{s=2}^k \left(\frac{N(1-\gamma_2)}{1-\gamma_1-\gamma_2} - s\right) \Gamma\left(\frac{N(1-\gamma_2)}{1-\gamma_1-\gamma_2} - k\right) \\ &= \prod_{s=1}^k \left(\frac{N(1-\gamma_2)}{1-\gamma_1-\gamma_2} - s\right) \Gamma\left(\frac{N(1-\gamma_2)}{1-\gamma_1-\gamma_2} - k\right). \end{aligned}$$

Y como

$$\Gamma\left(\frac{N(1-\gamma_2)}{1-\gamma_1-\gamma_2}\right) = \left(\frac{N(1-\gamma_2)}{1-\gamma_1-\gamma_2} - 1\right) \Gamma\left(\frac{N(1-\gamma_2)}{1-\gamma_1-\gamma_2} - 1\right),$$

entonces:

$$\Gamma\left(\frac{N(1-\gamma_2)}{1-\gamma_1-\gamma_2}\right) = \prod_{s=1}^k \left(\frac{N(1-\gamma_2)}{1-\gamma_1-\gamma_2} - s\right) \Gamma\left(\frac{N(1-\gamma_2)}{1-\gamma_1-\gamma_2} - k\right),$$

de donde se sigue que

$$\begin{aligned} \frac{\Gamma\left(\frac{N(1-\gamma_2)}{1-\gamma_1-\gamma_2}\right)}{\Gamma\left(\frac{N(1-\gamma_2)}{1-\gamma_1-\gamma_2} - k\right)} &= \prod_{s=1}^k \left(\frac{N(1-\gamma_2)}{1-\gamma_1-\gamma_2} - s\right) \\ &= \prod_{s=1}^k \left(\frac{N(1-\gamma_2) - s(1-\gamma_1-\gamma_2)}{1-\gamma_1-\gamma_2}\right) \end{aligned}$$

y queda así demostrada la igualdad (B.0.2). \square

Proposición 8. Sean $\gamma_1, \gamma_2 \geq 0$ tales que $\gamma_1 + \gamma_2 < 1$, N un entero mayor que cero y $k = 1, 2, \dots, N$. Entonces se satisface:

$$\left(\frac{N}{1 - \gamma_1 - \gamma_2} \right)^N \prod_{k=1}^N \left(1 - k \frac{(1 - \gamma_1 - \gamma_2)}{N} \right) = \frac{\Gamma \left(\frac{N}{1 - \gamma_1 - \gamma_2} \right)}{\Gamma \left(\frac{N(\gamma_1 + \gamma_2)}{1 - \gamma_1 - \gamma_2} \right)} \quad (\text{B.0.3})$$

$$\left(\frac{N}{1 - \gamma_1 - \gamma_2} \right)^N \prod_{k=1}^N \left((1 - \gamma_2) - k \frac{(1 - \gamma_1 - \gamma_2)}{N} \right) = \frac{\Gamma \left(\frac{N(1 - \gamma_2)}{1 - \gamma_1 - \gamma_2} \right)}{\Gamma \left(\frac{N\gamma_1}{1 - \gamma_1 - \gamma_2} \right)} \quad (\text{B.0.4})$$

Demostración. Utilizando el hecho de que $\Gamma(\alpha) = (\alpha - 1)\Gamma(\alpha - 1)$ para toda $\alpha > 1$ se tiene lo siguiente:

$$\begin{aligned} \Gamma \left(\frac{N}{1 - \gamma_1 - \gamma_2} \right) &= \left(\frac{N}{1 - \gamma_1 - \gamma_2} - 1 \right) \Gamma \left(\frac{N}{1 - \gamma_1 - \gamma_2} - 1 \right) \\ &= \left(\frac{N}{1 - \gamma_1 - \gamma_2} - 1 \right) \left(\frac{N}{1 - \gamma_1 - \gamma_2} - 2 \right) \Gamma \left(\frac{N}{1 - \gamma_1 - \gamma_2} - 2 \right) \\ &= \prod_{k=1}^N \left(\frac{N}{1 - \gamma_1 - \gamma_2} - k \right) \Gamma \left(\frac{N}{1 - \gamma_1 - \gamma_2} - N \right) \\ &= \prod_{k=1}^N \left(\frac{N}{1 - \gamma_1 - \gamma_2} - k \right) \Gamma \left(\frac{N - N(1 - \gamma_1 - \gamma_2)}{1 - \gamma_1 - \gamma_2} \right) \\ &= \prod_{k=1}^N \left(\frac{N}{1 - \gamma_1 - \gamma_2} - k \right) \Gamma \left(\frac{N(\gamma_1 + \gamma_2)}{1 - \gamma_1 - \gamma_2} \right). \end{aligned}$$

Entonces:

$$\frac{\Gamma \left(\frac{N}{1 - \gamma_1 - \gamma_2} \right)}{\Gamma \left(\frac{N(\gamma_1 + \gamma_2)}{1 - \gamma_1 - \gamma_2} \right)} = \prod_{k=1}^N \left(\frac{N}{1 - \gamma_1 - \gamma_2} - k \right).$$

Observación 41.

$$\begin{aligned} \text{(a)} \quad \prod_{k=1}^N \left(1 - k \frac{(1 - \gamma_1 - \gamma_2)}{N} \right) \left(\frac{N}{1 - \gamma_1 - \gamma_2} \right)^N &= \prod_{k=1}^N \left(\frac{N - k(1 - \gamma_1 - \gamma_2)}{1 - \gamma_1 - \gamma_2} \right) \\ &= \prod_{k=1}^N \left(\frac{N}{1 - \gamma_1 - \gamma_2} - k \right) \end{aligned}$$

y se sigue entonces la igualdad (B.0.3). Por otro lado,

$$\begin{aligned}
 \Gamma\left(\frac{N(1-\gamma_2)}{1-\gamma_1-\gamma_2}\right) &= \left(\frac{N(1-\gamma_2)}{1-\gamma_1-\gamma_2} - 1\right) \Gamma\left(\frac{N(1-\gamma_2)}{1-\gamma_1-\gamma_2} - 1\right) \\
 &= \left(\frac{N(1-\gamma_2)}{1-\gamma_1-\gamma_2} - 1\right) \left(\frac{N(1-\gamma_2)}{1-\gamma_1-\gamma_2} - 2\right) \Gamma\left(\frac{N(1-\gamma_2)}{1-\gamma_1-\gamma_2} - 2\right) \\
 &= \prod_{k=1}^N \left(\frac{N(1-\gamma_2)}{1-\gamma_1-\gamma_2} - k\right) \Gamma\left(\frac{N(1-\gamma_2)}{1-\gamma_1-\gamma_2} - N\right) \\
 &= \prod_{k=1}^N \left(\frac{N(1-\gamma_2)}{1-\gamma_1-\gamma_2} - k\right) \Gamma\left(\frac{N(1-\gamma_2) - N(1-\gamma_1-\gamma_2)}{N}\right) \\
 &= \prod_{k=1}^N \left(\frac{N(1-\gamma_2)}{1-\gamma_1-\gamma_2} - k\right) \Gamma\left(\frac{N\gamma_1}{1-\gamma_1-\gamma_2}\right).
 \end{aligned}$$

Entonces:

$$\frac{\Gamma\left(\frac{N(1-\gamma_2)}{1-\gamma_1-\gamma_2}\right)}{\Gamma\left(\frac{N\gamma_1}{1-\gamma_1-\gamma_2}\right)} = \prod_{k=1}^N \left(\frac{N(1-\gamma_2)}{1-\gamma_1-\gamma_2} - k\right).$$

$$\begin{aligned}
 \text{(b)} \quad & \prod_{k=1}^N \left((1-\gamma_2) - k \frac{(1-\gamma_1-\gamma_2)}{N} \right) \left(\frac{N}{1-\gamma_1-\gamma_2} \right)^N = \\
 &= \prod_{k=1}^N \left(\frac{N(1-\gamma_2) - k(1-\gamma_1-\gamma_2)}{1-\gamma_1-\gamma_2} \right) \\
 &= \prod_{k=1}^N \left(\frac{N(1-\gamma_2) - k}{1-\gamma_1-\gamma_2} \right)
 \end{aligned}$$

y se sigue entonces la igualdad (B.0.4).

□

Bibliografía

- [1] ASH, R.B. (1972), *Real Analysis and Probability*, EUA, Academic Press, Inc.
- [2] BARTLE, R.G. (1991), *Introducción al Análisis Matemático*, México, Ed. Limusa.
- [3] BILLINGSLEY, P. (1968), *Convergence of Probability Measures*, EUA, John Wiley & Sons.
- [4] BROWN, T.A. (1992), *Genetics: a molecular approach*, 2a. edición, EUA, Chapman and Hall.
- [5] DONNELLY, P. (1983), The Transient Behaviour of the Moran Model in Population Genetics, *Math. Proc. Cambridge Phil. Soc.*, **95**, p.349-350.
- [6] EWENS, W.J. (1979), *Mathematical Population Genetics*, EUA, Springer-Verlag.
- [7] EWENS, W.J. (1990), Population Genetics Theory - the past and the future, *Mathematical and Statistical Developments of Evolutionary Theory*, p.177-227.
- [8] FELLER, W. (1951), Diffusion Processes in Genetics, *Berkeley Symposium on Mathematics, Statistics and Probability*.
- [9] FELLER, W. (1973), *Introducción a la Teoría de Probabilidades y sus Aplicaciones*, vol. 1, México, Ed. Limusa.
- [10] FOLLAND, G.B. (1999), *Real Analysis, Modern Techniques and their Applications*, 2a. edición, EUA, John Wiley & Sons.

- [11] GRIMMET, G.R.; STIRZAKER, D.R. (1992), *Probability and Random Processes*, 2a. edición, EUA, Oxford Science Publications.
- [12] KARLIN, S.; MCGREGOR, J. (1962), On a Genetics Model of Moran, *Proc. Cambridge Phil. Soc.*, **58**, p.299-311.
- [13] KARLIN, S.; MCGREGOR, J. (1964), "On Some Stochastic Models in Genetics" en *Stochastic Models in Medicine and Biology*, EUA, The University of Wisconsin Press.
- [14] KARLIN, S.; TAYLOR, H. (1975), *A First Course in Stochastic Processes*, EUA, Academic Press, Inc.
- [15] KARLIN, S.; TAYLOR, H. (1981), *A Second Course in Stochastic Processes*, EUA, Academic Press, Inc.
- [16] KINGMAN, J.F.C. (1981), The Coalescent, *Stochastic Processes and their Applications*, **13**, p.235-248.
- [17] KINGMAN, J.F.C. (1982), On the Genealogy of Large Populations, *Journal of Applied Probability*, **19A**, p.27-43.
- [18] MORAN, P.A.P. (1957), Random Processes in Genetics, *Proc. Cambridge Phil. Soc.*, **54**, p.60-72.
- [19] OROZCO, M.I. (2001), *El uso del algoritmo EM para inferir secuencias de ADN*, Tesis de Licenciatura (UNAM), México.
- [20] RODRIGUES, E.R. (2001), *Modelos Estocásticos en Poblaciones Genéticas*, Notas de curso - Maestría.
- [21] ROSS, S. (1983), *Stochastic Processes*, EUA, John Wiley & Sons.
- [22] RUDIN, W. (1976), *Principles of Mathematical Analysis*, 3a. edición, EUA, McGraw-Hill.
- [23] SUSUKI, D.; KNUDTSON, P. (1990), *Genethics: the clash between the new genetics and human values*, EUA, Harvard University Press.
- [24] TUDOR, C. (1994) *Procesos estocásticos*, Colección Aportaciones Matemáticas, México, Sociedad Matemática Mexicana.

- [25] WATTERSON, G.A. (1983), Lines of Descent and the Coalescent, *Theoretical Population Biology*, **26**, p.77-92.