

00321



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

2

FACULTAD DE CIENCIAS

“Estimación de Líneas Ancestrales en Genética
utilizando el Coalescente de Kingman y
Métodos de Monte Carlo”

T E S I S

QUE PARA OBTENER EL TÍTULO DE:

A C T U A R I A

P R E S E N T A:

YESENIA ALFARO RAMÍREZ



DIRECTOR DE TESIS: ESTUDIOS R. RODRIGUES CALONI



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.

PAGINACION DISCONTINUA



UNIVERSIDAD NACIONAL
 AGENCIA DE
 MARAICÓN

DRA. MARÍA DE LOURDES ESTEVA PERALTA
Jefa de la División de Estudios Profesionales de la
Facultad de Ciencias
Presente

Comunicamos a usted que hemos revisado el trabajo escrito:

"Estimación de Líneas Ancestrales en Genética utilizando el Coalescente de Kingman y Métodos de Monte Carlo"

realizado por

Alfaro Ramírez Yesenia

con número de cuenta 09412126-3, quien cubrió los créditos de la carrera de: Actuaría

Dicho trabajo cuenta con nuestro voto aprobatorio.

Atentamente

Director de Tesis
 Propietario

Dra. Eliane R. Rodríguez Caloni

Eliane Regina Rodriguez

Propietario

Act. Jaime Vázquez Alamilla

Jaime Vázquez Alamilla

Propietario

Dr. Luis Antonio Rincón Solís

Luis Antonio Rincón Solís

Suplente

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

María del Pilar Alonso Reyes

Suplente

M. en C. José Antonio Flores Díaz

José Antonio Flores Díaz

Consejo Departamental de

Matemáticas

[Signature]
 FACULTAD DE CIENCIAS
 M. en C. JOSE ANTONIO FLORES DIAZ
 DE
 MATEMATICAS

" SI TENGO EL DON DE PROFECÍA, Y
PENETRO TODOS LOS MISTERIOS, Y
DOMINO TODA LA CIENCIA; SI TENGO UNA
FÉ TAN GRANDE, QUE AÚN CAMBIE
MONTAÑAS DE ACÁ PARA ALLÁ, PERO
SI NO TENGO AMOR, NO VALGO NADA"

I CORINTIOS 13, 2.

AGRADECIMIENTOS

A **DIOS** por la vida y el amor incondicional que me brinda en todo momento.

A mi **mamá**, por tu gran amor y porque nos has sabido aconsejar, cuidar y apoyar a lo largo de nuestra vida, a mis hermanas y a mí, dando lo mejor de ti. También te agradezco **papá** por el apoyo que nos das.

A mis hermanas:

Mari, porque siempre podemos contar contigo a pesar de que eres tan enojona.

Ara, por tu apoyo y porque siempre me contagias tu alegría.

Betino, por la comprensión y consejos que me das.

Claus, por tu confianza, y porque a pesar de que eres la más chica he aprendido mucho de ti.

A **Mirsa** y **Fernanda** porque són las nuevas y hermosas alegrías de la familia.

A mi tía **Concha**, mi tío **Ade**, y mis primos **Ana**, **Fenando**, **Mima** y **Caro**, porque a pesar de las adversidades, siempre contamos con su apoyo.

A mi asesora **Dra. Eliane**, por tu gran paciencia y nobleza.

A mis sinodales, que además también fueron mis profesores **Act. Jaime**, **M. en A. P. Pilar**, **Dr. Luis Antonio**, **M. en C. José Antonio**.

A mis amigos:

Frida, porque fuiste la primer persona en el CCH que me brindó su amistad y porque a pesar del tiempo seguimos siendo amigas.

Carlangas, por que eres un gran amigo, en quién toda la familia sabemos que podemos confiar y llamar y tener la seguridad de que siempre vas a estar ahí, apoyándonos como hasta ahora.

Pera, Ana y Mari, quienes con su apoyo, cariño y compañía, lograron buenos cambios en mí. También a **Dalia, Omar, Rodrigo, Omar Edgar, Mirra, Anabel, Lupita, Samuel, Héctor, Gonzalo, Charly, Pareja**.

Selene, Rosaura, Arminda, Ana Rosa, Erika, Olga, Toño, Gabriel, Liliana, Baltasar y Edgard, quienes compartieron conmigo sus ideales.

Alejandro, porque ya eres parte de la familia.

Álgebra, Maira, Enrique, Tonatihu, Gustavo, Eduardo, Karilia, Miguel Angel y Galo, a quienes a pesar de haber conocido al final de mi carrera, demostraron ser buenos amigos, y personas en quién confiar.

A **Hancy**, por la confianza, compañía y cariño que me brindas. Te Quiero Mucho.

A **Kika**, porque siempre me acompañabas por las mañanas, y me esperabas en la noche.

Al **Instituto de Matemáticas**, por la beca del lugar y a **PROBETEL**, por el apoyo económico para el desarrollo de mi tesis.

A **nuestra máxima casa de estudios, la UNAM**, en donde me formé, no solo profesionalmente, sino también humanamente, y aprendí a sustentar los conocimientos que actualmente poseo.

A mis profesores, por la formación académica que me otorgaron, en especial a **Javier Fernández, Rosa Maria T., Jesús López, Grabinsky, César Souza, César Cedillo, Carlos Mesa**.

Índice General

Introducción	iii
1 Cadenas de Markov	1
1.1 Resultados preliminares	1
1.2 Cadenas de Markov a tiempo discreto	4
1.2.1 Clasificación de estados	11
1.3 Cadenas de Markov a tiempo continuo	22
1.3.1 Ecuaciones diferenciales de Kolmogorov	29
1.3.2 Las ecuaciones diferenciales de Kolmogorov para proceso de nacimiento y muerte	33
1.3.3 Caso especial: Cadena de Markov de dos estados	35
1.3.4 Probabilidades límite	37
2 El modelo de Wright-Fisher y el modelo de Moran	45
2.1 Modelo de Wright-Fisher	46
2.2 Modelo de Moran	51
3 El proceso n-coalescente	65
3.1 El proceso del n -coalescente de Kingman y el modelo de Moran	66
4 La estadística bayesiana y el método de Monte Carlo	79
4.1 Introducción	79
4.2 Inferencia Bayesiana	83
4.3 Métodos de Monte Carlo	84
5 Inferencia ancestral	87
5.1 Introducción	87
5.2 Los datos de la tribu india Nuu-Chah-Nult	88

5.3	Modelo con un número infinito de sitios	90
5.4	Estimaciones obtenidas	103
Conclusiones		109
A	Algunos resultados de la teoría de la probabilidad	111
B	Resultados auxiliares	115
C	Algunos conceptos en genética	121
Bibliografía		123

Introducción

En casi todas las áreas científicas se hace uso de la probabilidad, ya que la mayoría de los fenómenos de la naturaleza no son deterministas. En particular en la genética se han hecho modelos con procesos estocásticos para poder inferir acerca del comportamiento de las poblaciones, logrando con ello un mayor conocimiento en la evolución de los seres vivos. En esta tesis se presentan algunos de estos modelos con la finalidad de mostrar como a través de ellos se obtiene una estimación del tiempo para obtener el ancestro en común de una población.

Tal estimación será a través del proceso n -coalescente propuesto por Kingman (ver Kingman (1982a, 1982b)), ya que él constituye una forma viable de como descubrir la relación ancestral de una población. Se tiene que éste modelo supone que el comportamiento de la población es descrito por otros dos modelos el de Wright-Fisher y el de Moran (ver Feller (1951), Karlin y McGregor (1964) y Moran (1958)), que indican la forma de reproducción de una población, por lo cual también estos modelos se describen.

Como el objetivo de la tesis es presentar de forma teórica como se utiliza el proceso n -coalescente de Kingman para estimar el tiempo para llegar al ancestro en común de una muestra obtenida a partir de una población en un cierto tiempo t_0 , esta tesis es presentada de la siguiente manera.

En el **Capítulo 1** primero se desarrolla los resultados principales sobre cadenas de Markov a tiempo discreto y posteriormente los correspondientes a tiempo continuo.

En el **Capítulo 2** se describirán los modelos de Wright-Fisher y de Moran. En ambos modelos se mostrarán resultados para el proceso sin mutación y con mutación. Por la estructura que tienen ambos modelos se tiene que si se realiza un cambio en la escala de tiempo y en el espacio donde se define el proceso que describe el comportamiento de la población estos dos modelos son

equivalentes (ver Karlin y Taylor (1981), Rodrigues (2002), Sánchez (2002)). Así que sólo describirá el modelo de Moran en tiempo continuo y también se presentarán sus probabilidades límite.

En el **Capítulo 3** se presentará el n -coalescente el cual como ya se había mencionado describe la genealogía de las poblaciones a partir de suponer que la población se comporta como la del modelo Moran. Se describe la forma en que se obtienen la probabilidades de obtener determinados ancestros en tiempo continuo, lo cual hace que se obtengan resultados acerca del tiempo para obtener el común antecesor.

En el **Capítulo 4** se presentan algunos resultados que se tienen en la inferencia Bayesiana y métodos de Monte Carlo, ya que se hace uso de estos para obtener algunas inferencias en el Capítulo 5.

En el **Capítulo 5** se describe un modelo llamado modelo con un número infinito de sitios, éste modelo es muy interesante porque a partir de la comparación de regiones homólogas en el ADN se logra determinar una aproximación a la probabilidad de tener determinada historia ancestral de una muestra de la población y con ello una estimación del tiempo esperado para obtener el más reciente común ancestro de la población mencionada, haciendo uso del modelo n -coalescente y de métodos de Monte Carlo. Adicionalmente se presenta como estimar la tasa de mutación genética en esta muestra.

Capítulo 1

Cadenas de Markov

Antes de iniciar la parte correspondiente a la teoría que se ha desarrollado en cadenas de Markov se describirán algunos conceptos básicos de la teoría de probabilidad, ya que se hará uso de ellos en la tesis (para mayores detalles ver Ash (1972) y Ross (1998)).

1.1 Resultados preliminares

Definición 1.1.1. Se le llama **espacio muestral** al conjunto de posibles resultados de un experimento aleatorio y se le denotará como Ω . Un **evento** es un subconjunto de Ω .

Ejemplo 1.1.1. Supóngase que en una muestra de genes el tipo de alelo sólo puede ser de tipo A_1 , A_2 ó A_3 . Si el experimento aleatorio consiste en elegir un gen y observar su tipo de alelo, entonces el espacio muestral estaría dado por $\Omega = \{A_1, A_2, A_3\}$, y un evento puede ser el siguiente:
 $A =$ La obtención del tipo de alelo A_1 .

Definición 1.1.2. Sea \mathcal{C} una colección no vacía de subconjuntos de un espacio Ω , \mathcal{C} es una σ -álgebra en Ω si satisface:

- i) $\Omega \in \mathcal{C}$.
- ii) Si $A \in \mathcal{C}$, entonces $A^c \in \mathcal{C}$.
- iii) Si $(A_n)_{n \geq 1}$ es una sucesión numerable de elementos en \mathcal{C} , entonces $\bigcup_{n=1}^{\infty} A_n \in \mathcal{C}$.

Definición 1.1.3. Sea S un conjunto no vacío y \mathcal{F} una σ -álgebra de S , a la pareja (S, \mathcal{F}) se le llama **espacio medible**.

Definición 1.1.4. Sea (Ω, \mathcal{F}) un espacio medible, una medida en (Ω, \mathcal{F}) , es una función conjuntista $\mu: \mathcal{F} \rightarrow \mathbb{R}$, con las siguientes propiedades:

- i) $\mu(\emptyset) = 0$,
- ii) $\mu(A) \geq 0$, $A \in \mathcal{F}$,
- iii) μ es σ -aditiva, es decir, si $(A_n)_{n \geq 1}$ es una sucesión numerable de elementos disjuntos de \mathcal{F} , entonces:

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n),$$

si además se tiene que $\mu(\Omega) = 1$, entonces a μ se le llama **medida de probabilidad** y se denotará por $P(\cdot)$.

Definición 1.1.5. A la terna (Ω, \mathcal{F}, P) donde Ω es el espacio muestral \mathcal{F} es una σ -álgebra en Ω y P es una medida de probabilidad en \mathcal{F} , se le llama **espacio de probabilidad**.

Definición 1.1.6. Sean (Ω, \mathcal{F}, P) un espacio de probabilidad y $A, B \in \mathcal{F}$ tal que $P(B) \neq 0$, se define la **probabilidad condicional** de A dado B como:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}.$$

Observación: Note que $P(\cdot|B)$ define una medida de probabilidad (ver Lema A.1 del Apéndice A).

Definición 1.1.7. Sean (Ω, \mathcal{F}, P) un espacio de probabilidad y $X: \Omega \rightarrow \mathbb{R}$ una función, se dice que X es \mathcal{F} -medible si:

$$X^{-1}\{(-\infty, x]\} \in \mathcal{F}, \text{ para toda } x \in \mathbb{R},$$

es decir,

$$\{\omega \in \Omega | X(\omega) \leq x\} \in \mathcal{F}.$$

Definición 1.1.8. Sea (Ω, \mathcal{F}, P) un espacio de probabilidad y $X : \Omega \rightarrow \mathbb{R}$ una función, se dice que X es **variable aleatoria** definida en (Ω, \mathcal{F}, P) si X es \mathcal{F} -medible.

Notas:

- i) La σ -álgebra generada por $(-\infty, x]$, $x \in \mathbb{R}$, se le llama **σ -álgebra de Borel** y se le denota por \mathcal{B} , por lo tanto se puede demostrar que X es variable aleatoria si $X^{-1}(B) \in \mathcal{F}$, $B \in \mathcal{B}$.
- ii) Como P es una medida en (Ω, \mathcal{F}) , entonces $P(X^{-1}((-\infty, x]))$ está bien definida. Por lo cual se define la medida inducida por X en \mathcal{B} indicada por P_X , de manera natural como:

$$P_X(B) = P(X^{-1}(B)), \quad B \in \mathcal{B}.$$

Definición 1.1.9. Sea (Ω, \mathcal{F}, P) un espacio de probabilidad y $X : \Omega \rightarrow \mathbb{R}$ una variable aleatoria definida en (Ω, \mathcal{F}, P) . Se define la **función de distribución** de X como la función $F_X : \mathbb{R} \rightarrow [0, 1]$, tal que:

$$F_X(x) = P_X((-\infty, x]) = P(\{\omega \in \Omega | X(\omega) \leq x\}) = P(X \leq x).$$

Definición 1.1.10. Un **proceso estocástico** $X = \{X_t : t \in T\}$ es una colección de variables aleatorias indexadas por un conjunto T , definidas en un mismo espacio de probabilidad (Ω, \mathcal{F}, P) , tomando valores en un mismo conjunto S . A S y a T se les llama **espacio de estados** y **conjunto de índices**, respectivamente.

Definición 1.1.11. Un proceso estocástico $X = \{X_t : t \in T\}$ con espacio de estado S es un **proceso de Markov** si satisface la propiedad de Markov, es decir, para $s, t \in T$, $s < t$ la distribución de X_t dado $\{X_u, u \leq s\}$ es igual a la distribución de X_t dado X_s , formalmente, para $x_u \in S$, $u \leq s$ y $A \subset S$, se tiene lo siguiente:

$$P(X_t \in A | X_u = x_u, u \leq s) = P(X_t \in A | X_s = x_s).$$

Cuando S es finito o infinito numerable, el proceso X es llamado **cadena de Markov**.

Definición 1.1.12. Cuando en una cadena de Markov $X = \{X_t : t \in T\}$ con espacio de estados S , el conjunto T es un intervalo real, X es llamada **cadena de Markov en tiempo continuo** o **cadena de Markov continua**. En el caso que T sea un subconjunto de \mathbb{Z} , X es llamada **cadena de Markov en tiempo discreto** o simplemente **cadena de Markov discreta**. En general se tomará a $T = [0, \infty)$ para el caso continuo y a $T = \{0, 1, 2, \dots\}$ para el caso discreto.

Definición 1.1.13. Sea S un conjunto de índices finito o infinito numerable. Una matriz $A = (a_{ij})_{i,j \in S}$ es llamada **estocástica** si satisface lo siguiente:

$$(i) \quad 0 \leq a_{ij} \leq 1, \quad i, j \in S.$$

$$(ii) \quad \sum_{j \in S} a_{ij} = 1, \quad i \in S.$$

En lo restante del capítulo se analizará primero el caso donde la cadena de Markov es discreta y después el caso donde es continua.

1.2 Cadenas de Markov a tiempo discreto

Definición 1.2.1. Sea $X = \{X_n : n \geq 0\}$ una cadena de Markov discreta con espacio de estados S . Se definen las **probabilidades de transición en un paso** (o **probabilidades de transición**) de X por:

$$P_{ij}^{(n,n+1)} = P(X_{n+1} = j | X_n = i), \quad i, j \in S, \quad n \geq 0,$$

la **probabilidad de transición en m pasos** está definida por:

$$P_{ij}^{(n,n+m)} = P(X_{n+m} = j | X_n = i), \quad i, j \in S, \quad m, n \geq 0,$$

donde para todo $n \geq 0$, se define:

$$P_{ij}^{(n,n)} = \delta_{ij} = \begin{cases} 1, & \text{si } j = i \\ 0, & \text{si } j \neq i. \end{cases} \quad (1.1)$$

Cuando las probabilidades de transición $P_{ij}^{(n,n+m)}$ no dependen de n , se dice que la cadena es **homogénea en el tiempo** y se denotan por:

$$P_{ij}^{(m)} = P(X_{n+m} = j | X_n = i), \quad i, j \in S, \quad n, m \geq 0.$$

En particular cuando $m = 1$ se tienen las probabilidades de transición y en este caso se tiene que:

$$P_{ij} = P(X_{n+1} = j | X_n = i), \quad i, j \in S, \quad n \geq 0.$$

Propiedades: Como las probabilidades de transición son probabilidades condicionales, entonces satisfacen para todo $n, m \geq 0$, lo siguiente:

$$(i) \quad 0 \leq P_{ij}^{(n, n+m)} \leq 1, \quad i, j \in S.$$

$$(ii) \quad \sum_{j \in S} P_{ij}^{(n, n+m)} = 1, \quad i \in S.$$

Definición 1.2.2. Sea $X = \{X_n : n \geq 0\}$ una cadena de Markov discreta con espacio de estados S , a la matriz que contenga las probabilidades de transición como elementos será llamada **matriz de transición** de X y será indicada por:

$$P^{(n, n+1)} = (P_{ij}^{(n, n+1)})_{i, j \in S}, \quad n \geq 0.$$

La matriz que tiene como elementos las probabilidades de transición en m pasos $m \geq 1$, será llamada **matriz de transición de m pasos** de X y será indicada por:

$$P^{(n, n+m)} = (P_{ij}^{(n, n+m)})_{i, j \in S}, \quad n \geq 0.$$

Cuando X es homogénea en el tiempo entonces $P^{(n, n+1)}$ y $P^{(n, n+m)}$ son indicadas por P y $P^{(m)}$, respectivamente.

Observación: Note que la matriz $P^{(n, n+m)}$, $m, n \geq 0$, satisface las propiedades de la Definición 1.1.13, por lo tanto es estocástica.

Definición 1.2.3. Sea S el espacio de estados de una cadena de Markov discreta $X = \{X_n : n \geq 0\}$ homogénea en el tiempo, con matriz de transición $P = (P_{ij})_{i, j \in S}$. Se dice que $i \in S$, es un **estado absorbente** de X si:

$$P_{ij} = \begin{cases} 1, & \text{si } j = i \\ 0, & \text{si } j \neq i. \end{cases} \quad (1.2)$$

Uno de los resultados más relevantes en cadenas de Markov son las ecuaciones del siguiente teorema, dado que proporcionan una forma recursiva de calcular las probabilidades de transición en m pasos, $m \geq 1$.

Teorema 1.2.1. Ecuaciones de Chapman-Kolmogorov. Sea $X = \{X_n : n \geq 0\}$ una cadena de Markov discreta con espacio de estados S , matriz de transición en m pasos $P^{(n,n+m)}$ y matriz de transición $P^{(n,n+1)}$, entonces para todo $i, j \in S$, $n \geq 0$ y $0 \leq r < m$, se vale:

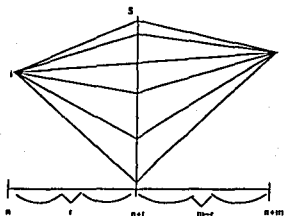
$$P_{ij}^{(n,n+m)} = \sum_{k \in S} P_{ik}^{(n,n+r)} P_{kj}^{(n+r,n+r+(m-r))}. \quad (1.3)$$

En particular cuando la cadena es homogénea se cumple:

$$P_{ij}^{(m)} = \sum_{k \in S} P_{ik}^{(r)} P_{kj}^{(m-r)}. \quad (1.4)$$

Demostración. La idea de la demostración resulta ser muy intuitiva, ya que (1.3) indica que la probabilidad de transición de m pasos se puede obtener a partir de las probabilidades de transición de r pasos ($0 \leq r \leq m$), es como si en lugar de ir de un estado a otro en un salto de tamaño por ejemplo m , se hiciera en dos saltos de tamaño menor que m , pero tales que suman m .

Lo anterior se puede observar en la siguiente figura:



Formalmente se tiene que:

$$\begin{aligned}
 P_{ij}^{(n,n+m)} &\stackrel{(1)}{=} P(X_{n+m} = j | X_n = i) \\
 &\stackrel{(2)}{=} \sum_{k \in S} P(X_{n+m} = j, X_{n+r} = k | X_n = i) \\
 &\stackrel{(3)}{=} \sum_{k \in S} P(X_{n+m} = j | X_{n+r} = k, X_n = i) P(X_{n+r} = k | X_n = i) \\
 &\stackrel{(4)}{=} \sum_{k \in S} P(X_{n+m} = j | X_r = k) P(X_r = k | X_n = i) \\
 &\stackrel{(1)}{=} \sum_{k \in S} P_{ik}^{(n,n+r)} P_{kj}^{(n+r,n+r+m-r)}
 \end{aligned}$$

- (1) Por definición de probabilidades de transición.
- (2) Por propiedades de probabilidad conjunta.
- (3) Es válido por la definición de probabilidad condicional.
- (4) Es válido porque X es cadena de Markov. □

Lema 1.2.2. Sea $X = \{X_n : n \geq 0\}$ una cadena de Markov discreta con espacio de estados S finito, se tiene que para $m \geq 1$, se vale lo siguiente:

- (a) Si X es homogénea en el tiempo y $P^{(m)}$ es su matriz de transición en m pasos, entonces:

$$P^{(m)} = P^m,$$

donde P^m es la matriz P a su m -ésima potencia.

- (b) Para X no homogénea en el tiempo con matriz de transición en m pasos $P^{(n,n+m)}$, se tiene que:

$$P^{(n,n+m)} = \prod_{j=0}^{m-1} P^{(n+j,n+j+1)}, \quad n \geq 0.$$

Demostración. Se utilizará inducción matemática para demostrar ambos apartados.

(a) Se tiene que la matriz de transición esta dada por $P = (P_{ij})_{i,j \in S}$.

Se demostrará primero que se cumple $P^{(2)} = P^2$.

Por las ecuaciones de Chapman-Kolmogorov, se vale lo siguiente:

$$P_{ij}^{(2)} = \sum_{k \in S} P_{ik} P_{kj},$$

por lo cual se tiene que para cada $i, j \in S$, $P_{ij}^{(2)}$ corresponde al elemento a_{ij} del producto de P y P . Por lo tanto $P^{(2)} = P^2$.

Supóngase ahora que para $m \geq 0$ se cumple, es decir, suponga que: $P^{(m)} = P^m$.

Finalmente se demostrará que la afirmación (a) del Lema se cumple para $m + 1$ y por lo tanto por inducción vale para todo $m \geq 0$.

De nuevo por las ecuaciones de Chapman-Kolmogorov

$$P_{ij}^{(m+1)} = \sum_{k \in S} P_{ik}^{(m)} P_{kj}, \quad (1.5)$$

dado que $P^{(m)} = P^m$ por hipótesis de inducción, entonces se tiene por (1.5) que $P_{ij}^{(m+1)}$ coincide con el elemento a_{ij} de la matriz $P^{(m)}P$. Por lo tanto $P^{(m+1)} = P^{m+1}$.

(b) Este apartado se demuestra de forma similar al apartado (a). □

Definición 1.2.4. Sea $X = \{X_n : n \geq 0\}$ una cadena de Markov discreta con espacio de estados S , se define la **probabilidad inicial** de X como:

$$\pi_0(i) = P(X_0 = i), \quad i \in S.$$

Lema 1.2.3. En una cadena de Markov discreta $X = \{X_n : n \geq 0\}$ con espacio de estados S , y matriz de transición en m pasos $P^{(n, n+m)}$. Se tiene que la función de probabilidad conjunta de X_n, X_{n-1}, \dots, X_0 , $n \geq 0$, se puede obtener en términos de la probabilidad inicial y las probabilidades de transición de X , de la siguiente manera:

$$P(X_n = i_n, X_{n-1} = i_{n-1}, \dots, X_0 = i_0) = \left(\prod_{k=0}^{n-1} P_{i_k i_{k+1}}^{(k, k+1)} \right) \pi_0(i_0), \quad (1.6)$$

donde $i_k \in S$, $k = 0, 1, 2, \dots, n$.

Demostración. Se utilizará inducción en n para demostrar este Lema.

i) Se tiene que para $n = 1$ vale, por la definición de la probabilidad condicional y de la inicial, dado que:

$$\begin{aligned} P(X_1 = i_1, X_0 = i_0) &= P(X_1 = i_1 | X_0 = i_0) P(X_0 = i_0) \\ &= P_{i_0 i_1} \pi_0(i_0). \end{aligned}$$

ii) Supóngase que el resultado vale para $n = k$, es decir,

$$P(X_k = i_k, \dots, X_1 = i_1, X_0 = i_0) = \left(\prod_{l=0}^{k-1} P_{i_l i_{l+1}}^{(l, l+1)} \right) \pi_0(i_0).$$

iii) Finalmente se demostrará que vale para $n = k + 1$, y por lo tanto se cumple para todo $n \geq 0$.

$$P(X_{k+1} = i_{k+1}, X_k = i_k, \dots, X_0 = i_0)$$

$$\stackrel{(1)}{=} P(X_{k+1} = i_{k+1} | X_k = i_k, \dots, X_0 = i_0) P(X_k = i_k, \dots, X_0 = i_0)$$

$$\stackrel{(2)}{=} P(X_{k+1} = i_{k+1} | X_k = i_k, \dots, X_0 = i_0) \left(\prod_{l=0}^{k-1} P_{i_l i_{l+1}}^{(l, l+1)} \right) \pi_0(i_0)$$

$$\stackrel{(3)}{=} P(X_{k+1} = i_{k+1} | X_k = i_k) \left(\prod_{l=0}^{k-1} P_{i_l i_{l+1}}^{(l, l+1)} \right) \pi_0(i_0)$$

$$\stackrel{(4)}{=} \left(\prod_{l=0}^k P_{i_l i_{l+1}}^{(l, l+1)} \right) \pi_0(i_0).$$

(1) Por definición de probabilidad condicional.

(2) Por hipótesis de inducción.

(3) Por que X es una cadena de Markov.

(4) Se vale por definición de probabilidades de transición. □

Observación: Cuando X es homogénea en el tiempo, entonces (1.6) se puede escribir como

$$P(X_n = i_n, X_{n-1} = i_{n-1}, \dots, X_0 = i_0) = \left(\prod_{k=0}^{n-1} P_{i_k i_{k+1}} \right) \pi_0(i_0). \quad (1.7)$$

Una de las características de las cadenas de Markov es que no sólo sus probabilidades de transición se pueden obtener a partir de otras probabilidades como se muestra en las ecuaciones de Chapman-Kolmogorov sino también las probabilidades $P(X_n = \cdot)$, como se podrá ver en el siguiente lema.

Lema 1.2.4. Sea $X = \{X_n : n \geq 0\}$ una cadena de Markov discreta homogénea en el tiempo con espacio de estados S , matriz de transición P y matriz de transición de m pasos $P^{(m)}$, si $\pi_n(\cdot) = P(X_n = \cdot)$ entonces para todo $n \geq 0$ y $k \in S$, se vale:

$$(a) \pi_n(k) = \sum_{j \in S} \pi_0(j) P_{jk}^{(n)},$$

o equivalentemente

$$(b) \pi_n(k) = \sum_{j \in S} \pi_{n-1}(j) P_{jk}.$$

Demostración. (a) Para este apartado note que:

$$\begin{aligned} \pi_n(k) = P(X_n = k) &\stackrel{(1)}{=} \sum_{j \in S} P(X_n = k, X_0 = j) \\ &\stackrel{(2)}{=} \sum_{j \in S} P(X_n = k | X_0 = j) P(X_0 = j) \\ &\stackrel{(3)}{=} \sum_{j \in S} P_{jk}^{(n)} \pi_0(j). \end{aligned}$$

(b) En forma análoga al anterior apartado se tiene que:

$$\begin{aligned} \pi_n(k) = P(X_n = k) &\stackrel{(1)}{=} \sum_{j \in S} P(X_n = k, X_{n-1} = j) \\ &\stackrel{(2)}{=} \sum_{j \in S} P(X_n = k | X_{n-1} = j) P(X_{n-1} = j) \\ &\stackrel{(4)}{=} \sum_{j \in S} P_{jk} \pi_{n-1}(j). \end{aligned}$$

(1) Por propiedades de función de probabilidad conjunta.

(2) Es válido por la definición de probabilidad condicional.

(3) Se vale por definición de probabilidades de transición y de la probabilidad inicial.

(4) Se vale por definición de probabilidades de transición y de $\pi_n(\cdot)$ □

Observación: Cuando X no es homogénea en el tiempo, entonces (i) y (ii) del Lema 1.2.4 estarían dadas de la siguiente manera:

$$(a) \pi_n(k) = \sum_{j \in S} \pi_0(j) P_{jk}^{(0,n)},$$

o equivalentemente

$$(b) \pi_n(k) = \sum_{j \in S} \pi_{n-1}(j) P_{jk}^{(n,n-1)}.$$

1.2.1 Clasificación de estados

En esta sección se presentarán las condiciones bajo las cuales una cadena de Markov converge para el estado de equilibrio cuando se hace tender el tiempo a infinito. En primer lugar se presentarán algunas definiciones y se finalizará con el resultado que garantiza la existencia de:

$$\lim_{n \rightarrow \infty} P_{ij}^{(n)},$$

y quien es este límite.

Definición 1.2.5. Sea S el espacio de estados para una cadena de Markov discreta $X = \{X_n : n \geq 0\}$, homogénea en el tiempo y con matriz de transición de m pasos $P^{(m)}$, se tiene para $i, j \in S$, lo siguiente:

(a) se dice que el estado j es accesible desde el estado i , si para alguna $m \geq 0$, $P_{ij}^{(m)} > 0$. Se denota por $i \rightarrow j$ al hecho de que j es accesible desde i ,

(b) si $i \rightarrow j$ y además $j \rightarrow i$, es decir, si existen $n, m \geq 0$, tales que $P_{ij}^{(n)} > 0$ y $P_{ji}^{(m)} > 0$, entonces se dice que los dos estados se comunican, y se denota como: $i \leftrightarrow j$.

Observación: Note que si dos estados i, j no se comunican entonces para toda $n \geq 0$, se tiene que:

$$P_{ij}^{(n)} = 0 \quad \text{o} \quad P_{ji}^{(n)} = 0.$$

Se tiene que la comunicación entre estados de una cadena de Markov es una relación de equivalencia, lo cual se formaliza en la siguiente proposición.

Proposición 1.2.5. Sea S el espacio de estados de una cadena de Markov discreta $X = \{X_n : n \geq 0\}$, homogénea en el tiempo con matriz de transición de m pasos $P^{(m)}$, entonces:

1. $i \leftrightarrow i$, $i \in S$,
2. $i \leftrightarrow j$, entonces $j \leftrightarrow i$, $i, j \in S$,
3. $i \leftrightarrow j$ y $j \leftrightarrow k$, entonces $i \leftrightarrow k$, $i, j, k \in S$.

Demostración. 1. Se sigue de (1.1), ya que $P_{ii}^{(0)} = 1 > 0$.

2. Es inmediato pues como el estado i se comunica con j , entonces existen $n_1, n_2 \geq 0$, tales que $P_{ji}^{(n_2)} > 0$ y $P_{ij}^{(n_1)} > 0$, se sigue que $j \leftrightarrow i$.
3. Como $i \leftrightarrow j$ entonces existen $n_1, n_2 \geq 0$, tales que $P_{ij}^{(n_1)} > 0$ y $P_{ji}^{(n_2)} > 0$ y como además $j \leftrightarrow k$, entonces también existen $m_1, m_2 \geq 0$, tal que $P_{jk}^{(m_1)} > 0$ y $P_{kj}^{(m_2)} > 0$, lo cual implica que:

$$P_{ik}^{(n_1+m_1)} \stackrel{(1)}{=} \sum_{l \in S} P_{il}^{(n_1)} P_{lk}^{(m_1)} \\ \stackrel{(2)}{\geq} P_{ij}^{(n_1)} P_{jk}^{(m_1)} > 0,$$

por lo tanto $i \rightarrow k$. De forma análoga se obtiene que $k \rightarrow i$. De esta forma se tiene que $i \leftrightarrow k$.

- (1) Por las ecuaciones de Chapman-Kolmogorov.
- (2) Es válido ya que $P_{ij}^{(m)} \geq 0$, para todo $i, j \in S$ y $m \geq 0$. □

Nota: En general se tiene que la relación de comunicación induce una partición del conjunto donde ella está definida, entonces por la Proposición 1.2.5, en una cadena de Markov se tiene una partición del espacio de estados S en clases de equivalencia, donde la relación de equivalencia es la de comunicación entre los estados de esta cadena de Markov.

Definición 1.2.6. Sea $X = \{X_n : n \geq 0\}$ una cadena de Markov discreta homogénea en el tiempo, con espacio de estados S . Se dice que X es **irreducible**, si todos los estados en S se comunican, es decir, X es irreducible si la relación de equivalencia dada en la Proposición 1.2.5 induce sólo una clase de equivalencia en S .

Definición 1.2.7. Sea $X = \{X_n : n \geq 0\}$ una cadena de Markov discreta homogénea en el tiempo con espacio de estados S y matriz de transición en m pasos $P^{(m)}$. Sea $i \in S$, se define el período de un estado i , denotado por $d(i)$, como:

$$d(i) = \text{m.c.d.}\{n \geq 1 : P_{ii}^{(n)} > 0\},$$

donde m.c.d. indica el máximo común divisor. Si $P_{ii}^{(n)} = 0$, para todo $n > 0$, se define $d(i) = 0$. Cuando $d(i) = 1$, se dice que el estado i es **aperiódico**.

Observación: El período $d(i)$ de un estado i es el mayor número con la propiedad de que para toda $m \neq d(i), 2d(i), \dots$, se tiene que $P_{ii}^{(m)} = 0$.

Definición 1.2.8. Una cadena de Markov discreta $X = \{X_n : n \geq 0\}$ homogénea en el tiempo con espacio de estados S , en la cual cada estado $i \in S$ es aperiódico, es decir, tiene período $d(i) = 1$, es llamada **aperiódica**.

Proposición 1.2.6. Sea $X = \{X_n : n \geq 0\}$ una cadena de Markov discreta, homogénea en el tiempo, con espacio de estados S . Si $i \longleftrightarrow j$, $i, j \in S$, entonces $d(i) = d(j)$, es decir, el período de un estado es una propiedad de clase.

Demostración. En el caso que $i = j$ es claro que se cumple. De esta forma considere el caso donde $i \neq j$. Como i se comunica con j , entonces existen $n_1, n_2 \geq 0$, tales que $P_{ji}^{(n_1)} > 0$ y $P_{ij}^{(n_2)} > 0$.

Sea $m^* \in M = \{m > 0 : P_{ii}^{(m)} > 0\}$, entonces:

$$\begin{aligned} P_{jj}^{(n_1+m^*+n_2)} &\stackrel{(1)}{=} \sum_{k \in S} P_{jk}^{(n_1+m^*)} P_{kj}^{(n_2)} \stackrel{(2)}{\geq} P_{ji}^{(n_1+m^*)} P_{ij}^{(n_2)} \\ &\stackrel{(1)}{=} \left(\sum_{k \in S} P_{jk}^{(n_1)} P_{ki}^{(m^*)} \right) P_{ij}^{(n_2)} \stackrel{(2)}{\geq} P_{ji}^{(n_1)} P_{ii}^{(m^*)} P_{ij}^{(n_2)} > 0. \end{aligned}$$

Análogamente se tiene que $P_{jj}^{(n_1+n_2)} > 0$. Entonces

$$d(j) \text{ divide a } (n_1 + n_2) \text{ y } d(j) \text{ divide a } (n_1 + n_2 + m^*),$$

entonces para todo $m^* \in M$, $d(j)$ divide a m^* . Por lo tanto $d(j) \leq d(i)$. De manera análoga se obtiene que $d(i) \leq d(j)$ y por lo tanto $d(j) = d(i)$.

(1) Por las ecuaciones de Chapman-Kolmogorov.

(2) Es válido ya que $P_{ij}^{(n)} \geq 0$, para todo $i, j \in S$ y $m \geq 0$. □

Definición 1.2.9. Sea $X = \{X_n : n \geq 0\}$ una cadena de Markov discreta, homogénea en el tiempo, con espacio de estados S , matriz de transición P y $P^{(m)}$, $m \geq 0$, su matriz de transición de m pasos. Para $i \in S$, i fijo, la probabilidad de que X entre por primera vez al estado j en el n -ésimo paso, dado que partirá del estado i , se define por:

$$f_{ij}^{(n)} = \begin{cases} P(X_n = j, X_k \neq j, k = 1, 2, \dots, n-1 | X_0 = i), & n \geq 1, i, j \in S \\ 0, & n = 0, i, j \in S. \end{cases} \quad (1.8)$$

En particular $f_{ij}^{(1)} = P_{ij}$, para todo $i, j \in S$.

Lema 1.2.7. Para X como en la Definición 1.2.9, se tiene que para todo $i, j \in S$ y $n \geq 0$, se vale lo siguiente:

$$P_{ij}^{(n)} = \sum_{k=0}^n f_{ij}^{(k)} P_{jj}^{(n-k)}.$$

Demostración. Note que si se denota por $E^{(n)}$ al evento de que al paso n , $n \geq 0$, X se encuentre en el estado j dado que partió del estado i , entonces $E^{(n)}$ comprende tanto el hecho de que la primera transición al estado j haya sido en el paso n , como en un paso anterior a n . Por lo tanto si se considera a E_k como el evento de retornar por primera vez al estado i en k -pasos, entonces el evento $E^{(n)}$ está dado por la unión de los E_k , $k = 1, \dots, n$. Como estos eventos son mutuamente excluyentes para valores diferentes de k y $P_{ij}^{(n)}$ es la probabilidad de $E^{(n)}$, se vale lo siguiente:

$$\begin{aligned}
 P_{ij}^{(n)} &\stackrel{(1)}{=} P(X_n = j | X_0 = i) \\
 &= \sum_{k=0}^n P(X_n = j, X_k = j, X_r \neq j, 0 < r < k | X_0 = i) \\
 &\stackrel{(2)}{=} \sum_{k=0}^n P(X_n = j | X_k = j, X_r \neq j, 0 < r < k, X_0 = i) \\
 &\quad \times P(X_k = j, X_r \neq j, 0 < r < k | X_0 = i) \\
 &\stackrel{(3)}{=} \sum_{k=0}^n P(X_n = j | X_k = j) P(X_k = j, X_r \neq j, 0 < r < k | X_0 = i) \\
 &\stackrel{(4)}{=} \sum_{k=0}^n P_{jj}^{(n-k)} f_{ij}^{(k)}.
 \end{aligned}$$

- (1) Por definición de probabilidades de transición.
- (2) Por la definición de probabilidad condicional.
- (3) Es válido porque X es cadena de Markov.
- (4) Por definición de probabilidades de transición y por la Definición 1.2.9. □

Definición 1.2.10. Sea $X = \{X_n : n \geq 0\}$ una cadena de Markov discreta, homogénea en el tiempo con espacio de estados S , se define la probabilidad de que partiendo del estado i se llegue al estado j en algún paso $n \geq 1$, por:

$$f_{ij} = \sum_{n=0}^{\infty} f_{ij}^{(n)}, \quad i, j \in S.$$

Definición 1.2.11. Sea S el espacio de estados de una cadena de Markov discreta $X = \{X_n : n \geq 0\}$, homogénea en el tiempo, entonces para $i \in S$:

- (i) se dice que i es recurrente si $f_{ii} = 1$,
- (ii) se dice que i es transitorio si $f_{ii} < 1$.

Observación: Note que la Definición 1.2.11 indica que si i es transitorio, entonces X tiene probabilidad positiva de no retornar al estado i .

Teorema 1.2.8. Sea $X = \{X_n : n \geq 0\}$ una cadena de Markov discreta, homogénea en el tiempo, con espacio de estados S y matriz de transición de n pasos $P^{(n)}$, entonces:

(a) i es un estado transitorio si y sólo si

$$\sum_{n=1}^{\infty} P_{ii}^{(n)} < \infty,$$

(b) i es un estado recurrente si y sólo si

$$\sum_{n=1}^{\infty} P_{ii}^{(n)} = \infty.$$

Demostración. La demostración de (a) estará dada en (a1) y (a2) y la de (b) en (b1) y (b2). Para la demostración de ambos apartados se denotará por M a la variable aleatoria que indica el número de veces que la cadena visita el estado i .

Note que la variable aleatoria M se puede expresar como:

$$M = \sum_{n=0}^{\infty} I_{(i)}(X_n),$$

donde

$$I_{(i)}(X_n) = \begin{cases} 1, & \text{si } X_n = i \\ 0, & \text{si } X_n \neq i. \end{cases} \quad (1.9)$$

Entonces

$$\begin{aligned} E(M|X_0 = i) &= E\left(\sum_{n=0}^{\infty} I_{(i)}(X_n)|X_0 = i\right) \\ &\stackrel{(1)}{=} \sum_{n=0}^{\infty} E(I_{(i)}(X_n)|X_0 = i) \\ &\stackrel{(2)}{=} \sum_{n=0}^{\infty} P(X_n = i|X_0 = i) \\ &\stackrel{(3)}{=} \sum_{n=0}^{\infty} P_{ii}^{(n)}. \end{aligned} \quad (1.10)$$

Por otro lado por la Definición 1.2.10 note que f_{ii} es la probabilidad de que la cadena retorne al estado i en algún paso n , lo cual es equivalente a que se retorne al menos una vez.

- (a1) Supóngase que el estado i es transitorio, entonces por la Definición 1.2.11 $f_{ii} < 1$. Por lo tanto con probabilidad $1 - f_{ii} > 0$ el proceso que empezó en el estado i jamás regresará a i . De esta forma

$$P(M = k | X_0 = i) = (f_{ii})^k (1 - f_{ii}), \quad k = 0, 1, \dots$$

es decir, M es una variable aleatoria con parámetro $1 - f_{ii}$. Entonces,

$$E(M | X_0 = i) = \frac{1}{1 - f_{ii}} \quad (1.11)$$

(ver Lema A.2 del Apéndice A). Entonces, por (1.11) y (1.10) se tiene que

$$\sum_{n=0}^{\infty} P_{ii}^{(n)} = \frac{1}{1 - f_{ii}} < \infty.$$

- (b1) Ahora supóngase que i es recurrente, entonces por la Definición 1.2.11 se tiene que $f_{ii} = 1$. Como se está trabajando con cadenas de Markov, el hecho que el proceso regrese al estado i , es equivalente a que el proceso inicie nuevamente en i y por lo tanto regresará a i en con probabilidad 1 y así sucesivamente. De esta forma, M tendrá valor infinito y por lo tanto

$$E(M | X_0 = i) = \infty, \quad (1.12)$$

y por (1.10), se tiene que

$$\sum_{n=0}^{\infty} P_{ii}^{(n)} = \infty.$$

Se demostrará ahora el inverso de los apartados anteriores.

- (a2) Supóngase que $\sum_{n=0}^{\infty} P_{ii}^{(n)} < \infty$, entonces necesariamente se tiene que i es transitorio. Si no fuera así, i tendría que ser recurrente y por lo tanto por (b1) debería ser tal que $\sum_{n=0}^{\infty} P_{ii}^{(n)} = \infty$, lo cual es una contradicción.

(b2) Se sigue por un argumento similar al dado en (a2).

(1) Por propiedades de esperanza.

(2) Es válido por (1.9).

(3) Es válido por la definición de probabilidad de transición. \square

Proposición 1.2.9. Sea $X = \{X_n : n \geq 0\}$ una cadena de Markov discreta, homogénea en el tiempo, con espacio de estados S y matriz de transición de m pasos $P^{(m)}$. Entonces para $i, j \in S$, se tiene que:

(a) si $i \longleftrightarrow j$ y además i es recurrente, entonces j es recurrente,

(b) si $i \longleftrightarrow j$ y además i es transitorio, entonces j es transitorio.

Demostración. (a) Como el estado i se comunica con j , entonces existen $n_1, n_2 \geq 1$, tales que $P_{ij}^{(n_1)} > 0$, $P_{ji}^{(n_2)} > 0$. Tome $m \geq 0$, entonces

$$\begin{aligned} P_{jj}^{(n_1+m+n_2)} &\stackrel{(1)}{\geq} \sum_{k \in S} P_{jk}^{(n_2+m)} P_{kj}^{(n_1)} \stackrel{(2)}{\geq} P_{ji}^{(n_2+m)} P_{ij}^{(n_1)} \\ &\stackrel{(1)}{=} \left(\sum_{k \in S} P_{jk}^{(n_2)} P_{ki}^{(m)} \right) P_{ij}^{(n_1)} \stackrel{(2)}{\geq} P_{ji}^{(n_2)} P_{ii}^{(m)} P_{ij}^{(n_1)}, \end{aligned}$$

entonces

$$\sum_{m=0}^{\infty} P_{jj}^{(n_2+m+n_1)} \geq \sum_{m=0}^{\infty} P_{ji}^{(n_2)} P_{ii}^{(m)} P_{ij}^{(n_1)} = P_{ji}^{(n_2)} P_{ij}^{(n_1)} \sum_{m=0}^{\infty} P_{ii}^{(m)}.$$

Como $\sum_{m=0}^{\infty} P_{ii}^{(m)}$ diverge, entonces $\sum_{m=0}^{\infty} P_{jj}^{(m)}$ también diverge.

(b) Supóngase que j no es transitorio, entonces es recurrente. Por otro lado dado que el estado j se comunica con el estado i , entonces i sería recurrente por (a) de esta proposición, lo cual contradice que i sea transitorio.

(1) Por las ecuaciones de Chapman-Kolmogorov.

(2) Es válido ya que $P_{ij}^{(m)} \geq 0$, para todo $i, j \in S$ y $m \geq 0$. \square

Definición 1.2.12. Sea $X = \{X_n : n \geq 0\}$ una cadena de Markov discreta homogénea en el tiempo, con espacio de estados S . Se dice que la clase $C \subseteq S$ es **recurrente**, si para $j \in C$, se tiene que j es recurrente y es **transitoria** si para $j \in C$, se tiene que j es transitorio.

Teorema 1.2.10. Sea S el espacio de estados de una cadena de Markov discreta $X = \{X_n : n \geq 0\}$, homogénea en el tiempo, con matriz de transición de m pasos $P^{(m)}$. Sea $j \in S$ un estado recurrente y $d(j)$ su período, entonces

(i) existe el siguiente límite

$$\lim_{n \rightarrow \infty} P_{jj}^{(nd(j))},$$

(ii) si además el estado j es aperiódico y se comunica con i , $i \in S$, entonces existe el siguiente límite

$$\lim_{n \rightarrow \infty} P_{ij}^{(n)}.$$

Demostración. Ver Ross (1996). □

Definición 1.2.13. Sea S el espacio de estados de una cadena de Markov $X = \{X_n : n \geq 0\}$ discreta homogénea en el tiempo, con matriz de transición de m pasos $P^{(m)}$. Sea $j \in S$, un estado recurrente y $d(j)$ su período, si se cumple que

$$\lim_{n \rightarrow \infty} P_{jj}^{(nd(j))}$$

es estrictamente positivo entonces se dice que el estado j es **recurrente positivo**, en el caso que sea cero se dice que el estado j es **recurrente nulo**.

Definición 1.2.14. Sea $X = \{X_n : n \geq 0\}$ una cadena de Markov discreta homogénea en el tiempo, con espacio de estados S . Se dice que la clase $C \subseteq S$ es **recurrente positiva** si para $j \in C$, se tiene que j es recurrente positivo. En el caso que para todo estado $j \in C$, j sea recurrente nulo, se dice que la clase es **recurrente nula**.

Definición 1.2.15. Sea S el espacio de estados de una cadena de Markov $X = \{X_n : n \geq 0\}$ discreta, homogénea en el tiempo, un estado $i \in S$, es llamado **ergódico** si i es recurrente positivo y aperiódico.

Teorema 1.2.11. Sea $X = \{X_n : n \geq 0\}$ una cadena de Markov discreta, homogénea en el tiempo con espacio de estados S finito y matriz de transición de m pasos $P^{(m)}$. Se tiene que si X es recurrente, irreducible y aperiódica, entonces X es recurrente positiva.

Demostración. Como X es aperiódica e irreducible, por el Teorema 1.2.10 se tiene que existe el límite de $P_{ij}^{(n)}$ cuando n diverge a infinito. Supóngase que:

$$\lim_{n \rightarrow \infty} P_{ij}^{(n)} = 0, \quad \text{para todo } j \in S. \quad (1.13)$$

Por propiedades de probabilidades de transición, se tiene que para todo $i \in S$, se vale:

$$\sum_{j \in S} P_{ij}^{(n)} = 1, \quad \text{para todo } n \geq 0.$$

De esta forma se tiene que:

$$\begin{aligned} \sum_{j \in S} \lim_{n \rightarrow \infty} P_{ij}^{(n)} &\stackrel{(1)}{=} \lim_{n \rightarrow \infty} \sum_{j \in S} P_{ij}^{(n)} \\ &= \lim_{n \rightarrow \infty} 1 = 1. \end{aligned}$$

Por (1.13) se tendría que:

$$\sum_{j \in S} \lim_{n \rightarrow \infty} P_{ij}^{(n)} = 0,$$

por lo cual se llega a una contradicción. Por lo tanto existe al menos un $j^* \in S$, tal que $\lim_{n \rightarrow \infty} P_{ij^*}^{(n)} > 0$.

Sólo resta ver que

$$\lim_{n \rightarrow \infty} P_{ij}^{(n)} > 0, \quad \text{para todo } j \in S.$$

Como la cadena es irreducible, se tiene que todos los estados en S se comunican, entonces para $j \in S$, existe un $m \geq 0$, tal que $P_{j^*j}^{(m)} > 0$. Por otro lado por las ecuaciones de Chapman-Kolmogorov, se tiene que para todo $i, j \in S$, vale:

$$P_{ij}^{(n)} = \sum_{k \in S} P_{ik}^{(n-m)} P_{kj}^{(m)}, \quad \text{para todo } n > m,$$

entonces para todo $n > m$,

$$\begin{aligned} \lim_{n \rightarrow \infty} P_{ij}^{(n)} &= \lim_{n \rightarrow \infty} \sum_{k \in S} P_{ik}^{(n-m)} P_{kj}^{(m)} \\ &\stackrel{(1)}{=} \sum_{k \in S} \lim_{n \rightarrow \infty} [P_{ik}^{(n-m)} P_{kj}^{(m)}] \\ &\stackrel{(2)}{\geq} \lim_{n \rightarrow \infty} [P_{ij}^{(n-m)} P_{jj}^{(m)}] > 0 \\ &= P_{jj}^{(m)} \lim_{n \rightarrow \infty} P_{ij}^{(n-m)} > 0. \end{aligned}$$

Por lo tanto la cadena es recurrente positiva.

(1) Como S es finito entonces el límite de la suma es la suma de los límites ya que cada uno de estos existe.

(2) Se tiene que $P_{ij}^{(n)} \geq 0$, para toda $i, j \in S$ y $n \geq 0$. \square

Definición 1.2.16. La *distribución estacionaria* de una cadena de Markov discreta homogénea en el tiempo con matriz de transición $P = (P_{ij})_{i,j \in S}$, es una colección de números $\{\pi(k), k \in S\}$, tal que:

- (a) $\pi(k) \geq 0, k \in S$,
- (b) $\sum_{k \in S} \pi(k) = 1$,
- (c) $\pi(k) = \sum_{i \in S} \pi(i) P_{ik}, k \in S$.

Nota: Las ecuaciones dadas en (a), (b) y (c) de la Definición 1.2.16 son llamadas ecuaciones de balance total.

Teorema 1.2.12. Sea S el espacio de estados de una cadena de Markov $X = \{X_n : n \geq 0\}$ discreta, homogénea en el tiempo, con matriz de transición de m pasos $P^{(m)}$. Si X es recurrente positiva y aperiódica, entonces existen números $\pi(k) > 0, k \in S$, tales que

$$\lim_{n \rightarrow \infty} P_{ij}^{(n)} = \pi(j), \text{ para todo estado } j \in S,$$

independientes del estado inicial i , donde $\pi(k)$ están determinadas por el siguiente sistema de ecuaciones

- (a) $\pi(k) > 0, k \in S$,

$$(b) \sum_{k \in S} \pi(k) = 1,$$

$$(c) \pi(k) = \sum_{i \in S} \pi(i) P_{ik}.$$

Demostración. Ver por ejemplo Karlin y Taylor (1975) y Ross (1996). \square

Definición 1.2.17. Sea $X = \{X_n : n \geq 0\}$ una cadena de Markov discreta, homogénea en el tiempo, con espacio de estados S y matriz de transición de m pasos $P^{(m)}$. Cuando existe el $\lim_{n \rightarrow \infty} P_{ij}^{(n)}$, para todo j en S , el límite es llamado la **distribución límite** de la cadena de Markov.

Observación: Por el Teorema 1.2.12 se sabe que si la cadena de Markov es ergódica, entonces la distribución límite existe y es la única distribución estacionaria de la cadena.

1.3 Cadenas de Markov a tiempo continuo

Definición 1.3.1. Sea $X = \{X_t : t \in [0, \infty)\}$ una cadena de Markov continua con espacio de estados S . Se definen las **probabilidades de transición de X en un espacio de tiempo t** , como:

$$P_{ij}(s, s+t) = P(X_{s+t} = j | X_s = i), \quad i, j \in S, \quad s, t \geq 0,$$

donde para todo $s \geq 0$, se define:

$$P_{ij}(s, s) = \delta_{ij} = \begin{cases} 1, & \text{si } j = i \\ 0, & \text{si } j \neq i. \end{cases} \quad (1.14)$$

Análogamente al caso discreto cuando las probabilidades de transición no dependen de s , se dice que X es homogénea en el tiempo. En este caso las probabilidades de transición al tiempo t , se denotan por:

$$P_{ij}(t) = P(X_{s+t} = j | X_s = i), \quad i, j \in S, \quad s, t \geq 0,$$

con

$$P_{ij}(0) = \delta_{ij} = \begin{cases} 1, & \text{si } j = i \\ 0, & \text{si } j \neq i. \end{cases} \quad (1.15)$$

Definición 1.3.2. Sea $X = \{X_t : t \in [0, \infty)\}$ una cadena de Markov continua con espacio de estados S , a la matriz que contenga como elementos las probabilidades de transición en un espacio de tiempo t será llamada **matriz de transición de X al tiempo t** y será indicada por:

$$P(0, t) = (P_{ij}(0, t))_{i, j \in S}, \quad t \geq 0.$$

en el caso que X sea homogénea:

$$P(t) = (P_{ij}(t))_{i, j \in S}, \quad t \geq 0.$$

En el caso de las ecuaciones de Chapman-Kolmogorov para las cadenas de Markov continuas, se tiene que son análogas al caso discreto como se podrá observar en el siguiente teorema.

Teorema 1.3.1. Ecuaciones de Chapman-Kolmogorov.

Sea $X = \{X_t : t \in [0, \infty)\}$ una cadena de Markov continua en el tiempo con espacio de estados S , y matriz de transición al tiempo $t \geq 0$, $P(0, t) = (P_{ij}(0, t))_{i, j \in S}$, entonces para todo $i, j \in S$ y para $s, t, r \geq 0$, $r < t$, se vale:

$$P_{ij}(s, s+t) = \sum_{k \in S} P_{ik}(s, s+r) P_{kj}(s+r, s+r+(t-r)),$$

en forma análoga al caso discreto cuando X es homogénea en el tiempo, se tiene:

$$P_{ij}(t) = \sum_{k \in S} P_{ik}(r) P_{kj}(t-r).$$

Demostración. Es análoga al caso discreto. □

En esta sección se presentará una forma alterna de definir una cadena de Markov continua en el tiempo. Para este fin se necesitará el concepto de proceso semi-Markov que se dará a continuación.

Definición 1.3.3. Un proceso estocástico a tiempo continuo $X = \{X_t : t \in [0, \infty)\}$, con espacio de estados $S = \{0, 1, 2, \dots\}$, se llama **proceso semi-Markov** si cada vez que entra al estado $i \in S$, se tiene que:

- (a) el siguiente estado que entra será $j \in S$, con probabilidad P_{ij} , $\sum_{j \in S} P_{ij} = 1$,

(b) dado que el siguiente estado al que entrará es $j \in S$, el tiempo de permanencia en i antes de cambiar a j tiene distribución F_{ij} .

Definición 1.3.4. Sea $X = \{X_t : t \in [0, \infty)\}$, un proceso semi-Markov, con espacio de estados S . Sea X_n^* la variable aleatoria que indica el n -ésimo estado visitado por el proceso X , entonces $X^* = \{X_n^* : n = 0, 1, 2, \dots\}$ es una cadena de Markov discreta y es llamada cadena de saltos de X , si satisface que:

$$P(X_{n+1}^* = j | X_n^* = i) = P_{ij}, \quad i, j \in S, n \geq 0.$$

Definición 1.3.5. Una cadena de Markov en tiempo continuo $X = \{X_t : t \in [0, \infty)\}$, con espacio de estados S es un proceso semi-Markov, tal que:

(a) dado que X está en el estado $i \in S$ el próximo estado será $j \in S$, $j \neq i$, con probabilidad P_{ij} , donde $\sum_{j \neq i} P_{ij} = 1$,

(b) F_{ij} es una distribución exponencial con parámetro ν_i , es decir, si τ_{ij} es el tiempo de permanencia en i antes de pasar a $j \neq i$, τ_{ij} no depende de j , así que se toma a $\tau_{ij} = \tau_i$ y se tiene:

$$P(\tau_i > t) = e^{-\nu_i t}.$$

Nota: Si $\nu_i = \infty$, $i \in S$, i es llamado estado instantáneo, en el caso que $\nu_i = 0$, i es llamado estado absorbente. De ahora en adelante se considerará que para $i \in S$, $0 \leq \nu_i < \infty$.

Definición 1.3.6. Sea $X = \{X_t : t \in [0, \infty)\}$ una cadena de Markov continua homogénea en el tiempo con espacio de estados S , cuya cadena de saltos tiene probabilidades de transición P_{ij} , $i, j \in S$. Para $i, j \in S$, $i \neq j$, se tiene que la tasa de transición del estado i al estado j , indicada por q_{ij} , es definida por:

$$q_{ij} = \nu_i P_{ij}, \quad (1.16)$$

donde ν_i es la tasa de permanencia en i antes de pasar a j .

Observación: Note que para $i, j \in S$, $i \neq j$, se tiene lo siguiente:

$$\sum_{j \neq i} q_{ij} = \nu_i, \quad (1.17)$$

dado que

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

Definición 1.3.7. Sea $X = \{X_t : t \in [0, \infty)\}$ una cadena de Markov continua, homogénea en el tiempo con espacio de estados S y tasas de transición q_{ij} , $i, j \in S$. Se definen las **probabilidades de transición infinitesimales de X** por:

$$P_{ij}(h) = \begin{cases} q_{ij}h + o(h), & j \neq i \\ 1 - \nu_i h + o(h), & j = i, \end{cases}$$

en donde el término $o(h)$ indica que $\lim_{h \rightarrow 0} \frac{o(h)}{h} = 0$.

Algunos ejemplos de cadenas de Markov continuas y homogéneas en el tiempo serán dados a continuación.

(a) **Proceso de nacimiento y muerte**

Definición 1.3.8. Sea $X = \{X_t : t \in [0, \infty)\}$ una cadena de Markov continua, homogénea en el tiempo con espacio de estados $S = \{0, 1, \dots\}$. Se dice que X es un **proceso de nacimiento y muerte** si para $i, j \in S$, $i \neq j$, se tiene que:

$$q_{ij} = 0 \quad \text{siempre que } |i - j| > 1.$$

Observación: Note que un proceso de nacimiento y muerte es una cadena de Markov, en donde las únicas transiciones que pueden tener probabilidad positiva de ocurrir se dan del estado i al $i + 1$ o al $i - 1$. Cuando pasa lo primero se dice que ocurre un **nacimiento** y en el otro caso ocurre una **muerte**, por lo cual se refiere a las tasas de transición como tasa de nacimiento y muerte.

Definición 1.3.9. Sea $X = \{X_t : t \in [0, \infty)\}$ un proceso de nacimiento y muerte, con espacio de estados S , tasa de transición q_{ij} , $i, j \in S$, $i \neq j$. Para $i \in S$, se definen la **tasa de nacimiento** λ_i y la **tasa de muerte** μ_i de X como sigue:

(a) Cuando $S = \{0, 1, \dots, N\}$

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

con $\lambda_N = \mu_0 = 0$.

(b) Si $S = \{0, 1, \dots\}$

$$\begin{aligned}\lambda_i &= q_{i, i+1}, \quad \text{para toda } i \in S, \\ \mu_i &= q_{i, i-1}, \quad i = 1, 2, \dots,\end{aligned}$$

con $\mu_0 = 0$.

(c) En el caso que $S = \{\dots, -1, 0, 1, \dots\}$

$$\lambda_i = q_{i, i+1}, \quad \text{y} \quad \mu_i = q_{i, i-1},$$

para toda $i \in S$.

Observación: Como se tiene que $\nu_i = \sum_{j \neq i} q_{ij}$, entonces:

(a) Cuando $S = \{0, 1, \dots, N\}$

$$\nu_i = \lambda_i + \mu_i, \quad \text{con } i = 1, \dots, N-1,$$

con $\nu_N = \lambda_N$ y $\nu_0 = \mu_0$. Entonces por (1.16), se tiene que las probabilidades de transición de la cadena de saltos X^s asociada al proceso de nacimiento y muerte X , están dadas para $i \in S$, por:

$$P_{i, i+1} = \frac{q_{i, i+1}}{\nu_i} = \frac{\lambda_i}{\lambda_i + \mu_i}, \quad i = 0, 1, 2, \dots, N-1 \quad (1.19)$$

$$P_{i, i-1} = \frac{q_{i, i-1}}{\nu_i} = \frac{\mu_i}{\lambda_i + \mu_i}, \quad i = 1, 2, \dots, N. \quad (1.20)$$

(b) Si $S = \{0, 1, \dots\}$

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

con $\nu_0 = \lambda_0$. Análogamente al apartado (a), se tiene que las probabilidades de transición de la cadena de saltos X^s asociada al proceso de nacimiento y muerte X , están dadas para $i \in S$, por:

$$P_{i, i+1} = \frac{\lambda_i}{\lambda_i + \mu_i}, \quad i = 0, 1, 2, \dots$$

$$P_{i, i-1} = \frac{\mu_i}{\lambda_i + \mu_i}, \quad i = 1, 2, \dots,$$

(c) En el caso que $S = \{\dots, -1, 0, 1, \dots\}$

$$\nu_i = \lambda_i + \mu_i, \quad i \in S.$$

Por lo tanto las probabilidades de transición de la cadena de saltos X^* asociada al proceso de nacimiento y muerte X , estarían dadas para $i \in S$, por:

$$P_{i \ i+1} = \frac{\lambda_i}{\lambda_i + \mu_i}, \quad P_{i \ i-1} = \frac{\mu_i}{\lambda_i + \mu_i}, \quad i \in S.$$

Definición 1.3.10. Sea $X = \{X_t : t \in [0, \infty)\}$ un proceso de nacimiento y muerte, con espacio de estados S . Sea $i \in S$, cuando las únicas transiciones que pueden tener probabilidad positiva de ocurrir se dan del estado i al $i+1$, se tiene que a X se le llama proceso de **nacimiento puro**. En el caso que las únicas transiciones que pueden tener probabilidad positiva de ocurrir se den del estado i al $i-1$, a X se le llama proceso de **muerte pura**.

Con lo que respecta al tiempo de permanencia de la cadena en el estado $i \in S$, antes de saltar a un estado $j \in S$, $j \neq i$, se tiene que la distribución de este tiempo de permanencia es exponencialmente distribuido con parámetro $\nu_i = \mu_i + \lambda_i$, por la definición de cadena de Markov continua con espacio de estados S .

Un ejemplo de proceso de nacimiento y muerte será estudiado en el Capítulo 2 de esta tesis.

(b) Proceso de Conteo

Definición 1.3.11. Un proceso estocástico $N = \{N(t) : t \geq 0\}$, se dice que es un **proceso de conteo** si $N(t)$ representa el número total de eventos que han ocurrido al tiempo t . Entonces un proceso de conteo satisface:

- (i) $N(t) \geq 0$, para $t \geq 0$.
- (ii) $N(t)$ es un valor entero, para todo $t \geq 0$.
- (iii) Si $s < t$, con $s, t \geq 0$, entonces $N(s) \leq N(t)$.
- (iv) Para $s < t$, $s, t \geq 0$, $N(t) - N(s)$ es igual al número de eventos ocurridos en el intervalo $(s, t]$

Definición 1.3.12. Se dice que un proceso de conteo posee **incrementos independientes** si el número de eventos que ocurren en intervalos de tiempo disjuntos son independientes. Un proceso de conteo se dice que posee **incrementos homogéneos en el tiempo** si la distribución del número de eventos que ocurren en algún intervalo de tiempo sólo depende de la longitud del intervalo.

(c) Proceso Poisson

Definición 1.3.13. El **proceso Poisson** es una cadena de Markov continua $X = \{X_t : t \in [0, \infty)\}$, con espacio de estados $S = \{0, 1, 2, \dots\}$, con tasa de cambio $\lambda > 0$, en donde las probabilidades de transición infinitesimales están determinadas de la siguiente manera:

- (i) $P(X_{t+h} = k + 1 | X_t = k) = \lambda h + o(h)$, $k \in S$.
- (ii) $P(X_{t+h} = k | X_t = k) = 1 - \lambda h + o(h)$, $k \in S$.
- (iii) $P(X_{t+h} = j | X_t = k) = o(h)$, si $j - k \geq 2$, $j, k \in S$.

Adicionalmente X debe cumplir que $X_0 = 0$.

Considérese una cadena de Markov X con espacio de estados $S = \{0, 1, \dots\}$ en donde el tiempo de permanencia en todos los estados es el mismo, es decir, tome $\nu_i = \nu$ para todos los estados $i \in S$. En este caso la cantidad de tiempo en cada estado durante una visita se distribuye exponencial con tasa ν , se sigue que si $N(t)$ denota el número de transiciones de X en el tiempo t , entonces $N = \{N(t) : t \geq 0\}$ será un proceso Poisson con tasa ν .

Definición 1.3.14. El **proceso de conteo** $N = \{N(t) : t \geq 0\}$, con espacio de estado $S = \{0, 1, 2, \dots\}$ es un **proceso Poisson con tasa** $\lambda \in \mathbb{R}$, con $\lambda > 0$, si:

- (i) $N(0) = 0$.
- (ii) El proceso tiene incrementos independientes.
- (iii) El número de eventos en algún intervalo de longitud t se distribuye Poisson con media λt , es decir, para toda $s, t \geq 0$, se tiene

$$P(N_{t+s} - N_s = n) = e^{-\lambda t} \frac{(\lambda t)^n}{n!}, \quad n = 0, 1, 2, \dots$$

1.3.1 Ecuaciones diferenciales de Kolmogorov

Lema 1.3.2. Sea $X = \{X_t : t \in [0, \infty)\}$, una cadena de Markov continua homogénea en el tiempo con espacio de estados S y $P(t) = (P_{ij}(t))_{i,j \in S}$ su matriz de transición al tiempo t , para $i, j \in S$ se tiene que:

$$(a) \lim_{t \rightarrow 0} \frac{1 - P_{ii}(t)}{t} = \nu_i,$$

$$(b) \lim_{t \rightarrow 0} \frac{P_{ij}(t)}{t} = q_{ij}, \quad i \neq j.$$

Demostración. Ver Karlin y Taylor (1981). □

Teorema 1.3.3. Ecuaciones hacia atrás (backward) de Kolmogorov.

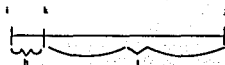
Sea $X = \{X_t : t \in [0, \infty)\}$, una cadena de Markov continua homogénea en el tiempo con espacio de estados S y matriz de transición al tiempo t , $P(t) = (P_{ij}(t))_{i,j \in S}$, entonces para toda $i, j \in S$ y $t \geq 0$, se cumple que:

$$P'_{ij}(t) = \sum_{k \neq i, k \in S} q_{ik} P_{kj}(t) - \nu_i P_{ij}(t). \quad (1.21)$$

Demostración. Sean $i, j \in S$, $t \geq 0$, se tiene que:

$$P'_{ij}(t) = \lim_{h \rightarrow 0} \frac{P_{ij}(t+h) - P_{ij}(t)}{h}$$

Note que uno de los términos involucrados para obtener $P'_{ij}(t)$ es $P_{ij}(t+h)$, el cual puede ser escrito de manera conveniente utilizando las ecuaciones de Chapman-Kolmogorov para obtener las ecuaciones (1.21). Note también que en estas ecuaciones las tasas de transición son del estado i a un estado k . De esta forma se considerará primero la probabilidad de transición en un espacio de tiempo h y después en un espacio de tiempo t , como lo indica la siguiente figura:



Por las ecuaciones de Chapman-Kolmogorov se tiene que:

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

Entonces

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

Ahora si se divide por h , se tiene que:

$$\frac{P_{ij}(t+h) - P_{ij}(t)}{h} = \frac{1}{h} \left[\sum_{k \neq i, k \in S} P_{ik}(h)P_{kj}(t) + (P_{ii}(h) - 1)P_{ij}(t) \right].$$

Ahora se tomará el límite cuando h tiende a cero. Sin embargo como el espacio de estados S de X es finito o contable, se analizarán ambos casos por separado.

Caso 1. Si el espacio de estados S es finito

$$\begin{aligned} \lim_{h \rightarrow 0} \frac{P_{ij}(t+h) - P_{ij}(t)}{h} &= \lim_{h \rightarrow 0} \frac{1}{h} \left(\sum_{k \neq i, k \in S} P_{kj}(t)P_{ik}(h) - (1 - P_{ii}(h))P_{ij}(t) \right) \\ &\stackrel{(1)}{=} \sum_{k \neq i, k \in S} P_{kj}(t) \lim_{h \rightarrow 0} \frac{P_{ik}(h)}{h} - \lim_{h \rightarrow 0} \frac{(1 - P_{ii}(h))}{h} P_{ij}(t) \\ &\stackrel{(2)}{=} \sum_{k \neq i, k \in S} P_{kj}(t)q_{ik} - \nu_i P_{ij}(t). \end{aligned}$$

Caso 2. Si el espacio de estados S es infinito numerable

$$\lim_{h \rightarrow 0} \frac{P_{ij}(t+h) - P_{ij}(t)}{h} = \lim_{h \rightarrow 0} \frac{1}{h} \left(\sum_{k \neq i, k \in S} P_{kj}(t)P_{ik}(h) - (1 - P_{ii}(h))P_{ij}(t) \right). \quad (1.22)$$

Por el Lema 1.3.2 se tiene que existe el siguiente límite:

$$\frac{(1 - P_{ii}(h))}{h},$$

así que sólo se tiene que analizar el primer término de lado derecho de (1.22).

Se tiene por definición que

$$\liminf_{h \rightarrow 0} a_h \leq \limsup_{h \rightarrow 0} a_h.$$

Se dice que $\lim_{h \rightarrow 0} a_h$ existe cuando $\liminf_{h \rightarrow 0} a_h = \limsup_{h \rightarrow 0} a_h$, en tal caso

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

Por lo cual si se demuestra que

$$\liminf_{h \rightarrow 0} \frac{1}{h} \sum_{\substack{k \neq i \\ k \in S}} P_{ik}(h) P_{kj}(t) \geq \sum_{\substack{k \neq i \\ k \in S}} q_{ik} P_{kj}(t) \geq \limsup_{h \rightarrow 0} \frac{1}{h} \sum_{\substack{k \neq i \\ k \in S}} P_{ik}(h) P_{kj}(t),$$

se tendrá lo que se desea.

Tome $N > 0$, N fijo, entonces:

$$\begin{aligned} \liminf_{h \rightarrow 0} \frac{1}{h} \sum_{\substack{k \neq i \\ k \in S}} P_{ik}(h) P_{kj}(t) &\geq \liminf_{h \rightarrow 0} \frac{1}{h} \sum_{k \leq N, k \neq i} P_{ik}(h) P_{kj}(t) \\ &\stackrel{(3)}{=} \sum_{k \leq N, k \neq i} \left(\liminf_{h \rightarrow 0} \frac{P_{ik}(h)}{h} \right) P_{kj}(t) \\ &\stackrel{(4)}{=} \sum_{k \leq N, k \neq i} \lim_{h \rightarrow 0} P_{ik}(h) P_{kj}(t) \\ &\stackrel{(2)}{=} \sum_{k \leq N, k \neq i} q_{ik} P_{kj}(t). \end{aligned}$$

Como se tomó N arbitrario, entonces para todo $N > 0$, vale

$$\liminf_{h \rightarrow 0} \frac{1}{h} \sum_{k \in S, k \neq i} P_{ik}(h) P_{kj}(t) \geq \sum_{k \leq N, k \neq i} q_{ik} P_{kj}(t).$$

Por lo cual

$$\liminf_{h \rightarrow 0} \frac{1}{h} \sum_{k \in S, k \neq i} P_{ik}(h) P_{kj}(t) \geq \sum_{k \in S, k \neq i} q_{ik} P_{kj}(t).$$

(1) Como S es finito, el límite de la suma es la suma de los límites si cada uno de estos existe.

(2) Por el Lema 1.3.2.

(3) Por propiedades de límite inferior.

(4) Cuando un límite existe es igual a su límite inferior.

Por otro lado para $N > 0$, N fijo, $N > i$,

$$\begin{aligned} \sum_{k \neq i, k \in S} \frac{P_{ik}(h)}{h} P_{kj}(t) &= \sum_{k \leq N, k \neq i} \frac{P_{ik}(h)}{h} P_{kj}(t) + \sum_{k \geq N+1} \frac{P_{ik}(h)}{h} P_{kj}(t) \\ &\stackrel{(5)}{\leq} \sum_{k \leq N, k \neq i} \frac{P_{ik}(h)}{h} P_{kj}(t) + \sum_{k \geq N+1} \frac{P_{ik}(h)}{h} \\ &\stackrel{(6)}{=} \sum_{k \leq N, k \neq i} \frac{P_{ik}(h)}{h} P_{kj}(t) + \frac{1}{h} \left(1 - \sum_{k \leq N} P_{ik}(h) \right) \\ &= \sum_{k \leq N, k \neq i} \frac{P_{ik}(h)}{h} P_{kj}(t) + \frac{1 - P_{ii}(h)}{h} - \sum_{k \leq N, k \neq i} \frac{P_{ik}(h)}{h}. \end{aligned}$$

Entonces

$$\begin{aligned} \limsup_{h \rightarrow 0} \sum_{k \neq i, k \in S} \frac{P_{ik}(h)}{h} P_{kj}(t) \\ &\stackrel{(7)}{\leq} \sum_{k \leq N, k \neq i} \limsup_{h \rightarrow 0} \frac{P_{ik}(h)}{h} P_{kj}(t) + \limsup_{h \rightarrow 0} \frac{1 - P_{ii}(h)}{h} - \sum_{k \leq N, k \neq i} \limsup_{h \rightarrow 0} \frac{P_{ik}(h)}{h} \\ &\stackrel{(8)}{=} \sum_{k \leq N, k \neq i} q_{ik} P_{kj}(t) + \nu_i - \sum_{k \leq N, k \neq i} q_{ik}. \end{aligned}$$

Entonces para toda $N > i$, se tiene

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

De nuevo tomando el límite cuando N tiende a infinito, se obtiene lo siguiente:

$$\begin{aligned} \limsup_{h \rightarrow 0} \sum_{k \neq i, k \in S} \frac{P_{ik}(h)}{h} P_{kj}(t) &\leq \sum_{k \in S, k \neq i} q_{ik} P_{kj}(t) + \nu_i - \sum_{k \in S, k \neq i} q_{ik} \\ &\stackrel{(6)}{=} \sum_{k \in S, k \neq i} q_{ik} P_{kj}(t) + \nu_i - \nu_i. \end{aligned}$$

Entonces

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

Por lo tanto

$$\begin{aligned} \limsup_{h \rightarrow 0} \frac{1}{h} \sum_{k \neq i, k \in S} P_{ik}(h) P_{kj}(t) &\leq \sum_{k \neq i, k \in S} q_{ik} P_{kj}(t) \\ &\leq \liminf_{h \rightarrow 0} \frac{1}{h} \sum_{k \neq i, k \in S} P_{ik}(h) P_{kj}(t). \end{aligned}$$

Por lo tanto

$$\lim_{h \rightarrow 0} \frac{1}{h} \sum_{k \neq i, k \in S} P_{ik}(h) P_{kj}(t) = \sum_{k \neq i, k \in S} q_{ik} P_{kj}(t).$$

(5) Es válido porque $0 \leq P_{ij}(t) \leq 1$, $i, j \in S$, $t \geq 0$.

(6) Por propiedades de probabilidad de transición.

(7) Propiedades de límite superior.

(8) Válido por (1.17) □

Teorema 1.3.4. Ecuaciones hacia adelante (forward) de Kolmogorov. Sea $X = \{X_t : t \in [0, \infty)\}$ una cadena de Markov continua homogénea en el tiempo con espacio de estados S y matriz de transición al tiempo t $P(t) = (P_{ij}(t))_{i,j \in S}$, entonces para toda $i, j \in S$ y $t \geq 0$ se cumple que:

$$P'_{ij}(t) = \sum_{k \neq j, k \in S} P_{ik}(t) q_{kj} - \nu_j P_{ij}(t). \quad (1.23)$$

Demostración. Se hace de forma similar al Teorema 1.3.3, pero tomando supuestos extras, ver por ejemplo, Karlin y Taylor (1981), Norris (1997), Ross (1996). □

1.3.2 Las ecuaciones diferenciales de Kolmogorov para proceso de nacimiento y muerte

Ahora se analizarán las ecuaciones diferenciales hacia atrás y hacia adelante de Kolmogorov, para el proceso de nacimiento y muerte.

Tomando a $S = \{0, 1, \dots\}$, se tiene por (1.21) que:

$$\begin{aligned}
 P'_{ij}(t) &= \sum_{k \neq i, k \in S} q_{ik} P_{kj}(t) - \nu_i P_{ij}(t) \\
 &\stackrel{(1)}{=} q_{i, i-1} P_{i-1, j}(t) + q_{i, i+1} P_{i+1, j}(t) - \nu_i P_{ij}(t) \\
 &\stackrel{(2)}{=} \mu_i P_{i-1, j}(t) + \lambda_i P_{i+1, j}(t) - (\lambda_i + \mu_i) P_{ij}(t).
 \end{aligned}$$

Observe que cuando $i = 0$ la ecuación se simplifica ya que por la Definición 1.3.9 $\mu_0 = 0$, de esta forma, se vale lo siguiente:

$$\begin{aligned}
 P'_{0j}(t) &= \lambda_0 P_{1j}(t) - \lambda_0 P_{0j}(t), \\
 P'_{ij}(t) &= \mu_i P_{i-1, j}(t) + \lambda_i P_{i+1, j}(t) - (\lambda_i + \mu_j) P_{ij}(t), \quad i \geq 1.
 \end{aligned}$$

En cuanto a las ecuaciones hacia adelante de Kolmogorov, se tiene por el Teorema 1.3.4 lo siguiente:

$$\begin{aligned}
 P'_{ij}(t) &= \sum_{k \neq j, k \in S} P_{ik}(t) q_{kj} - \nu_j P_{ij} \\
 &\stackrel{(1)}{=} q_{j-1, j} P_{i, j-1}(t) + q_{j+1, j} P_{i, j+1}(t) - \nu_j P_{ij} \\
 &\stackrel{(2)}{=} \lambda_{j-1} P_{i, j-1}(t) + \mu_{j+1} P_{i, j+1}(t) - (\lambda_j + \mu_j) P_{ij}(t).
 \end{aligned}$$

De forma análoga a las ecuaciones hacia atrás de Kolmogorov, cuando $i = 0$ la ecuación se simplifica, por lo tanto se vale:

$$\begin{aligned}
 P'_{i0}(t) &= \mu_1 P_{i1}(t) - \lambda_i P_{i0}(t), & (1.24) \\
 P'_{ij}(t) &= \lambda_{i-1} P_{i, j-1}(t) + \mu_{j+1} P_{i, j+1}(t) - (\lambda_j + \mu_j) P_{ij}(t), \quad i \geq 1 & (1.25)
 \end{aligned}$$

(1) Como X es un proceso de nacimiento y muerte, se tiene que para $i \neq k$, $i, k \in S$, $q_{ik} = 0$, cuando $|i - k| > 1$.

(2) Es válido por la Definición 1.3.9.

Una de las razones por las cuales las ecuaciones diferenciales de Kolmogorov son importantes es que a partir de ellas se puede obtener las probabilidades de transición al tiempo t de una cadena de Markov continua, como se podrá observar en el siguiente ejemplo.

1.3.3 Caso especial: Cadena de Markov de dos estados

Considérese una cadena de Markov continua homogénea en el tiempo, con espacio de estados $S = \{0, 1\}$, donde el tiempo de permanencia en el estado 0 antes de hacer una transición al estado uno es exponencial con tasa λ , y el tiempo de permanencia en estado uno antes de entrar al estado 0 es exponencial con tasa μ , entonces

$$q_{10} = \mu, \quad q_{01} = \lambda,$$

dado que $\sum_{j \neq i} q_{ij} = \nu_i$, se tiene que:

$$\nu_0 = \sum_{j \neq 0} q_{0j} = q_{01} = \lambda \quad y \quad \nu_1 = \sum_{j \neq 1} q_{1j} = q_{10} = \mu.$$

Observe que esta cadena es un proceso de nacimiento y muerte por lo cual las ecuaciones hacia adelante de Kolmogorov están dadas por:

$$\begin{aligned} P'_{00}(t) &\stackrel{(1)}{=} \mu P_{01}(t) - \lambda P_{00}(t) \\ &\stackrel{(2)}{=} \mu(1 - P_{00}(t)) - \lambda P_{00}(t) \\ &= -(\lambda + \mu)P_{00}(t) + \mu, \end{aligned}$$

ya que $P_{01}(t) = 1 - P_{00}(t)$. Por lo cual para obtener $P_{00}(t)$ se resolverá la ecuación diferencial ordinaria

$$P'_{00}(t) = -(\lambda + \mu)P_{00}(t) + \mu,$$

de la cual se tiene que:

$$P'_{00}(t) + (\lambda + \mu)P_{00}(t) = \mu.$$

De esta forma multiplicando ambos lados la ecuación por $e^{(\lambda+\mu)t}$, se tiene que:

$$e^{(\lambda+\mu)t} [P'_{00}(t) + (\lambda + \mu)P_{00}(t)] = \mu e^{(\lambda+\mu)t}. \quad (1.26)$$

Note que:

$$\begin{aligned} \frac{d}{dt} [e^{(\lambda+\mu)t} P_{00}(t)] &= e^{(\lambda+\mu)t} \frac{d}{dt} [P_{00}(t)] + P_{00}(t) \frac{d}{dt} [e^{(\lambda+\mu)t}] \\ &= e^{(\lambda+\mu)t} P'_{00}(t) + P_{00}(t) (\lambda + \mu) e^{(\lambda+\mu)t} \\ &= e^{(\lambda+\mu)t} [P'_{00}(t) + (\lambda + \mu)P_{00}(t)]. \end{aligned}$$

Por lo tanto (1.26), se puede escribir como:

$$\frac{d}{dt}[e^{(\lambda+\mu)t} P_{00}(t)] = \mu e^{(\lambda+\mu)t}.$$

Integrando a ambos lados con respecto al tiempo, se tiene que:

$$\int \frac{d}{dt}(e^{(\lambda+\mu)t} P_{00}(t)) dt = \int \mu(e^{(\lambda+\mu)t}) dt$$

y por lo tanto

$$e^{(\lambda+\mu)t} P_{00}(t) = \frac{\mu}{\lambda + \mu} e^{(\lambda+\mu)t} + c,$$

entonces

$$P_{00}(t) = \frac{\mu}{\lambda + \mu} + ce^{-(\lambda+\mu)t}. \quad (1.27)$$

Por (1.15) se tiene que que $P_{00}(0) = 1$, entonces (1.27) evaluada en $t = 0$ es:

$$1 = \frac{\mu}{\lambda + \mu} + c.$$

de donde

$$c = \frac{\lambda}{\lambda + \mu}.$$

Por lo tanto

$$P_{00}(t) = \frac{\mu}{\lambda + \mu} + \frac{\lambda}{\lambda + \mu} e^{-(\lambda+\mu)t}.$$

De forma similar, se obtiene que:

$$P_{11}(t) = \frac{\lambda}{\lambda + \mu} + \frac{\mu}{\lambda + \mu} e^{-(\lambda+\mu)t}.$$

Con respecto a $P_{01}(t)$ y $P_{10}(t)$ quedan determinadas de la siguiente manera:

$$\begin{aligned} P_{01}(t) & \stackrel{(2)}{=} 1 - P_{00}(t) \\ & = 1 - \left(\frac{\mu}{\lambda + \mu} + \frac{\lambda}{\lambda + \mu} e^{-(\lambda+\mu)t} \right) \\ & = \frac{\lambda}{\lambda + \mu} - \frac{\lambda}{\lambda + \mu} e^{-(\lambda+\mu)t} \\ & = \frac{\lambda}{\lambda + \mu} (1 - e^{-(\lambda+\mu)t}). \end{aligned}$$

(1.28)

Y en forma similar, se tiene:

$$P_{10}(t) = \frac{\mu}{\lambda + \mu} (1 - e^{-(\lambda + \mu)t}).$$

(1) Por (1.24).

(2) Por propiedades de probabilidades de transición.

1.3.4 Probabilidades límite

En esta sección se presentará una forma de obtener la probabilidad límite, cuando ésta existe, en el caso de cadenas de Markov continuas. Se empezará con algunas definiciones.

Definición 1.3.15. Sea $X \geq 0$ una variable aleatoria. Se dice que X es reticulada (lattice), si existe un entero $\ell \geq 0$, tal que:

$$\sum_{n=0}^{\infty} P(X = n\ell) = 1,$$

es decir, X es reticulada si sólo toma valores múltiplos de ℓ , para alguna ℓ no negativa. La ℓ más grande que cumple esta propiedad es el período de X . Si X es reticulada y F es la función de distribución de X , entonces se dice que F es reticulada.

Definición 1.3.16. Sea $X = \{X_t : t \in [0, \infty)\}$ un proceso semi-Markov, se dice que el proceso es irreducible si la cadena de saltos $X^s = \{X_n^s : n = 0, 1, 2, \dots\}$ asociada a X es irreducible.

En los siguientes teoremas r_i denotará la esperanza de tiempo de permanencia por el proceso semi-Markov en el estado i antes de pasar a un estado j . Adicionalmente T_{ij} indicará el tiempo entre transiciones sucesivas en el estado i , y $r_{ii} = E[T_{ii}]$ indicará el tiempo esperado de regreso al estado i .

Teorema 1.3.5. Sea $X = \{X_t : t \in [0, \infty)\}$ un proceso semi-Markov homogéneo irreducible con espacio de estados S y con T_{ii} , $i \in S$, no reticulada con media finita, entonces:

$$P(i) = \lim_{t \rightarrow \infty} P(X(t) = i | X(0) = j),$$

existe y es independiente del estado inicial $j \in S$, además para $i \in S$,

$$P(i) = \frac{r_i}{r_{ii}}.$$

Demostración. Ver Ross (1996). □

Teorema 1.3.6. Sea $X = \{X_t : t \in [0, \infty)\}$, una cadena de Markov continua, homogénea en el tiempo, irreducible con espacio de estados S , tal que T_{ii} , $i \in S$, es no reticulada. Supóngase también que la cadena de Markov anidada ó cadena de saltos $X^s = \{X_n^s : n \geq 0\}$ asociada a X , es recurrente positiva y con distribución estacionaria $\pi^s(i)$, $i \in S$, entonces:

$$P(i) = \lim_{t \rightarrow \infty} P(X_t = i | X_0 = j) = \frac{\pi^s(i)r_i}{\sum_{j \in S} \pi^s(j)r_j}.$$

Demostración. Ver Ross (1996). □

Teorema 1.3.7. Sea $X = \{X_t : t \in [0, \infty)\}$ una cadena de Markov continua homogénea en el tiempo, con cadena de saltos $X^s = \{X_n^s : n \geq 0\}$ con espacio de estados S , y suponga que las condiciones del Teorema 1.3.6 son válidas para X , entonces existe el límite de probabilidades:

$$P(j) = \lim_{t \rightarrow \infty} P_{ij}(t),$$

y está dado por

$$P(j) = \frac{\pi^s(j)}{\sum_{i \in S} \frac{\pi^s(i)}{\nu_i}}. \quad (1.29)$$

Demostración. Es inmediata ya que se tiene que se cumplen las condiciones del Teorema 1.3.6, y además $r_i = 1/\nu_i$ para toda $i \in S$, por lo cual

$$P(i) = \lim_{t \rightarrow \infty} P(X_t = i | X_0 = j) = \frac{\pi^s(i)r_i}{\sum_{j \in S} \pi^s(j)r_j} = \frac{\pi^s(i)}{\sum_{j \in S} \frac{\pi^s(j)}{\nu_j}}.$$

□

Proposición 1.3.8. Sea $X = \{X_t : t \in [0, \infty)\}$, una cadena de Markov continua homogénea en el tiempo, con espacio de estados S , y suponga que las condiciones de los Teoremas 1.3.6 y 1.3.7 son válidas para X , entonces:

- (a) $\{P(j)\}_{j \in S}$ cumple condiciones análogas a las de la distribución estacionaria para cadenas discretas, es decir,

- (i) $P(j) \geq 0, j \in S,$
 (ii) $\sum_{j \in S} P(j) = 1,$
 (iii) $P(j) = \sum_{i \in S} P(i)P_{ij}(t), j \in S.$

(b) Se tiene la siguiente identidad:

$$\nu_j P(j) = \sum_{i \in S} \nu_i P(i) P_{ij}, \quad (1.30)$$

o en forma equivalente

$$\nu_j P(j) = \sum_{i \in S} P(i) q_{ij}. \quad (1.31)$$

Demostración. (a) Por (1.29) se tiene que (i) y (ii) son inmediatos, ya que se cumple que $\pi^*(j) \geq 0, j \in S$ y también $\nu_i > 0, i \in S$, entonces:

$$P(j) = \frac{\pi^*(j)}{\sum_{i \in S} \frac{\pi^*(i)}{\nu_i}} \geq 0,$$

y también se tiene que

$$\sum_{j \in S} P(j) = \sum_{j \in S} \frac{\pi^*(j)/\nu_j}{\sum_{i \in S} \pi^*(i)/\nu_i} = \frac{\sum_{j \in S} \pi^*(j)/\nu_j}{\sum_{i \in S} \pi^*(i)/\nu_i} = 1$$

En cuanto a (iii) se tiene que la demostración es análoga a la realizada para las ecuaciones hacia atrás de Kolmogorov.

(b) Como se cumple las condiciones del Teorema 1.3.7 se tiene que:

$$\begin{aligned} \nu_j P(j) &\stackrel{(1)}{=} \frac{\pi^*(j)}{\sum_{k \in S} \pi^*(k)/\nu_k} \stackrel{(2)}{=} \frac{\sum_{i \in S} \pi^*(i) P_{ij}}{\sum_{k \in S} \pi^*(k)/\nu_k} \\ &= \frac{\sum_{i \in S} \frac{\pi^*(i)}{\nu_i} P_{ij} \nu_i}{\sum_{k \in S} \pi^*(k)/\nu_k} = \sum_{i \in S} \frac{\pi^*(i)/\nu_i}{\sum_{k \in S} \pi^*(k)/\nu_k} P_{ij} \nu_i \\ &\stackrel{(1)}{=} \sum_{i \in S} P(i) P_{ij} \nu_i. \end{aligned}$$

Como $q_{ij} = \nu_i P_{ij}$ entonces se tiene la equivalencia

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

(1) Es válido por (1.29).

(2) Es válido ya que $\pi^s(i)$, $i \in S$, es la distribución estacionaria de la cadena de saltos asociada a X . \square

En particular, las probabilidades límite del proceso de nacimiento y muerte se verán en el siguiente teorema.

Teorema 1.3.9. Sea $X = \{X_t : t \geq 0\}$ un proceso de nacimiento y muerte con espacio de estados $S = \{0, 1, \dots\}$, tasa de nacimiento λ_i , $i \in S$ y tasa de muerte μ_i , $i \in S$, dadas como en la Definición 1.3.9. Si se cumple que:

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

entonces las probabilidades límite de X están dadas por:

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

$$P(n) = \frac{\lambda_0 \lambda_1 \dots \lambda_{n-1}}{\mu_1 \mu_2 \dots \mu_n \left(1 + \sum_{m=1}^{\infty} \frac{\lambda_0 \lambda_1 \dots \lambda_{m-1}}{\mu_1 \mu_2 \dots \mu_m} \right)}; \quad n \geq 1.$$

Demostración. Por (1.31) para X es un proceso de nacimiento y muerte, se tiene que:

$$\nu_j P(j) = P(j-1) q_{j-1, j} + P(j+1) q_{j+1, j}, \quad j \in S$$

$$(\lambda_j + \mu_j) P(j) = \lambda_{j-1} P(j-1) + \mu_{j+1} P(j+1), \quad j \in S.$$

Como $S = \{0, 1, \dots\}$, entonces por la Definición 1.3.9 se tiene que $\mu_0 = 0$ y por lo tanto:

$$\lambda_0 P(0) = \mu_1 P(1); \tag{1.32}$$

$$(\lambda_n + \mu_n) P(n) = \lambda_{n-1} P(n-1) + \mu_{n+1} P(n+1), \quad n \geq 1. \tag{1.33}$$

Para obtener las probabilidades límite del proceso de nacimiento y muerte se hará uso de la siguiente igualdad, la cual se demostrará por inducción en n .

$$\lambda_n P(n) = \mu_{n+1} P(n+1) \quad n \geq 0. \quad (1.34)$$

Para $n = 0$ es verdad por (1.32). Se demostrará que (1.34) se cumple para $n = 1$, por (1.33) se tiene que:

$$\begin{aligned} (\lambda_1 + \mu_1)P(1) &= \lambda_0 P(0) + \mu_2 P(2) \\ &= \mu_1 P(1) + \mu_2 P(2), \end{aligned}$$

y por lo tanto

$$\lambda_1 P(1) = \mu_1 P(1) + \mu_2 P(2) - \mu_1 P(1) = \mu_2 P(2).$$

Suponga ahora que se cumple (1.34) para $n = k$, es decir,

$$\lambda_k P(k) = \mu_{k+1} P(k+1).$$

Se demostrará que (1.34) se cumple para $n = k+1$. De nuevo por (1.33), se tiene que:

$$(\lambda_{k+1} + \mu_{k+1})P(k+1) = \lambda_k P(k) + \mu_{k+2} P(k+2),$$

entonces

$$\begin{aligned} \lambda_{k+1} P(k+1) &= \mu_{k+2} P(k+2) + \lambda_k P(k) - \mu_{k+1} P(k+1) \\ &\stackrel{(1)}{=} \mu_{k+2} P(k+2) + \mu_{k+1} P(k+1) - \mu_{k+1} P(k+1) \\ &= \mu_{k+2} P(k+2). \end{aligned}$$

Por lo tanto se tiene que:

$$P(n) = \frac{\lambda_{n-1}}{\mu_n} P(n-1), \quad n \geq 1. \quad (1.35)$$

Utilizando (1.35) se demostrará de nuevo por inducción lo siguiente:

$$P(n) = \frac{\lambda_{n-1} \lambda_{n-2} \dots \lambda_0}{\mu_n \mu_{n-1} \dots \mu_1} P(0), \quad n \geq 1. \quad (1.36)$$

(i) Para $n = 1$ por (1.34) se vale, ya que

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

(ii) Se supone válido para $n = k$, es decir,

$$P(k) = \frac{\lambda_{k-1} \lambda_{k-2} \dots \lambda_0}{\mu_k \mu_{k-1} \dots \mu_1} P(0).$$

(iii) Ahora se demostrará que se vale para $n = k + 1$, se tiene por (1.34)

$$\begin{aligned} P(k+1) &= \frac{\lambda_k}{\mu_{k+1}} P(k) \\ &\stackrel{(1)}{=} \frac{\lambda_k}{\mu_{k+1}} \frac{\lambda_{k-1} \lambda_{k-2} \dots \lambda_0}{\mu_k \mu_{k-1} \dots \mu_1} P(0). \end{aligned}$$

Por lo tanto se cumple (1.36). Por otro lado se tiene que $\sum_{j \in S} P(j) = 1$ por la Proposición 1.3.8, entonces

$$\begin{aligned} \sum_{n=0}^{\infty} P(n) &= P(0) + \sum_{m=1}^{\infty} \frac{\lambda_{m-1} \lambda_{m-2} \dots \lambda_0}{\mu_m \mu_{m-1} \dots \mu_1} P(0) \\ &= P(0) \left(1 + \sum_{m=1}^{\infty} \frac{\lambda_{m-1} \lambda_{m-2} \dots \lambda_0}{\mu_m \mu_{m-1} \dots \mu_1} \right), \end{aligned}$$

entonces

$$1 = P(0) \left(1 + \sum_{m=1}^{\infty} \frac{\lambda_{m-1} \lambda_{m-2} \dots \lambda_0}{\mu_m \mu_{m-1} \dots \mu_1} \right),$$

y como

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

se sigue que

$$\begin{aligned} P(0) &= \frac{1}{1 + \sum_{m=1}^{\infty} \frac{\lambda_0 \lambda_1 \dots \lambda_{m-1}}{\mu_1 \mu_2 \dots \mu_m}} \\ P(n) &= \frac{\lambda_0 \lambda_1 \dots \lambda_{n-1}}{\mu_1 \mu_2 \dots \mu_n} \left(\frac{1}{1 + \sum_{m=1}^{\infty} \frac{\lambda_0 \lambda_1 \dots \lambda_{m-1}}{\mu_1 \mu_2 \dots \mu_m}} \right). \end{aligned}$$

(1) Hipótesis de inducción. □

Observación: Si $\sum_{m=1}^{\infty} \frac{\lambda_{m-1}\lambda_{m-2}\dots\lambda_1\lambda_0}{\mu_m\mu_{m-1}\dots\mu_2\mu_1} = \infty$, entonces $P(0) = 0$ y por lo tanto $P(j) = 0$, para toda $j \in S$. De esta forma, para que se tenga $P(j) > 0$, $j \in S$, es necesario que:

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

Capítulo 2

El modelo de Wright-Fisher y el modelo de Moran

Uno de los temas de gran interés en el estudio de poblaciones en genética es la forma en que se dan los cambios en la composición de poblaciones, por lo cual se han hecho numerosas investigaciones y varios modelos han sido presentados. Entre las aportaciones importantes al respecto se encuentran dos modelos uno introducido por S. Wright y R. Fisher y el otro introducido por Moran, conocidos por el modelo de Wright-Fisher y el modelo de Moran, respectivamente (ver Feller (1951), Karlin y McGregor (1964) y Moran (1958)). Ambos modelos son de nuestro interés pues aún son utilizados en el estudio de la población en genética (ver por ejemplo, la parte desarrollada en el Capítulo 5 de este trabajo). Adicionalmente se tiene que la base matemática de los modelos son las cadenas de Markov.

En este capítulo se describirán primeramente las hipótesis comunes a los dos modelos. Posteriormente se describirá el modelo de Wright-Fisher seguido de la descripción del modelo de Moran.

En ambos modelos se considerará una población cuyos individuos poseen un único progenitor (individuos haploides), que pueden ser por ejemplo los gametos, las células entre otros. Además se considerará un único locus en un cromosoma con dos alelos posibles que se indicará por a y A . En una primera versión de los modelos no se permitirá que un alelo mute a otro, pero posteriormente se permitirá que el alelo a mute al alelo A con probabilidad α_1 y que el alelo A mute al alelo a con probabilidad α_2 , con $0 \leq \alpha_1, \alpha_2 < 1$ (note que $\alpha_1 = \alpha_2 = 0$ corresponde al caso sin mutación). En ambos modelos

la población tendrá tamaño fijo $N \geq 1$. También se asumirá que los modelos son neutrales, es decir, el tipo de alelo del individuo no interviene en su capacidad de reproducción.

Se considerará primero la versión discreta de ambos modelos y posteriormente se pasará a la versión continua del modelo de Moran.

El objetivo de estos modelos es obtener información acerca de la composición alelica de la población a medida que el tiempo pasa, es decir, se desea saber cuanto alelos existen de cada tipo después de algunas generaciones o después de un cierto número de eventos de reproducción. Sin embargo se debe notar que si se tiene el tamaño de la población y el número de individuos de tipo **a**, se obtiene automáticamente el número de individuos del otro tipo (el de tipo **A**), por lo cual se concentrará en contar solo el número de individuos de tipo **a** después de cada evento de reproducción.

Para el caso discreto en los dos modelos se denotará como X_n , $n = 0, 1, 2, \dots$ a la variable aleatoria que indica el número de individuos de tipo **a** después del n -ésimo evento de reproducción. Para el caso continuo X_t , $t \geq 0$, indicará el número de individuos de tipo **a** al tiempo t .

Bajo las hipótesis que siguen los procesos demográfico y de mutación en ambos modelos hacen que el proceso $X = \{X_n : n = 0, 1, \dots\}$ y $X = \{X_t : t \geq 0\}$ sea una cadena de Markov homogénea en el tiempo con espacio de estados $S = \{0, 1, \dots, N\}$.

2.1 Modelo de Wright-Fisher

En esta sección se describirá el modelo de Wright-Fisher considerando que el tiempo es discreto.

La demografía de la población que se considera para este modelo evolucionaria de la siguiente forma: considere una generación dada, los individuos de esta población son seleccionados con reemplazo uniformemente e independientemente para producir un descendiente, cuando N descendientes hayan sido producidos, entonces el proceso de reproducción termina y los individuos de la presente generación mueren y son reemplazados por los descendientes.

El proceso de mutación es dado por: un individuo de tipo **a** es seleccionado para reproducirse y produce un descendiente de tipo **A** con probabilidad α_1 , en otro caso el descendiente es de tipo **a**. Por otro lado un individuo de tipo

A produce un descendiente de tipo a con probabilidad α_2 y en otro caso el descendiente tiene tipo A. Se tomará $0 \leq \alpha_1, \alpha_2 \leq 1$.

Sea $X = \{X_n : n = 0, 1, \dots\}$ la cadena de Markov que cuenta el número de individuos de tipo a. Note que X tiene espacio de estados $S = \{0, 1, \dots, N\}$.

a) Modelo sin mutación

Considere primeramente el caso donde no se permite la mutación, es decir, el caso donde $\alpha_1 = \alpha_2 = 0$. A continuación se obtendrán las probabilidades de transición de X.

Este proceso se puede ilustrar a través de la siguiente figura:

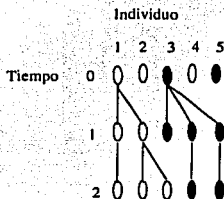


Figura 2.1: El color negro indica un individuo de tipo a y el color blanco un individuo de tipo A. En este caso se tiene que sólo el 1er y 3er gen se reprodujeron en la primera transición (generación de 0 a 1), y para la segunda (generación 1 a 2) todos se reprodujeron excepto uno.

Considere una generación dada, se dice que el sistema al tiempo n ($n = 0, 1, 2, \dots$) está en el estado j , si se tienen j individuos de tipo a al tiempo n .

Si esta generación progenitora tiene exactamente j individuos de tipo a (con $j \in S$), entonces en cada selección la probabilidad p_j de que un individuo de tipo a sea seleccionado para reproducirse (lo cual significa que un individuo de tipo a sea producido), está dada por:

$$p_j = \frac{j}{N}, \quad (2.1)$$

y la probabilidad $q_j, j \in S$, para que sea de tipo A (es decir, un individuo de tipo A sea producido) es:

$$q_j = \frac{N-j}{N} = 1 - \frac{j}{N} = 1 - p_j, \quad (2.2)$$

dado que los individuos son seleccionados uniformemente para reproducirse.

Suponga que al tiempo n se tienen j individuos de tipo a , entonces las probabilidades de transición quedan determinadas de la siguiente forma:

$$\begin{aligned} P_{jk} &= P(X_{n+1} = k | X_n = j) = \binom{N}{k} p_j^k q_j^{N-k} \\ &= \binom{N}{k} \left(\frac{j}{N}\right)^k \left(\frac{N-j}{N}\right)^{N-k}, \end{aligned} \quad (2.3)$$

para $j, k \in S = \{0, 1, \dots, N\}$.

Observación: La forma de las probabilidades de transición dadas en (2.3) se debe a que si se considera un suceso como la selección de un individuo de tipo a para que se reproduzca y se tienen j ($j \in S$) individuos de tipo a para elegir, entonces éste es un evento Bernoulli con probabilidad del suceso p_j , dada en (2.1). Por lo tanto dado que los individuos son seleccionados independientemente y con remplazo, hasta que N descendientes son obtenidos, se tienen N ensayos de Bernoulli con probabilidad del suceso p_j , $j \in S$. De esta forma la probabilidad de que se tengan k individuos de tipo a en la siguiente generación queda determinada por la Binomial (N, p_j) , pues equivale a tener k sucesos en N ensayos independientes de Bernoulli con probabilidad de suceso p_j .

Teorema 2.1.1. Sea $X = \{X_n : n = 0, 1, 2, \dots\}$ la cadena de Markov con espacio de estado $S = \{0, 1, 2, \dots, N\}$ y probabilidades de transición P_{ik} , $i, k \in S$, dadas por (2.3), entonces los estados 0 y N son los únicos estados absorbentes y $1, 2, 3, \dots, N-1$ son estados transitorios.

Demostración. Note que las probabilidades de transición (2.3) de la cadena X están dadas por la distribución Binomial (N, p_j) y por tanto cuando se suman sobre todos los valores de $k \in S$, se tiene que:

$$\sum_{k \in S} P_{jk} = \sum_{k=0}^N P_{jk} = 1,$$

entonces

$$P_{00} = 1 - \sum_{i=1}^N P_{0i},$$

y por (2.1) para $j \in S$, se tiene que $p_0 = 0$, entonces para todo $i = 1, 2, \dots, N$, se tiene que:

$$P_{0i} = \binom{N}{i} p_0^i q_0^{N-i} = 0,$$

es decir, $P_{0i} = 0$ para todo $i \neq 0$ y por lo tanto $P_{00} = 1$. De esta forma por la Definición 1.2.3 se tiene que el estado 0 es un estado absorbente.

De forma análoga se demuestra que N es un estado absorbente pero ahora se utiliza el hecho que $q_N = 0$.

Como 0 es un estado absorbente entonces por definición, se tiene que para $i \in S$, $i \neq 0, N$, el estado i no es accesible desde 0, pero éste estado si es accesible desde i , con $0 < i < N$. La razón es la siguiente:

$$P_{i0} = \binom{N}{0} \left(\frac{i}{N}\right)^0 \left(\frac{N-i}{N}\right)^N = \left(\frac{N-i}{N}\right)^N > 0, \quad 0 < i < N.$$

Finalmente como el estado 0 es accesible desde j , $j = 1, 2, \dots, N - 1$ y j no es accesible desde 0, entonces se tiene probabilidad positiva de no retornar al estado j , por lo tanto el estado j ($1 \leq j \leq N - 1$) es transitorio. \square

b) Modelo con mutación

A continuación se considerará el caso donde se permite la mutación, es decir, el caso donde $\alpha_1 > 0$ y $\alpha_2 > 0$. Este proceso se puede ilustrar a través de la siguiente figura:

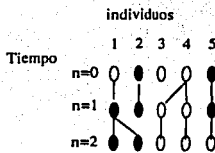


Figura 2.2: El color negro indica un individuo de tipo a y el color blanco un individuo de tipo A, igual que en la Figura 2.1. En este caso se reprodujeron todos excepto uno en la primera transición (generación de 0 a 1) y ocurrió una mutación en el 1er. gen. Para la segunda transición (generación 1 a 2) también ocurrió una mutación en el 5to. gen y no se reprodujo el segundo gen.

Se indicará nuevamente a p_j y q_j , $j \in S$, a la probabilidad de que un descendiente de tipo a y tipo A sean producidos respectivamente.

En este caso se tiene que cuando existen j , $j = 0, 1, \dots, N$, individuos de tipo a en la población vale:

$$p_j = \frac{j(1 - \alpha_1) + (N - j)\alpha_2}{N}, \quad (2.4)$$

$$q_j = \frac{j\alpha_1 + (N - j)(1 - \alpha_2)}{N}. \quad (2.5)$$

Observación: La forma de p_j , $j \in S$, se debe a que el evento de producir un gen (individuo) de tipo a es la unión del evento:

$B = \{\text{un individuo de tipo } a \text{ seleccionado para reproducirse y reproduce un individuo no mutante}\}$,
con el evento:

$C = \{\text{un individuo de tipo } A \text{ seleccionado para reproducirse y reproduce un individuo mutante}\}$.

Note que los eventos B y C son mutuamente excluyentes dado que a cada selección de individuos para reproducir solamente uno es seleccionado. Además, se tiene que para j , $j = 0, 1, \dots, N$, vale

$$P(B) = \frac{j}{N}(1 - \alpha_1) \quad \text{y} \quad P(C) = \frac{N - j}{N}\alpha_2.$$

Por tanto $p_j = P(B) + P(C)$.

Para el caso de q_j , $j \in S$, el argumento es análogo.

Si se considera un suceso como la reproducción de un descendiente de tipo a y si se tienen j , $j = 0, 1, 2, \dots, N$ individuos en la generación progenitora, entonces p_j es la probabilidad de que un suceso ocurra. De esta forma por la misma justificación dada en el caso sin mutación, se tiene que las probabilidades de transición para X son:

$$P_{jk} = P(X_{n+1} = k | X_n = j) = \binom{N}{k} (p_j)^k (q_j)^{N-k}, \quad j, k \in S. \quad (2.6)$$

donde p_j y q_j están determinadas por (2.4) y (2.5) respectivamente.

Teorema 2.1.2. *La cadena de Markov $X = \{X_n : n = 0, 1, 2, \dots\}$ con espacio de estados $S = \{0, 1, 2, \dots, N\}$ y probabilidades de transición dadas por (2.6), es irreducible, aperiódica y recurrente positiva.*

Demostración. Para demostrar que X es irreducible se tiene que demostrar por la Definición 1.2.6 que todos los estados se comunican, es decir, que para $i, j \in S$ existe una $m \geq 0$, tal que $P_{ij}^{(m)} > 0$. En este caso se demostrará que para $m = 1$ se cumple, es decir, $P_{ij} > 0$, para todo $i, j \in S$. Por (2.6) se tiene que

$$P_{jk} = \binom{N}{k} (p_j)^k (q_j)^{N-k}, \quad j, k \in S.$$

Note que $\binom{N}{k} > 0$ y si $j = 1, 2, \dots, N-1$, entonces

$$\frac{j(1-\alpha_1)}{N} > 0 \quad \text{y} \quad \frac{N-j}{N} > 0,$$

entonces

$$p_j = \frac{j(1-\alpha_1) + (N-j)\alpha_2}{N} > 0, \quad j = 1, 2, \dots, N-1.$$

Para $j = 0$ se tiene que $p_0 = \alpha_2 > 0$ y para $j = N$ se tiene que $p_j = \alpha_1 > 0$, así que

$$p_j = \frac{j(1-\alpha_1) + (N-j)\alpha_2}{N} > 0, \quad j \in S.$$

Similarmente, se argumenta que $q_j > 0$, $j \in S$. Entonces $P_{jk} > 0$, para todo $j, k \in S$. Por lo tanto X es irreducible.

Por lo anterior $P_{jk} > 0$, $j, k \in S$, en particular $P_{jj} > 0$, $j \in S$, entonces $d(j) = 1$ y por lo tanto X es aperiódica por la Definición 1.2.8.

Por último como X es irreducible y aperiódica y además el espacio de estados S es finito, entonces por el Teorema 1.2.11 se tiene que X es recurrente positiva. \square

2.2 Modelo de Moran

El desarrollo para este modelo será similar al anterior. Se describirá primero el modelo en el caso discreto y posteriormente se describirá el modelo a tiempo continuo.

a) Tiempo discreto

La evolución de la demografía de una población que se reproduce de acuerdo con el modelo de Moran es la siguiente: a cada unidad de tiempo $n = 0, 1, 2, \dots$ un individuo es seleccionado uniformemente, independientemente y con reemplazo de la población para producir un descendiente y un individuo es seleccionado uniformemente e independientemente para morir. El descendiente producido reemplaza el individuo que muere. Con lo que respecta al proceso de mutación es idéntico al descrito para el modelo de Wright-Fisher.

Sea $X = \{X_n : n = 0, 1, 2, \dots\}$ la cadena de Markov con espacio de estados $S = \{0, 1, \dots, N\}$, que al tiempo n indica cuántos individuos de tipo a tiene la población después de n eventos de reproducción. De esta forma si al tiempo n existen i individuos de tipo a , entonces después del próximo evento de reproducción las únicas posibilidades para el número de individuos de tipo a son i , $i + 1$ ó $i - 1$. A continuación se describirán las probabilidades de transición para X .

(a.1) Modelo sin mutación

El caso sin mutación, como ya se aludió es donde $\alpha_1 = \alpha_2 = 0$. La ilustración del proceso sin mutación queda de la siguiente manera:

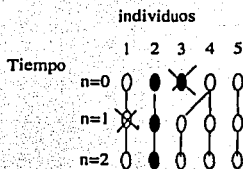


Figura 2.3: Como en las figuras anteriores los ovals negros representan individuos de tipo a . En este caso en la primera transición (generación de 0 a 1) el individuo tres es seleccionado para morir el cual es reemplazado por un descendiente del individuo cuatro. En cambio en la segunda transición (generación de 1 a 2) el mismo individuo que es elegido para reproducirse es elegido para morir.

En cuanto a las probabilidades de transición del proceso están dadas como sigue:

$$P_{j, j+1} = \frac{(N-j)j}{N^2}, \quad \text{si } j = 0, 1, \dots, N-1 \quad (2.7)$$

$$P_{j, j-1} = \frac{j(N-j)}{N^2}, \quad \text{si } j = 1, \dots, N \quad (2.8)$$

$$P_{jj} = \left(\frac{N-j}{N}\right)^2 + \left(\frac{j}{N}\right)^2, \quad \text{si } j \in S \quad (2.9)$$

$$P_{jk} = 0, \quad \text{si } |k-j| > 1, j, k \in S. \quad (2.10)$$

Para ver porque (2.7) es válido, basta notar que la cadena pasa del estado j al $j+1$ si y sólo si un individuo de tipo **A** es seleccionado para morir y un individuo de tipo **a** es seleccionado para reproducirse. Note que

$$P(\text{individuo de tipo a es seleccionado para reproducirse}) = \frac{j}{N},$$

y que

$$P(\text{individuo de tipo A es seleccionado para morir}) = \frac{N-j}{N},$$

y dado que la selección de individuos es independiente el resultado se cumple. La justificación es análoga para (2.8). Para el caso donde no hay cambios, es decir, para el caso de (2.9), note que no hay cambios en el número de individuos de tipo **a**, lo cual sucede cuando el individuo que es seleccionado para reproducirse es del mismo tipo que el seleccionado para morir, es decir, el evento:

$D = \{\text{en la generación } n+1 \text{ existen } j \text{ individuos a, dado que en la generación } n \text{ existen } j \text{ individuos}\},$

es la unión de

$B = \{\text{un individuo de tipo a es seleccionado para reproducir y un individuo de tipo a es seleccionado para morir}\}$

y

$C = \{\text{un individuo de tipo A es seleccionado para reproducir y un individuo de tipo A es seleccionado para morir}\}.$

Note que estos eventos son mutuamente excluyentes y además:

$$P(B) = \binom{j}{N} \frac{j}{N} \quad \text{y} \quad P(C) = \binom{N-j}{N} \frac{N-j}{N},$$

y el resultado sigue.

Teorema 2.2.1. *La cadena de Markov $X = \{X_n : n = 0, 1, \dots\}$ que registra el número de individuos de tipo a en una población que evoluciona de acuerdo con el modelo de Moran sin mutación, con probabilidades de transición dadas por (2.7), (2.8), (2.9) y (2.10), es tal que los estados 0 y N son absorbentes y $1, 2, \dots, N-1$, son estados transitorios.*

Demostración. Primero se demostrará que los estados i , $0 < i < N$, forman una clase, es decir, estos estados se comunican entre sí. Sean $i, j \in S$, $0 < i, j < N$. Note que para $i = 1, 2, \dots, N-1$, se tiene que $P_{i, i-1}$, $P_{i, i+1}$ son estrictamente positivos. Suponga que $i < j$ entonces existe una $m > 0$, tal que $i + m = j$. Se demostrará por inducción que:

$$P_{ij}^{(m)} \geq \prod_{k=0}^{m-1} P_{i+k, i+k+1} > 0. \quad (2.11)$$

Para $j = i+1$ (es decir, para $m = 1$) el resultado es inmediato. Suponga que vale para $j = i+k$ (es decir, $m = k$) y se demostrará que vale para $j = i+k+1$ (es decir, $m = k+1$). Por las ecuaciones de Chapman Kolmogorov, se tiene que

$$\begin{aligned} P_{ij}^{(k+1)} &\stackrel{(1)}{=} \sum_{l \in S} P_{il}^{(k)} P_{lj} \stackrel{(2)}{\geq} P_{ii+k}^{(k)} P_{i+k, i+k+1} \\ &\stackrel{(3)}{\geq} \left(\prod_{n=0}^{k-1} P_{i+n, i+n+1} \right) P_{i+k, i+k+1} \\ &= \prod_{n=0}^k P_{i+n, i+n+1} > 0. \end{aligned}$$

Por lo tanto el estado i se comunica con el estado j , para todo $i, j \in S$, tal que $0 < i, j < N$.

Por otro lado se tiene por (2.9) que $P_{00} = 1$ y $P_{NN} = 1$. De nuevo por la Definición 1.2.3 los estados 0 y N son absorbentes, por lo cual el estado

i ($0 < i < N$) no es accesible desde cero. En particular el estado $i = 1$ no es accesible desde 0, sin embargo el inverso si se cumple, es decir, 0 si es accesible desde 1 entonces el estado $i = 1$ es transitorio. Como $i = 1$ se comunica con i ($0 < i < N$), entonces por la Proposición 1.2.9 el estado i ($0 < i < N$) es transitorio.

- (1) Por las ecuaciones de Chapman-Kolmogorov.
- (2) Es válido ya que $P_{ij}^{(n)} \geq 0$, para todo $i, j \in S$, $n \geq 0$.
- (3) Por hipótesis de inducción. □

Observación: En el modelo de Moran una generación equivaldría a N eventos de nacimiento y muerte, aunque usualmente no todos los N individuos son reemplazados en este tiempo.

(a.2) Modelo con mutación

Para el caso con mutación (es decir cuando α_1 y α_2 son estrictamente positivos), el proceso queda ilustrado de la siguiente manera:

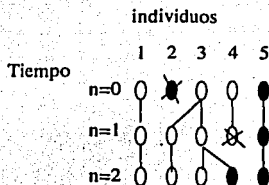


Figura 2.4: De nuevo los ovalos negros representan los individuos de tipo a y los blancos los individuos de tipo A . En esta figura en la primera transición (generación de 0 a 1) el individuo dos es seleccionado para morir el cual es reemplazado por un descendiente del individuo tres. En cambio en la segunda transición (generación de 1 a 2) el individuo 3 es seleccionado para reproducirse y produce un mutante que reemplaza al individuo cuatro.

En lo que se refiere a las probabilidades de transición de X son dadas por:

$$P_{j, j+1} = P(X_{n+1} = j+1 | X_n = j) \\ = \frac{j(N-j)}{N^2} (1 - \alpha_1) + \left(\frac{N-j}{N}\right)^2 \alpha_2, \quad j = 0, \dots, N-1 \quad (2.12)$$

$$P_{j, j-1} = P(X_{n+1} = j-1 | X_n = j) \\ = \frac{j(N-j)}{N^2} (1 - \alpha_2) + \left(\frac{j}{N}\right)^2 \alpha_1, \quad j = 1, \dots, N \quad (2.13)$$

$$P_{jj} = P(X_{n+1} = j | X_n = j) \\ = \left(\frac{N-j}{N}\right)^2 (1 - \alpha_2) + \left(\frac{j}{N}\right)^2 (1 - \alpha_1) \\ + \frac{j(N-j)}{N^2} (\alpha_1 + \alpha_2), \quad j \in S. \quad (2.14)$$

$$P_{jk} = P(X_{n+1} = j-1 | X_n = j) = 0, \quad |k-j| > 1, \quad j, k \in S. \quad (2.15)$$

Observación: La probabilidad de transición dada por (2.12) está determinada por la probabilidad de que haya un incremento en los individuos de tipo **A** y esto se da de dos formas: una es la de que se elija un individuo de tipo **A** para morir y se elija del otro tipo para que se reproduzca y no mute y la otra posibilidad es que se elija el tipo **A** para que se reproduzca y mute y se elige otro individuo de tipo **A** para que muera (puede ser el mismo). De forma similar se obtiene las otras probabilidades de transición.

El teorema para el modelo de Moran análogo al Teorema 2.1.2 es el siguiente.

Teorema 2.2.2. *La cadena de Markov $X = \{X_n : n = 0, 1, 2, \dots\}$ con espacio de estados $S = \{0, 1, 2, \dots, N\}$ y probabilidades de transición dadas por (2.12), (2.13), (2.14) y (2.15) es irreducible, aperiódica y recurrente positiva.*

Demostración. Por (2.14) se tiene que $P_{jj} > 0$, $j \in S$, entonces X es aperiódica. Por otro lado se tiene que cada estado i se comunica con sus estados contiguos, es decir, con $i-1$ y con el $i+1$. Por lo tanto se tiene que todos los estados se comunican, por la misma justificación dada para obtener (2.11). Finalmente como el espacio de estados es finito, entonces por el Teorema 1.2.11, X es recurrente positiva. \square

b) Tiempo continuo

Ahora se analizará el caso en el tiempo continuo. Dado que el tiempo no influye en la elección de un individuo, en el tiempo continuo el proceso de demografía y el proceso de mutación son los mismos que para el tiempo discreto, pero ahora los eventos de nacimiento y muerte ocurren en intervalos de tiempo exponencialmente distribuidos con parámetro $\nu_i > 0$, $i \in S$. Adicionalmente se asume que la probabilidad de que más de un individuo muera o nazca en un intervalo de tiempo $[t, t + h)$ es $o(h)$, donde $\lim_{h \rightarrow 0} \frac{o(h)}{h} = 0$. En este caso, la cadena de Markov en tiempo continuo $X = \{X_t : t \geq 0\}$, que cuenta el número de individuos de tipo \mathbf{a} , permanece en un estado un tiempo τ_i , tal que $P(\tau_i > t) = e^{-\nu_i t}$, $t \geq 0$.

Dado que el modelo utilizado en los Capítulos 4 y 5 de esta tesis es el de con mutación, entonces se fijará atención solamente en el modelo con mutación.

Las probabilidades de transición infinitesimales son dadas para $i \in S$, por

$$\begin{aligned}
 P(X(t+h) = j+1 | X(t) = j) &= \lambda \left(\frac{j(N-j)}{N^2} (1 - \alpha_1) + \left(\frac{N-j}{N} \right)^2 \alpha_2 \right) h + o(h) \\
 P(X(t+h) = j-1 | X(t) = j) &= \lambda \left(\frac{j(N-j)}{N^2} (1 - \alpha_2) + \left(\frac{j}{N} \right)^2 \alpha_1 \right) h + o(h) \\
 P(X(t+h) = j | X(t) = j) &= \lambda \left[\left(\frac{N-j}{N} \right)^2 (1 - \alpha_2) + \left(\frac{j}{N} \right)^2 (1 - \alpha_1) \right. \\
 &\quad \left. + \frac{j(N-j)}{N^2} (\alpha_1 + \alpha_2) \right] h + o(h).
 \end{aligned} \tag{2.16}$$

Las justificaciones para las formas que tienen estas probabilidades de transición son similares a las que condujeron a la obtención de las probabilidades para el caso discreto (ver Donnelly (1984)).

Entonces para el caso con mutación las tasas de transición de X están dadas por:

$$\lambda_i = \lambda \frac{i(N-i)(1-\alpha_1) + (N-i)^2\alpha_2}{N^2} \quad i = 0, 1, \dots, N-1$$

$$\mu_i = \lambda \frac{i(N-i)(1-\alpha_2) + i^2\alpha_1}{N^2} \quad i = 1, 2, \dots, N$$

Por lo tanto de nuevo por (1.19) y (1.20), las probabilidades de transición de la cadena de saltos asociada a X , están dadas de la siguiente manera:

$$P_{i \rightarrow i+1} = \frac{\lambda_i}{\lambda_i + \mu_i}$$

$$= \frac{i(N-i)(1-\alpha_2) + i^2\alpha_1}{i(N-i)(2-\alpha_1-\alpha_2) + (N-i)^2\alpha_2 + i^2\alpha_1}, \quad i = 0, 1, 2, \dots, N-1$$

$$P_{i \rightarrow i-1} = \frac{\mu_i}{\lambda_i + \mu_i}$$

$$= \frac{i(N-i)(1-\alpha_1) + (N-i)^2\alpha_2}{i(N-i)(2-\alpha_1-\alpha_2) + (N-i)^2\alpha_2 + i^2\alpha_1}, \quad i = 1, 2, \dots, N.$$

Antes de obtener la expresión para la distribución estacionaria de la cadena $X = \{X_t : t \geq 0\}$, note que las tasas de transición λ_j , $j = 0, 1, \dots, N-1$ y μ_j , $j = 1, \dots, N$, son tales que

$$\begin{aligned} \lambda_j &= \lambda \frac{(N-j)}{N} \left[\frac{j}{N}(1-\alpha_1) + \left(\frac{N-j}{N} \right) \alpha_2 \right] \\ &= \lambda \frac{(N-j)}{N} \left[\frac{j}{N} - \frac{j}{N}\alpha_1 - \frac{j}{N}\alpha_2 + \alpha_2 \right] \\ &= \lambda \frac{(N-j)}{N} \left[\alpha_2 + \frac{j}{N}(1-\alpha_1-\alpha_2) \right], \end{aligned} \quad (2.17)$$

$$\begin{aligned}
 \mu_j &= \lambda \frac{j}{N} \left[\left(\frac{N-j}{N} \right) (1 - \alpha_2) + \frac{j}{N} \alpha_1 \right] \\
 &= \lambda \frac{j}{N} \left[\frac{N-j}{N} - \left(\frac{N-j}{N} \right) \alpha_2 + \frac{j}{N} \alpha_1 \right] \\
 &= \lambda \frac{j}{N} \left[1 - \alpha_2 - \frac{j}{N} + \frac{j}{N} \alpha_2 + \frac{j}{N} \alpha_1 \right] \\
 &= \lambda \frac{j}{N} \left[1 - \alpha_2 - \frac{j}{N} (1 - \alpha_1 - \alpha_2) \right]. \tag{2.18}
 \end{aligned}$$

Teorema 2.2.3. Sea $X = \{X_t : t \geq 0\}$ la cadena de Markov continua con espacio de estados $S = \{0, 1, \dots, N\}$, que describe el comportamiento de una población que evoluciona de acuerdo con el modelo de Moran con mutación. Entonces la distribución estacionaria de X y consecuentemente su distribución límite está dada por:

$$P(k) = \frac{\Gamma(N+1)}{\Gamma(k+1)\Gamma(N-k+1)} \frac{\Gamma\left(\frac{N(\alpha_1+\alpha_2)}{1-\alpha_1-\alpha_2}\right)\Gamma\left(\frac{N\alpha_2}{1-\alpha_1-\alpha_2} + k\right)\Gamma\left(\frac{N(1-\alpha_2)}{1-\alpha_1-\alpha_2} - k\right)}{\Gamma\left(\frac{N\alpha_2}{1-\alpha_1-\alpha_2}\right)\Gamma\left(\frac{N}{1-\alpha_1-\alpha_2}\right)\Gamma\left(\frac{N\alpha_1}{1-\alpha_1-\alpha_2}\right)}, \tag{2.19}$$

para todo $k \in S$, donde

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

Demostración. Dado que X es un proceso de nacimiento y muerte. Por el Teorema 1.3.9, se tiene que la distribución límite de X , cuando ésta existe, esta dada por:

$$\begin{aligned}
 P(0) &= \frac{1}{1 + \sum_{n=1}^N \frac{\lambda_0 \lambda_1 \dots \lambda_{n-1}}{\mu_1 \mu_2 \dots \mu_n}} \\
 P(j) &= \frac{\lambda_0 \lambda_1 \dots \lambda_{j-1}}{\mu_1 \mu_2 \dots \mu_j \left(1 + \sum_{n=1}^N \frac{\lambda_0 \lambda_1 \dots \lambda_{n-1}}{\mu_1 \mu_2 \dots \mu_n} \right)}. \tag{2.21}
 \end{aligned}$$

Note que el caso del modelo de Moran con mutación la cadena X es irreducible y recurrente positiva, adicionalmente dado que el espacio de estados S es

finito, se tiene que:

$$1 + \sum_{n=1}^N \left(\frac{\lambda_0 \lambda_1 \dots \lambda_{n-1}}{\mu_1 \mu_2 \dots \mu_n} \right) < \infty.$$

Por lo tanto basta verificar que las tasas de nacimiento y muerte producen la expresión (2.19).

Por (2.17) se tiene que

$$\begin{aligned} \prod_{k=0}^{j-1} \lambda_k &= \prod_{k=0}^{j-1} \left(\lambda \frac{N-k}{N} \left[\alpha_2 + \frac{k}{N} (1 - \alpha_1 - \alpha_2) \right] \right) \\ &= \left(\frac{\lambda}{N} \right)^j \left[\prod_{k=0}^{j-1} (N-k) \right] \left[\prod_{k=0}^{j-1} \left(\alpha_2 + \frac{k}{N} (1 - \alpha_1 - \alpha_2) \right) \right], \end{aligned}$$

y por (2.18)

$$\begin{aligned} \prod_{k=0}^{j-1} \mu_{k+1} &= \prod_{k=0}^{j-1} \lambda \frac{k+1}{N} \left[1 - \alpha_2 - \frac{k+1}{N} (1 - \alpha_1 - \alpha_2) \right] \\ &= \left(\frac{\lambda}{N} \right)^j \left[\prod_{k=0}^{j-1} (k+1) \right] \prod_{k=0}^{j-1} \left[1 - \alpha_2 - \frac{(k+1)}{N} (1 - \alpha_1 - \alpha_2) \right], \end{aligned}$$

y por lo tanto

$$\begin{aligned} \prod_{k=0}^{j-1} \frac{\lambda_k}{\mu_{k+1}} &= \prod_{k=0}^{j-1} \frac{N-k}{k+1} \prod_{k=0}^{j-1} \left[\frac{\alpha_2 + \frac{k}{N} (1 - \alpha_1 - \alpha_2)}{1 - \alpha_2 - \frac{k+1}{N} (1 - \alpha_1 - \alpha_2)} \right] \\ &= \binom{N}{j} \prod_{k=0}^{j-1} \left[\frac{\alpha_2 + \frac{k}{N} (1 - \alpha_1 - \alpha_2)}{1 - \alpha_2 - \frac{k+1}{N} (1 - \alpha_1 - \alpha_2)} \right]. \end{aligned}$$

Tomando

$$v_j = \alpha_2 + \frac{j}{N} (1 - \alpha_1 - \alpha_2) \quad y \quad w_j = 1 - \alpha_2 - \frac{j}{N} (1 - \alpha_1 - \alpha_2) \quad (2.22)$$

se tiene que:

$$\prod_{k=0}^{j-1} \frac{\lambda_k}{\mu_{k+1}} = \binom{N}{j} \frac{v_0 v_1 \dots v_{j-1}}{w_1 \dots w_j}$$

y por tanto

$$\begin{aligned}
 P(0) &= \left(1 + \sum_{k=1}^N \binom{N}{k} \frac{v_0 v_1 \dots v_{k-1}}{w_1 \dots w_k} \right)^{-1} \\
 &= \left(\frac{1}{w_1 \dots w_N} \left(w_1 \dots w_N + \sum_{k=1}^N \binom{N}{k} \frac{v_0 v_1 \dots v_{k-1}}{w_1 \dots w_k} w_1 \dots w_N \right) \right)^{-1} \\
 &= w_1 \dots w_N \left(w_1 \dots w_N + \sum_{k=1}^N \binom{N}{k} v_0 v_1 \dots v_{k-1} w_{k+1} \dots w_N \right)^{-1} \\
 &= w_1 \dots w_N \left(w_1 \dots w_N + \binom{N}{1} v_0 w_2 \dots w_n + \dots + v_0 \dots v_{N-1} \right)^{-1}. \quad (2.23)
 \end{aligned}$$

Note que el segundo factor de (2.23) es una suma de la forma $S_N(\alpha_2, \frac{1-\alpha_1-\alpha_2}{N})$ (ver Lema B.1 Apéndice B), entonces

$$\begin{aligned}
 S_N \left(\alpha_2, \frac{1-\alpha_1-\alpha_2}{N} \right) &= w_1 \dots w_N + \binom{N}{1} v_0 w_2 \dots w_n + \dots + v_0 \dots v_{N-1} \\
 &= \prod_{k=1}^N \left(1 - \frac{k}{N} (1 - \alpha_1 - \alpha_2) \right).
 \end{aligned}$$

Por otro lado por (2.22), se tiene

$$w_1 \dots w_N = \prod_{k=1}^N \left(1 - \alpha_2 - \frac{k}{N} (1 - \alpha_1 - \alpha_2) \right),$$

entonces:

$$\begin{aligned}
 P(0) &= \frac{\prod_{k=1}^N (1 - \alpha_2 - \frac{k}{N}(1 - \alpha_1 - \alpha_2))}{\prod_{k=1}^N (1 - \frac{k}{N}(1 - \alpha_1 - \alpha_2))} \\
 &= \frac{\prod_{k=1}^N \left(\frac{N(1-\alpha_2)}{N} - \frac{k}{N}(1 - \alpha_1 - \alpha_2) \right)}{\prod_{k=1}^N \left(\frac{N}{N} - \frac{k}{N}(1 - \alpha_1 - \alpha_2) \right)} \\
 &= \frac{\frac{1-\alpha_1-\alpha_2}{N} \prod_{k=1}^N \left(\frac{N(1-\alpha_2)}{1-\alpha_1-\alpha_2} - k \right)}{\frac{1-\alpha_1-\alpha_2}{N} \prod_{k=1}^N \left(\frac{N}{1-\alpha_1-\alpha_2} - k \right)} = \frac{\prod_{k=1}^N \left(\frac{N(1-\alpha_2)}{1-\alpha_1-\alpha_2} - k \right)}{\prod_{k=1}^N \left(\frac{N}{1-\alpha_1-\alpha_2} - k \right)} \\
 &= \frac{\Gamma \left(\frac{N(\alpha_1+\alpha_2)}{1-\alpha_1-\alpha_2} \right) \prod_{k=1}^N \left(\frac{N(1-\alpha_2)}{1-\alpha_1-\alpha_2} - k \right) \Gamma \left(\frac{N\alpha_1}{1-\alpha_1-\alpha_2} \right)}{\Gamma \left(\frac{N(\alpha_1+\alpha_2)}{1-\alpha_1-\alpha_2} \right) \prod_{k=1}^N \left(\frac{N}{1-\alpha_1-\alpha_2} - k \right) \Gamma \left(\frac{N\alpha_1}{1-\alpha_1-\alpha_2} \right)}.
 \end{aligned}$$

Se tiene por el Lema B.3 del Apéndice B que

$$\Gamma \left(\frac{N(1-\alpha_2)}{1-\alpha_1-\alpha_2} \right) = \prod_{k=1}^N \left(\frac{N(1-\alpha_2)}{1-\alpha_1-\alpha_2} - k \right) \Gamma \left(\frac{N\alpha_1}{1-\alpha_1-\alpha_2} \right)$$

y

$$\Gamma \left(\frac{N}{1-\alpha_1-\alpha_2} \right) = \prod_{k=1}^N \left(\frac{N}{1-\alpha_1-\alpha_2} - k \right) \Gamma \left(\frac{N(\alpha_1+\alpha_2)}{1-\alpha_1-\alpha_2} \right).$$

Por lo tanto

$$P(0) = \frac{\Gamma \left(\frac{N(\alpha_1+\alpha_2)}{1-\alpha_1-\alpha_2} \right) \Gamma \left(\frac{N(1-\alpha_2)}{1-\alpha_1-\alpha_2} \right)}{\Gamma \left(\frac{N}{1-\alpha_1-\alpha_2} \right) \Gamma \left(\frac{N\alpha_1}{1-\alpha_1-\alpha_2} \right)}.$$

Ahora se analizará $P(j)$, $j = 1, \dots, N$.

$$\begin{aligned}
 P(j) &= \frac{\lambda_0 \dots \lambda_{j-1}}{\mu_1 \dots \mu_j} P(0) \\
 &= \frac{N! v_0 v_1 \dots v_{j-1}}{j!(N-j)! w_1 \dots w_j} P(0) \\
 &= \frac{\Gamma(N+1) v_0 v_1 \dots v_{j-1}}{\Gamma(j+1) \Gamma(N-j+1) w_1 \dots w_j} P(0).
 \end{aligned}$$

Como

$$\begin{aligned} w_1 \dots w_j &= \prod_{k=1}^j \left(1 - \alpha_2 - \frac{k}{N} (1 - \alpha_1 - \alpha_2) \right) \\ &= \frac{1 - \alpha_1 - \alpha_2}{N} \prod_{k=1}^j \left(\frac{N(1 - \alpha_2)}{1 - \alpha_1 - \alpha_2} - k \right) \\ &= \frac{1 - \alpha_1 - \alpha_2}{N} \frac{\prod_{k=1}^j \left(\frac{N(1 - \alpha_2)}{1 - \alpha_1 - \alpha_2} - k \right) \Gamma \left(\frac{N(1 - \alpha_2)}{1 - \alpha_1 - \alpha_2} - j \right)}{\Gamma \left(\frac{N(1 - \alpha_2)}{1 - \alpha_1 - \alpha_2} - j \right)}. \end{aligned}$$

Por el Lema B.3 del Apéndice B, se tiene que

$$\Gamma \left(\frac{N(1 - \alpha_2)}{1 - \alpha_1 - \alpha_2} \right) = \prod_{k=1}^j \left(\frac{N(1 - \alpha_2)}{1 - \alpha_1 - \alpha_2} - k \right) \Gamma \left(\frac{N(1 - \alpha_2)}{1 - \alpha_1 - \alpha_2} - j \right),$$

entonces

$$w_1 \dots w_j = \frac{1 - \alpha_1 - \alpha_2}{N} \frac{\Gamma \left(\frac{N(1 - \alpha_2)}{1 - \alpha_1 - \alpha_2} \right)}{\Gamma \left(\frac{N(1 - \alpha_2)}{1 - \alpha_1 - \alpha_2} - j \right)}.$$

Por otro lado

$$\begin{aligned} v_0 v_1 \dots v_{j-1} &= \frac{1 - \alpha_1 - \alpha_2}{N} \left(\frac{N\alpha_2}{1 - \alpha_1 - \alpha_2} \right) \left(\frac{N\alpha_2}{1 - \alpha_1 - \alpha_2} + 1 \right) \dots \left(\frac{N\alpha_2}{1 - \alpha_1 - \alpha_2} + j - 1 \right) \\ &= \frac{1 - \alpha_1 - \alpha_2}{N} \prod_{k=1}^j \left(\frac{N\alpha_2}{1 - \alpha_1 - \alpha_2} + j - k \right) \\ &= \frac{1 - \alpha_1 - \alpha_2}{N} \frac{\prod_{k=1}^j \left(\frac{N\alpha_2}{1 - \alpha_1 - \alpha_2} + j - k \right) \Gamma \left(\frac{N\alpha_2}{1 - \alpha_1 - \alpha_2} \right)}{\Gamma \left(\frac{N\alpha_2}{1 - \alpha_1 - \alpha_2} \right)}. \end{aligned}$$

También por el Lema B.3 del Apéndice B, se cumple

$$\Gamma \left(\frac{N\alpha_2}{1 - \alpha_1 - \alpha_2} + j \right) = \left[\prod_{k=1}^j \left(\frac{N\alpha_2}{1 - \alpha_1 - \alpha_2} + j - k \right) \right] \Gamma \left(\frac{N\alpha_2}{1 - \alpha_1 - \alpha_2} \right).$$

Entonces

$$\begin{aligned} v_0 v_1 \dots v_{j-1} &= \frac{\frac{1-a_1-a_2}{N} \prod_{k=1}^j \left(\frac{Na_2}{1-a_1-a_2} + j - k \right) \Gamma \left(\frac{Na_2}{1-a_1-a_2} \right)}{\Gamma \left(\frac{Na_2}{1-a_1-a_2} \right)} \\ &= \frac{\frac{1-a_1-a_2}{N} \Gamma \left(\frac{Na_2}{1-a_1-a_2} + j \right)}{\Gamma \left(\frac{Na_2}{1-a_1-a_2} \right)}. \end{aligned}$$

Se sigue que

$$\begin{aligned} \frac{v_0 \dots v_{j-1}}{w_1 \dots w_n} &= \frac{\frac{1-a_1-a_2}{N} \Gamma \left(\frac{Na_2}{1-a_1-a_2} + j \right) / \Gamma \left(\frac{Na_2}{1-a_1-a_2} \right)}{\frac{1-a_1-a_2}{N} \Gamma \left(\frac{N(1-a_2)}{1-a_1-a_2} \right) / \Gamma \left(\frac{N(1-a_2)}{1-a_1-a_2} - j \right)} \\ &= \frac{\Gamma \left(\frac{Na_2}{1-a_1-a_2} + j \right) / \Gamma \left(\frac{Na_2}{1-a_1-a_2} \right)}{\Gamma \left(\frac{N(1-a_2)}{1-a_1-a_2} \right) / \Gamma \left(\frac{N(1-a_2)}{1-a_1-a_2} - j \right)}. \end{aligned}$$

Por lo tanto

$$\begin{aligned} P(j) &= \frac{\lambda_0 \dots \lambda_{j-1}}{\mu_1 \dots \mu_j} P(0) \\ &= \frac{\Gamma(N+1)}{\Gamma(j+1)\Gamma(N-j+1)} \frac{\Gamma \left(\frac{Na_2}{1-a_1-a_2} + j \right) / \Gamma \left(\frac{Na_2}{1-a_1-a_2} \right)}{\Gamma \left(\frac{N(1-a_2)}{1-a_1-a_2} \right) / \Gamma \left(\frac{N(1-a_2)}{1-a_1-a_2} - j \right)} P(0) \\ &= \frac{\Gamma(N+1)}{\Gamma(j+1)\Gamma(N-j+1)} \\ &\quad \times \frac{\Gamma \left(\frac{Na_2}{1-a_1-a_2} + j \right) \Gamma \left(\frac{N(1-a_2)}{1-a_1-a_2} - j \right) \Gamma \left(\frac{N(a_1+a_2)}{1-a_1-a_2} \right) \Gamma \left(\frac{N(1-a_2)}{1-a_1-a_2} \right)}{\Gamma \left(\frac{Na_2}{1-a_1-a_2} \right) \Gamma \left(\frac{N(1-a_2)}{1-a_1-a_2} \right) \Gamma \left(\frac{N}{1-a_1-a_2} \right) \Gamma \left(\frac{Na_1}{1-a_1-a_2} \right)} \\ &= \frac{\Gamma(N+1)}{\Gamma(j+1)\Gamma(N-j+1)} \frac{\Gamma \left(\frac{Na_2}{1-a_1-a_2} + j \right) \Gamma \left(\frac{N(1-a_2)}{1-a_1-a_2} - j \right) \Gamma \left(\frac{N(a_1+a_2)}{1-a_1-a_2} \right)}{\Gamma \left(\frac{Na_2}{1-a_1-a_2} \right) \Gamma \left(\frac{N}{1-a_1-a_2} \right) \Gamma \left(\frac{Na_1}{1-a_1-a_2} \right)}. \end{aligned}$$

□

Capítulo 3

El proceso n -coalescente

En el capítulo anterior se analizó un estudio de la evolución de poblaciones de acuerdo con el modelo de Moran y el modelo de Wright-Fisher. En este Capítulo se considerará otro proceso relevante para el estudio de la ecología evolutiva, se trata del estudio de la genealogía de poblaciones. En este tema se considerará un modelo matemático para su análisis propuesto por Kingman en 1980, llamado el n -coalescente (ver Kingman (1982a, 1982b)).

Como se hizo en el Capítulo 2, primeramente se describirá las hipótesis del modelo. Se supondrá que la reproducción de la población para la cual se utilizará el n -coalescente sigue el modelo de Moran ó el de Wright-Fisher. Por lo tanto el tipo de individuos considerados será igual que en el capítulo anterior, es decir, individuos haploides con un único locus, también se asumirá que son neutrales. Dado que el interés se concentrará en el estudio de las relaciones de parentesco entre los individuos, en este análisis no se considerará mutación. En cuanto al tamaño de la población se considerará constante, aunque también se puede analizar con tamaño variable que conlleva a algunos supuestos adicionales y el uso de cadenas de Markov no homogéneas en el tiempo.

El coalescente es una cadena de Markov que describe las relaciones familiares entre lo miembros de una población, así que el tipo de alelo que se tenga es irrelevante.

Si se tiene un cambio adecuado en la escala de tiempo y espacio en los modelos de Moran y Wright-Fisher, se observaría que ellos se comportan de forma idéntica (ver Karlin y Taylor (1981), Rodrigues (2002), Sánchez (2002)), así que se describirá el n -coalescente de Kingman solamente para

el modelo de Moran (por su sencillez) y también en la inferencia para las poblaciones en genética en el Capítulo 5.

3.1 El proceso del n -coalescente de Kingman y el modelo de Moran

Las condiciones necesarias para este modelo son las siguientes: En un tiempo t_0 se toma una muestra de tamaño n de una población haploide de tamaño N ($n < N$), y se localizan sus ancestros inmediatos. En el caso de una población evolucionando de acuerdo con el modelo de Wright-Fisher por ejemplo, estos ancestros se encuentran en una generación anterior. Para estos ancestros también se localizan sus más recientes ancestros y se repite este proceso sucesivamente hasta que el total de ancestros conste de un sólo individuo. El interés está en encontrar en el más reciente común antecesor de una muestra de tamaño n y el tiempo que tarda hasta llegar a este ancestro, entre otras cosas.

Como el tiempo en que se localizan los ancestros es anterior al tiempo donde se toma la muestra (t_0), entonces el tiempo en el proceso que cuenta el número de los ancestros de la muestra contará del presente al pasado con respecto al tiempo que describe la evolución de la población.

La descripción de la composición de los ancestros en un tiempo cualquiera en el pasado será a través de una relación de equivalencia en $\{1, 2, \dots, n\}$, descrita en la siguiente definición.

Definición 3.1.1. *Se dice que el individuo i se encuentra en la misma clase de equivalencia que el individuo j al tiempo s , siempre que los miembros i y j tengan en común el antecesor al tiempo $t_0 - s$. Se indicará que los individuos i y j están en la misma clase de equivalencia por (i, j) . En general si k -individuos tienen el mismo ancestro en un tiempo s , se denotará análogamente por (i_1, i_2, \dots, i_k) . Al conjunto finito que contiene todas las clases de equivalencia posibles en $\{1, 2, \dots, n\}$ se denotará por \mathcal{E}_n .*

Entonces para cada tiempo $t_0 - s$, se tiene un conjunto de clases de equivalencia de la muestra de tamaño n , que será denotado por R_s . Para $R_s \in \mathcal{E}_n$, se indicará por $|R_s|$ al número de clases de equivalencia de R_s .

Sean $\xi, \eta \in \mathcal{E}_n$, se introduce la notación $\xi \prec \eta$, que indica que η es obtenido de ξ , por la inmersión de dos de sus clases de equivalencia, es decir,

$$\xi \prec \eta \text{ si y sólo si } \xi \subset \eta \text{ y } |\xi| = |\eta| + 1. \quad (3.1)$$

Kingman propone el modelo que indica la probabilidad de que en un tiempo la muestra tenga determinados ancestros y lo llamo n -coalescente (ver Kingman (1982a), (1982b)).

Definición 3.1.2. Se le llama proceso n -coalescente a la cadena de Markov en tiempo continuo $R = \{R_t : t \geq 0\}$ con espacio de estados \mathcal{E}_n , que indica los predecesores al tiempo $t_0 - t$ de una población haploide de tamaño n al tiempo t_0 , donde

(a) R_0 es la relación identidad, es decir,

$$R_0 = \Delta = \{(1), (2), \dots, (n)\}. \quad (3.2)$$

indicando que al tiempo t_0 cada individuo es el ancestro de él mismo.

(b) Para $\xi, \eta \in \mathcal{E}_n$, R tiene tasa de transición $q_{\xi\eta}$, $\xi \neq \eta$ dada por:

$$q_{\xi\eta} = \begin{cases} 1, & \text{si } \xi \prec \eta \\ 0, & \text{en otro caso.} \end{cases} \quad (3.3)$$

Ejemplo 3.1.1. En la siguiente figura se puede mostrar el proceso n -coalescente para una muestra de 6 individuos.

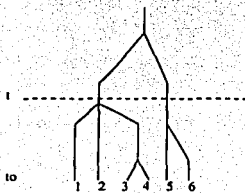


Figura 3.1: Note que al tiempo $t_0 - t$ se tienen dos ancestros.

Teorema 3.1.1. Sea $R = \{R_t : t \geq 0\}$ el n -coalescente con espacio de estados \mathcal{E}_n , donde la dinámica de la población se comporta como el modelo de Moran, entonces:

- (a) R permanece en el estado ξ , $\xi \in \mathcal{E}_n$, por un tiempo exponencialmente distribuido con parámetro

$$\nu_\xi = \binom{|\xi|}{2} = \frac{|\xi|(|\xi| - 1)}{2}. \quad (3.4)$$

- (b) R tiene probabilidades de transición infinitesimales de ξ a η , $\xi, \eta \in \mathcal{E}_n$, determinadas de la siguiente manera:

$$\begin{aligned} P(R_{t+h} = \eta | R_t = \xi) &= h + o(h), & \text{si } \xi \prec \eta, \\ P(R_{t+h} = \eta | R_t = \xi) &= 1 - \left(\frac{|\xi|(|\xi| - 1)}{2} \right) h + o(h), & \text{si } \xi = \eta, \end{aligned} \quad (3.5)$$

$$P(R_{t+h} = \eta | R_t = \xi) = o(h), \quad \text{si } \xi \neq \eta \text{ y } \xi \not\prec \eta.$$

- (c) El único estado absorbente de R es el que tiene sólo una clase de equivalencia, que es:

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

Demostración. (a) Como R es una cadena de Markov en tiempo continuo por la Definición 1.3.5, el tiempo de permanencia en un estado ξ , $\xi \in \mathcal{E}_n$, es exponencial con parámetro ν_ξ y por (1.17) se tiene que:

$$\nu_\xi = \sum_{\eta \neq \xi} q_{\xi\eta},$$

como $q_{\xi\eta}$ es diferente de cero (es igual a uno) sólo en el caso que $\xi \prec \eta$, lo cual sucede si existe una coalescencia de dos de las clases de ξ , entonces sólo se debe contar todas las posibles coalescencias. Por lo tanto:

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

(b) Por la Definición 1.3.7, las probabilidades de transición infinitesimales de R de ξ a η , $\xi, \eta \in \mathcal{E}_n$, están dadas por

$$P(R_{\tau+h} = \eta | R_\tau = \xi) = q_{\xi\eta}h + o(h), \quad \eta \neq \xi,$$

y

$$P(R_{\tau+h} = \eta | R_{\tau} = \xi) = 1 - \nu_{\xi} h + o(h), \quad \eta = \xi.$$

Por (3.3) se tiene que cuando $\xi \neq \eta$ y $\xi \neq \eta$.

$$P(R_{t+h} = \eta | R_t = \xi) = o(h).$$

En el caso que $\xi \prec \eta$, se tiene que

$$P(R_{t+h} = \eta | R_t = \xi) = h + o(h).$$

Por otro lado por (a) de este teorema, se tiene que $\nu_{\xi} = \frac{k\xi(|\xi|-1)}{2}$. Entonces

$$P(R_{t+h} = \eta | R_t = \xi) = 1 - \left(\frac{|\xi|(|\xi|-1)}{2} \right) h + o(h), \quad \text{si } \xi = \eta.$$

(c) La cardinalidad de Θ es igual a uno y es el único estado con esta característica. Entonces por (a) se tiene que $\nu_{\Theta} = 0$, se sigue de la nota de la Definición 1.3.5 que Θ es un estado absorbente. \square

Con lo que respecta a la cadena de saltos de R (por la Definición 1.3.4), se tiene que está dada por la variable aleatoria que indica los estados visitados por R . Dado que R indica los ancestros de la muestra y que el comportamiento de la población es como en el modelo de Moran, se tiene que si ocurre una transición entonces hay inmersión de dos de las clases de R . De esta forma la sucesión de estados visitados por la cadena de saltos indicada por $\mathcal{R} = \{\mathcal{R}_k : k = n, n-1, \dots, 1\}$ será:

$$\mathcal{R}_n, \mathcal{R}_{n-1}, \dots, \mathcal{R}_1.$$

Observaciones:

- $\mathcal{R}_n \prec \mathcal{R}_{n-1} \prec \dots \prec \mathcal{R}_1$,
- $|\mathcal{R}_k| = k$,
- R permanece en \mathcal{R}_k un período de tiempo T_k , exponencialmente distribuido con parámetro $|\mathcal{R}_k|$.

Recuerde que para cadenas de Markov en tiempo continuo, la cadena de saltos queda determinada por las probabilidades de transición P_{ij} , $i \neq j$. Kingman precisa estas probabilidades al relacionar la cadena de saltos de R y una cadena de muerte pura (ver Kingman (1982a), (1982b)), la cual se describirá en el siguiente teorema.

Teorema 3.1.2. Sea $D = \{D_t : t \geq 0\}$ una cadena de Markov continua con espacio de estados $S = \{1, 2, \dots, n\}$, donde D_t indica la cardinalidad de clases de equivalencia, al tiempo t , es decir, $D_t = |R_t|$ entonces:

(a) D es un proceso de muerte pura.

(b) Si $D_t = k$, entonces el tiempo de permanencia en el estado k tiene distribución exponencial con parámetro

$$d_k = \binom{|k|}{2} = \frac{k(k-1)}{2}. \quad (3.6)$$

(c) D tiene probabilidades de transición infinitesimales dadas por:

$$P(D_{t+h} = j | D_t = i) = \begin{cases} d_i h + o(h), & j = i - 1 \\ 1 - d_i h + o(h), & j = i \\ 0, & j \neq i, i - 1. \end{cases} \quad (3.7)$$

Observación: Como $D_t = |R_t|$ entonces ocurre una transición en D si y sólo si una transición también ocurre en R . Esto se debe a que si $R_t = \xi$ y $R_{t+h} = \eta$, con $\eta \prec \xi$, se tiene que $|\eta| = |\xi| - 1$, por lo cual si $D_t = k$ y una coalescencia ocurre en $t + h$, entonces $D_{t+h} = k - 1$.

Demostración. (a) Si la cadena D está en el estado j en el tiempo t , por la observación anterior se tiene que lo único que puede suceder en $t + h$ (h pequeño) es que D permanezca en j o realice una transición a $j - 1$. Por lo tanto D es un proceso de muerte pura.

(b) También por la misma observación se tiene que el tiempo de permanencia para D en el estado k (si $|R_t| = k$) es el mismo que el del proceso R en $R_t = \xi$, $|\xi| = k$, por lo tanto:

$$d_k = \nu_\xi = \binom{|\xi|}{2} = \frac{k(k-1)}{2}.$$

(c) Se cumple por (b) de este teorema y por la Definición (1.3.7). \square

Uno de los aspectos importantes en la genealogía de poblaciones es el momento en que se llega al común antecesor de la población, el cual se puede analizar a través del proceso D , ya que este momento es cuando D llega al estado 1 que es su único estado absorbente.

Como D es una cadena de Markov continua entonces el tiempo de permanencia en el estado k que se denotará por T_k , es exponencial con parámetro d_k (dado por (3.6)). Entonces:

$$P(T_k > t) = d_k e^{-d_k t}.$$

Note que T_k , $k = 2, 3, \dots, n$, son mutuamente independientes.

Si se designa por T al tiempo en que D llega a la absorción, entonces:

$$T = \inf\{t \geq 0 : R_t = \Theta\} = \inf\{t \geq 0 : D_t = 1\},$$

y

$$T = \sum_{k=2}^n T_k,$$

entonces

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

Por lo tanto

$$E(T) = 2\left(1 - \frac{1}{n}\right). \quad (3.8)$$

Se tiene entonces que (3.8) indica que para tamaño de muestra muy grande, se espera que en dos generaciones el ancestro en común sea alcanzado. Dado que para el modelo de Moran cada generación equivale a N eventos de

nacimiento y muerte, se tiene que se necesitan $2N$ eventos de nacimiento y muerte para llegar al ancestro en común de la población.

- (1) Se cumple por propiedades de la esperanza.
- (2) Se tiene que la esperanza de una exponencial es uno entre su parámetro por lo tanto $E[T_k] = 1/d_k$.

Teorema 3.1.3. Sea $R = \{R_t : t \geq 0\}$ el n -coalescente con espacio de estados \mathcal{E}_n , y $D = \{D_t : t \geq 0\}$ el proceso de muerte con espacio de estados $S = \{1, 2, \dots, n\}$, donde $D_t = |R_t|$ y sea $\mathcal{R} = \{\mathcal{R}_k : k = n, n-1, \dots, 1\}$ la cadena de saltos de R , entonces:

- (a) El proceso de muerte pura D y la cadena de saltos son independientes y $R_t = \mathcal{R}_{D_t}$, $t \geq 0$.
- (b) Las probabilidades de transición del estado ξ al estado η ($\xi, \eta \in \mathcal{E}_n$) para la cadena de saltos \mathcal{R} son:

$$P(\mathcal{R}_{k-1} = \eta | \mathcal{R}_k = \xi) = \begin{cases} \frac{2}{k(k-1)}, & \text{si } \xi \prec \eta \\ 0, & \text{en otro caso,} \end{cases} \quad (3.9)$$

donde $|\xi| = k$, $2 \leq k \leq n$.

- (c) Las probabilidades de $\mathcal{R}_k = \xi$ quedan determinadas de la siguiente forma:

$$P(\mathcal{R}_k = \xi) = \frac{(n-k)!k!(k-1)!}{n!(n-1)!} n_1!n_2!\dots n_k!, \quad (3.10)$$

donde n_1, n_2, \dots, n_k son los tamaños de las clases de equivalencia de ξ .

Demostración. (a) Se tiene que D_t indica el número de clases de equivalencia de R_t . Note que si $R_t = \xi$ y $D_t = k$. Adicionalmente se tiene que $\nu_\xi = \nu_k = d_k$ y los T_k son independientes con distribución exponencial con parámetro d_k , $k = n, n-1, \dots, 2$, entonces para cada $t \geq 0$, la distribución de D_t dado \mathcal{R} es la misma que su distribución no condicionada a \mathcal{R} . Por lo tanto D y \mathcal{R} son independientes.

Por definición $D_t = |R_t|$ y si $|R_t| = k$ entonces también por definición $R_t = \mathcal{R}_k$, por lo tanto para todo $t \geq 0$ se cumple

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

(b) Las probabilidades de transición de la cadena de saltos \mathcal{R} pueden ser obtenidas por (1.16). Note que por la definición de la tasa de transición de una cadena de Markov continua, se tiene que para $\xi, \eta \in \mathcal{E}_n$, vale

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

Entonces si $|\xi| = k$, se tiene que

$$q_{\xi\eta} = 1$$

y

$$\nu_\xi = \sum_{\xi \prec \eta} q_{\xi\eta} = \frac{k(k-1)}{2} = \binom{k}{2},$$

dado que existen $\binom{k}{2}$ formas de dos clases que coalescen en una colección de k clase, se tiene que el resultado se cumple.

(c) La demostración se realizará por inducción. Para $k = n$, se tiene que

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

donde Δ está dada por (3.2). Entonces se cumple (3.10).

Suponga que el resultado vale para $k = \ell$, es decir,

$$P(\mathcal{R}_\ell = \xi) = \frac{(n-\ell)! \ell! (\ell-1)!}{n!(n-1)!} n_1! n_2! \dots n_\ell!$$

Se demostrará que se vale para $k = \ell - 1$,

$$\begin{aligned} P(\mathcal{R}_{\ell-1} = \eta) &\stackrel{(1)}{=} \sum_{\xi \in \mathcal{E}_n} P(\mathcal{R}_{\ell-1} = \eta | \mathcal{R}_\ell = \xi) P(\mathcal{R}_\ell = \xi) \\ &\stackrel{(2)}{=} \sum_{\xi \prec \eta} \frac{2}{\ell(\ell-1)} P(\mathcal{R}_\ell = \xi). \end{aligned}$$

(1) Se tiene que para ℓ fijo, las clases que puede tomar \mathcal{R} son mutuamente excluyentes así que se puede aplicar el teorema de probabilidad total (ver Teorema A.1 del Apéndice A).

(2) Es válido por (3.9).

Para obtener la probabilidad de $\mathcal{R}_\ell = \xi$ se debe tomar en cuenta que si $\mathcal{R}_{\ell-1} = \eta$, entonces ξ difiere de η necesariamente en las clases de equivalencia que coalescen, ya que la cadena \mathcal{R} pasó de un estado con ℓ clases de equivalencia a uno con $\ell - 1$ clases. Sean $\eta_1, \eta_2, \dots, \eta_{k-1}$ las clases de equivalencia de η y n_1, n_2, \dots, n_{k-1} los respectivos tamaños de las clases. Entonces existe m , $1 \leq m \leq k - 1$, tal que η_m es formada por la coalescencia de dos clases de ξ , por lo tanto n_m es la suma del tamaño de estas dos clases. Sea M tal que $1 \leq M \leq n_m - 1$ y sean $n_1, n_2, \dots, n_{m-1}, M, n_m - M, n_{m+1}, \dots, n_{k-1}$ los tamaños de las clases de ξ . Entonces por hipótesis de inducción, se tiene que:

$$P(\mathcal{R}_\ell = \xi) = \frac{(n - \ell)! \ell! (\ell - 1)!}{n! (n - 1)!} n_1! n_2! \dots n_{m-1}! M! (n_m - M)! n_{m+1}! \dots n_{\ell-1}!$$

$$P(\mathcal{R}_{\ell-1} = \eta)$$

$$\begin{aligned} & \stackrel{(3)}{=} \sum_{m=1}^{\ell-1} \sum_{M=1}^{n_m-1} \frac{2}{\ell(\ell-1)} \frac{(n-\ell)! \ell! (\ell-1)!}{n! (n-1)!} \\ & \quad \times n_1! n_2! \dots n_{m-1}! M! (n_m - M)! n_{m+1}! \dots n_{\ell-1}! \frac{1}{2} \binom{n_m}{M} \\ & = \sum_{m=1}^{\ell-1} \sum_{M=1}^{n_m-1} \frac{(n-\ell)! (\ell-1)! (\ell-2)!}{n! (n-1)!} \\ & \quad \times n_1! n_2! \dots n_{m-1}! M! (n_m - M)! n_{m+1}! \dots n_{\ell-1}! \frac{n_m!}{M! (n_m - M)!} \\ & = \sum_{m=1}^{\ell-1} \sum_{M=1}^{n_m-1} \frac{(n-\ell)! (\ell-1)! (\ell-2)!}{n! (n-1)!} n_1! n_2! \dots n_{m-1}! n_m! n_{m+1}! \dots n_{\ell-1}! \\ & = \frac{(n-\ell)! (\ell-1)! (\ell-2)!}{n! (n-1)!} n_1! n_2! \dots n_{m-1}! n_m! n_{m+1}! \dots n_{\ell-1}! \sum_{m=1}^{\ell-1} \sum_{M=1}^{n_m-1} 1 \\ & \stackrel{(4)}{=} \frac{(n-\ell)! (\ell-1)! (\ell-2)!}{n! (n-1)!} n_1! n_2! \dots n_{m-1}! n_m! n_{m+1}! \dots n_{\ell-1}! (n - (\ell - 1)) \\ & = \frac{(n - (\ell - 1))! (\ell - 1)! (\ell - 2)!}{n! (n - 1)!} n_1! n_2! \dots n_{\ell-1}! \end{aligned}$$

(3) Es válido por lo siguiente:

- $|\eta| = |\xi| - 1 = \ell - 1$ y $1 \leq M \leq n_M - 1$.
- Por hipótesis de inducción.
- Por que $\frac{1}{2} \binom{n_M}{M}$, es el número de posibles formas sin importar el orden que se puede formar un conjunto de tamaño M y otro de tamaño $n_M - M$, a partir de un conjunto de tamaño n_m .

(4) Es válido ya que

$$\sum_{m=1}^{\ell-1} \sum_{M=1}^{n_m-1} 1 = \sum_{m=1}^{\ell-1} (n_m - 1) = \sum_{m=1}^{\ell-1} n_m - \sum_{m=1}^{\ell-1} 1 = n - (\ell - 1).$$

□

Proposición 3.1.4. Sea $R = \{R_t : t \geq 0\}$ el n -coalescente con espacio de estados \mathcal{E}_n y $\mathcal{A} = \{\mathcal{A}_k : k = n, n-1, \dots, 1\}$ la cadena de saltos de R , entonces:

1. Para $\xi, \eta \in \mathcal{E}_n$, $\xi \subset \eta$, tal que $|\xi| = k$, $|\eta| = \ell$, se cumple:

$$P(\mathcal{A}_\ell = \eta | \mathcal{A}_k = \xi) = \frac{(k - \ell)! \ell! (\ell - 1)!}{k! (k - 1)!} n_1! n_2! \dots n_\ell!, \quad (3.13)$$

donde n_1, n_2, \dots, n_ℓ son los tamaños de las clases de equivalencia en η en \mathcal{E}_n inducida por las clases de equivalencia de $\xi \in \mathcal{E}_n$.

2. Para $\xi \in \mathcal{E}_n$, donde $|\xi| = k$, se tiene que:

$$P(R_t = \xi) = P(\mathcal{A}_k = \xi) P(D_t = k).$$

3.

$$P(D_t = k) = P\left(\sum_{s=k+1}^n T_s \leq t\right) - P\left(\sum_{s=k+1}^n T_s \leq t\right).$$

Demostración. 1. Se demostrará por inducción en ℓ . Para $\ell = k - 1$ se cumple por el Teorema 3.1.3. Se supone que el resultado vale para $\ell = m$ y se demuestra que vale para $\ell = m - 1$ de manera análoga a la hecha en la demostración del Teorema 3.1.3.

2. También por el Teorema 3.1.3, se tiene lo siguiente:

$$\begin{aligned}
 P(R_t = \xi) &= P(\mathcal{A}_{D_t} = \xi) \stackrel{(1)}{=} \sum_{\ell \in S} P(\mathcal{A}_t = \xi, D_t = \ell) \\
 &\stackrel{(2)}{=} \sum_{\ell \in S} P(\mathcal{A}_t = \xi | D_t = \ell) P(D_t = \ell) \\
 &\stackrel{(3)}{=} P(\mathcal{A}_k = \xi | D_t = k) P(D_t = k) \\
 &= P(\mathcal{A}_k = \xi) P(D_t = k).
 \end{aligned}$$

(1) Se vale por propiedades de probabilidad conjunta.

(2) Válido por la definición de probabilidad condicional.

(3) Se vale ya que $D_t = |R_t| = |\xi| = k$.

3. Recuerde que T_k es el tiempo de permanencia de D en el estado k . Entonces al tiempo t se tienen k clases si y sólo si el total del tiempo de permanencia de D en los estados $n, n-1, \dots, k$ es menor que t . Entonces se tiene que el evento $\{D_t = k\}$, es equivalente a:

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

que equivale al evento

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

De esta forma

$$\begin{aligned}
 & P(D_t = k) \\
 &= P \left\{ \sum_{s=k+1}^n T_s \leq t, \sum_{s=k}^n T_s \geq t \right\} \\
 &= P \left\{ \sum_{s=k+1}^n T_s \leq t, T_k + \sum_{s=k+1}^n T_s \geq t \right\} \\
 &\stackrel{(1)}{=} P \left\{ T_k + \sum_{s=k+1}^n T_s \geq t \mid \sum_{s=k+1}^n T_s \leq t, \right\} P \left\{ \sum_{s=k+1}^n T_s \leq t, \right\} \\
 &= \left[1 - P \left\{ T_k + \sum_{s=k+1}^n T_s \leq t \mid \sum_{s=k+1}^n T_s \leq t, \right\} \right] P \left\{ \sum_{s=k+1}^n T_s \leq t, \right\} \\
 &\stackrel{(1)}{=} P \left\{ \sum_{s=k+1}^n T_s \leq t \right\} - P \left\{ T_k + \sum_{s=k+1}^n T_s \leq t, \sum_{s=k+1}^n T_s \leq t, \right\} \\
 &= P \left(\sum_{s=k+1}^n T_s \leq t \right) - P \left(\sum_{s=k}^n T_s \leq t \right).
 \end{aligned}$$

(1) Por definición de probabilidad condicional.

□

Capítulo 4

La estadística bayesiana y el método de Monte Carlo

4.1 Introducción

En la naturaleza se encuentran fenómenos aleatorios para los cuales no se puede obtener de manera directa e inmediata información respecto de ellos. De esta forma se puede hacer uso de técnicas estadísticas como el análisis Bayesiano que propociona métodos para estimar los parámetros que determinan el comportamiento del fenómeno aleatorio que produce los resultados que se observan (datos).

Si el modelo que registra el fenómeno observado es especificado, entonces es posible hacerse predicciones al respecto de la ocurrencia de resultados de interés.

Sea θ el parámetro que regula el modelo considerado y sea y_1, \dots, y_n un conjunto de observaciones producido por el fenómeno aleatorio que se desea estudiar. Note que dado θ , y teniendo en cuenta la hipótesis del modelo considerado, es posible calcular la probabilidad de obtener los datos y_1, \dots, y_n , es decir, si se repite el fenómeno aleatorio n veces y si se indica por $\mathbf{Y} = (Y_1, \dots, Y_n)$ al vector aleatorio que registra los resultados obtenidos donde Y_i , $i = 1, 2, \dots, n$, es la variable aleatoria que indica el resultado de la i -ésima realización del experimento. Entonces dado θ es posible calcular para

$y = (y_1, \dots, y_n)$:

$$P(\mathbf{Y} = \mathbf{y}|\theta) = P((Y_1, Y_2, \dots, Y_n) = (y_1, y_2, \dots, y_n)|\theta) = p(y_1, y_2, \dots, y_n|\theta) \quad (4.1)$$

Definición 4.1.1. Sean Y_1, Y_2, \dots, Y_n variables aleatorias idénticamente distribuidas con función de probabilidad $p(\cdot|\theta)$, las cuales registran los resultados de n repeticiones de un experimento aleatorio, es decir, $Y_i, i = 1, 2, \dots, n$ es la variable aleatoria que indica el resultado de la i -ésima realización del experimento. Se denotará por $\mathbf{Y} = (Y_1, Y_2, \dots, Y_n)$ al vector aleatorio que registra el resultado de n repeticiones del experimento. A la función $p((y_1, y_2, \dots, y_n)|\theta)$, dada por (4.1) se le llama **función de verosimilitud** para el modelo con parámetro θ .

Observación: Note que si el experimento aleatorio es repetido n veces independientemente, entonces

$$p((y_1, y_2, \dots, y_n)|\theta) = \prod_{j=1}^n p(y_j|\theta),$$

donde

$$p(y_i|\theta) = P(Y_i = y_i|\theta), \quad i = 1, 2, \dots, n.$$

Por lo descrito hasta el momento se puede observar que si se tiene el parámetro θ y conocimiento del modelo, se puede describir cómo se distribuyen los resultados del experimento. Sin embargo, este conocimiento no siempre está disponible. De esta forma, se puede considerar θ como una variable aleatoria y suponer que se distribuye de acuerdo con alguna distribución (ver Carlin y Louis (1996)).

Definición 4.1.2. Sea θ el parámetro aleatorio que caracteriza un modelo bajo el cual se realiza un experimento y sea $y = (y_1, \dots, y_n)$ los resultados de n repeticiones de este experimento. A la distribución de θ antes de que se tome en cuenta los resultados del experimento es llamada **distribución a priori** de θ y es indicada por $p_{\text{priori}}(\theta)$. A la distribución de θ dado los resultados del experimento, es decir, $p(\theta|y)$ es llamada **distribución a posteriori** de θ y es indicada por $p_{\text{posteriori}}(\theta)$.

Observación: Note que la distribución $p_{\text{priori}}(\theta)$ depende de algún parámetro η , el cual puede ser determinista o aleatorio. En el caso que se quiera llamar la atención por este hecho se usará $p_{\text{priori}}(\theta|\eta)$.

La relación que existe entre $p_{\text{priori}}(\theta)$, $p_{\text{posteriori}}(\theta)$ y $p(y|\theta)$, está dada en el siguiente teorema.

Teorema 4.1.1. Sea $Y = (Y_1, Y_2, \dots, Y_n)$ el vector aleatorio como en la Definición 4.1.1, donde θ también es una variable aleatoria con espacio parametral Θ y función de probabilidad a priori $p_{\text{priori}}(\theta|\eta)$, donde η es un vector de hiperparámetros que caracterizan la distribución de θ . Entonces la distribución posteriori de θ queda determinada de la siguiente manera:

a) Si η es conocido:

$$p_{\text{posteriori}}(\theta) = p(\theta|y, \eta) = \frac{p(y|\theta)p_{\text{priori}}(\theta|\eta)}{\sum_{u \in \Theta} p(y|u)p_{\text{priori}}(u|\eta)}. \quad (4.2)$$

b) Si η es desconocido con espacio hiperparametral Λ , entonces:

$$p_{\text{posteriori}}(\theta) = p(\theta|y) = \frac{\sum_{\eta \in \Lambda} p(y|\theta)p_{\text{priori}}(\theta|\eta)h(\eta)}{\sum_{\eta \in \Lambda} \sum_{u \in \Theta} p(y|u)p_{\text{priori}}(u|\eta)h(\eta)}. \quad (4.3)$$

donde $h(\cdot)$ es la función de probabilidad de η (llamada hiperpriori).

Demostración. a) Como η es conocido se tiene lo siguiente:

$$\begin{aligned} p_{\text{posteriori}}(\theta) = p(\theta|y, \eta) &\stackrel{(1)}{=} \frac{p(\theta, y, \eta)}{p(y, \eta)} \\ &\stackrel{(1)}{=} \frac{p(y|\theta, \eta)p(\theta, \eta)p(\eta)}{p(y, \eta)p(\eta)} \\ &\stackrel{(1)}{=} \frac{p(y|\theta, \eta)p(\theta|\eta)}{p(y|\eta)} \stackrel{(2)}{=} \frac{p(y|\theta, \eta)p(\theta|\eta)}{\sum_{u \in \Theta} p(y, u|\eta)} \\ &\stackrel{(1)}{=} \frac{p(y|\theta, \eta)p(\theta|\eta)}{\sum_{u \in \Theta} p(y|\eta, u)p(u|\eta)} \\ &\stackrel{(3)}{=} \frac{p(y|\theta)p(\theta|\eta)}{\sum_{u \in \Theta} p(y|u)p(u|\eta)} \\ &= \frac{p(y|\theta)p_{\text{priori}}(\theta|\eta)}{\sum_{u \in \Theta} p(y|u)p_{\text{priori}}(u|\eta)} \end{aligned}$$

(1) Por densidad condicional.

(2) Se tiene que los hiperparámetros η influyen en la distribución de Y sólo a través de θ por lo tanto $p(y|\eta, \theta) = p(y|\theta)$.

(3) Propiedades de probabilidad marginal.

- b) Si η es una variable aleatoria con distribución h y espacio hiperparametral Λ , entonces se tiene

$$\begin{aligned}
 p_{\text{posteriori}}(\theta) &= p(\theta|y) \stackrel{(1)}{=} \frac{p(\theta, y)}{p(y)} \stackrel{(2)}{=} \frac{p(y, \theta)}{\sum_{u \in \Theta} p(y, u)} \\
 &\stackrel{(3)}{=} \frac{\sum_{\eta \in \Lambda} p(y, \theta, \eta)}{\sum_{\eta \in \Lambda} \sum_{u \in \Theta} p(y, u, \eta)} \\
 &\stackrel{(1)}{=} \frac{\sum_{\eta \in \Lambda} p(y|\theta, \eta)p(\theta, \eta)}{\sum_{\eta} \sum_{u \in \Theta} p(y|\eta, u)p(u, \eta)} \\
 &\stackrel{(2)}{=} \frac{\sum_{\eta \in \Lambda} p(y|\theta)p(\theta|\eta)p(\eta)}{\sum_{\eta \in \Lambda} \sum_{u \in \Theta} p(y|u)p(u|\eta)p(\eta)} \\
 &= \frac{\sum_{\eta \in \Lambda} p(y|\theta)p_{\text{priori}}(\theta|\eta)h(\eta)}{\sum_{\eta \in \Lambda} \sum_{u \in \Theta} p(y|u)p_{\text{priori}}(u|\eta)h(\eta)}
 \end{aligned}$$

□

Observación: En general la determinación de la distribución a priori se basa en la información acumulada de estudios pasados y opiniones de expertos en el área. En el caso que esto no sea posible, es decir, cuando no se tenga información fiable se considera la alternativa de una distribución que no favorezca alguna θ , la cual se precisa en la siguiente definición.

Definición 4.1.3. Sea θ el parámetro aleatorio que caracteriza un modelo bajo el cual se realiza un experimento, con espacio parametral Θ . Se define la función distribución a priori **no informativa** de θ como sigue:

- a) Si el espacio parametral es discreto y finito, es decir, $\Theta = \{\theta_1, \theta_2, \dots, \theta_n\}$.

$$p_{\text{priori}}(\theta) = P(\theta = \theta_i) = \frac{1}{n}, \quad \theta_i \in \Theta.$$

- b) En el caso que el espacio parametral sea continuo y acotado, es decir, $\Theta = [a, b]$, $-\infty < a < b < \infty$, entonces, la distribución a priori no informativa sería la uniforme continua, en el intervalo $[a, b]$. De esta forma

$$p_{\text{priori}}(\theta) = \frac{1}{b-a}, \quad \theta' \in [a, b].$$

Observación: Si el espacio parametral no es acotado pero discreto, se tiene que la situación no es tan clara, ya que la distribución a priori no informativa apropiada cumpliría que

$$p(\theta) = c \quad \text{para alguna } c > 0,$$

pero esta es inadecuada ya que:

$$\sum_{\theta' \in \Theta} p(\theta = \theta') = \infty.$$

Esto se resuelve si la suma de $p(x|\theta)$ con respecto a todos los posibles valores de θ es algún valor finito M , ya que:

$$p_{\text{posteriori}} = p(\theta|y) = \frac{p(y|\theta)c}{\sum_{\theta} p(y|\theta)c} = \frac{p(y|\theta)}{M}.$$

Entonces toda la información obtenida en la posterior surge sólo de los datos.

4.2 Inferencia Bayesiana

Como ahora ya se cuenta con una distribución para θ , entonces θ puede estimarse utilizando inferencia Bayesiana. Por ejemplo si θ es univariado, un estimador puntual de θ indicado por $\hat{\theta}$ puede ser la media, la mediana o la moda obtenidas a partir de la distribución posteriori de θ .

Como es indudable, se desea precisión en el estimador, así que una forma de estimar la exactitud del estimador puntual $\hat{\theta}$ de θ , es a través de la varianza posteriori con respecto al estimador $\hat{\theta}$, es decir,

$$E[(\theta - \hat{\theta}(y))^2 | y]$$

donde $E(\theta|y)$, es la esperanza con respecto a la distribución posteriori de θ .

Del Lema A.3 del Apéndice A, se tiene que

$$E[(\theta - \hat{\theta}(y))^2 | y] = \text{var}(\theta|y) + (E[\theta|y] - \hat{\theta}(y))^2 \quad (4.4)$$

Dado que se busca encontrar el valor de θ que minimiza (4.4), se tiene que:

$$\begin{aligned} \frac{d}{d\theta} E[(\theta - \hat{\theta}(y))^2 | y] &= 0 + 2(E[\theta|y] - \hat{\theta}(y))(-1) \\ &= -2E[(\theta|y) - \hat{\theta}(y)]. \end{aligned}$$

De esta forma

$$\frac{d}{d\theta} E[(\theta - \hat{\theta}(y))^2 | y] = 0 \quad \text{si y sólo si} \quad \hat{\theta} = E[\theta | y].$$

Note que

$$\frac{d^2}{d\theta^2} \hat{\theta} E[(\theta - \hat{\theta}(y))^2 | y] = 2 > 0.$$

Por lo tanto $\hat{\theta}(y) = E(\theta | y)$ es el valor del estimador de θ que minimiza $E[(\theta - \hat{\theta}(y))^2 | y]$. Además note que el valor mínimo alcanzado es $Var(\theta | y) = E[(\theta - E_{\theta|y}(\theta))^2]$.

Para mayor información con respecto a la inferencia Bayesiana ver por ejemplo Casella y Robert (1999), Carlin, Gelman y Rudin (1994) y Carlin y Louis (1996).

4.3 Métodos de Monte Carlo

De la Sección 4.2 se tiene que se hace uso de la estadística para aproximar ciertos resultados los cuales no son triviales de obtener. Una de las expresiones importantes en probabilidad y que tampoco siempre es posible obtener es la siguiente:

$$E[f(X)] = \sum_{x \in S} f(x)h(x),$$

donde X es una variable aleatoria con espacio de estados S y función de probabilidad h .

En este caso la $E[f(X)]$ se puede aproximar utilizando el Teorema de la Ley Fuerte de los Números grandes (ver Teorema A.2 del Apéndice A), como se describirá a continuación.

Suponga que X es una variable aleatoria definida en un espacio de probabilidad (Ω, \mathcal{F}, h) , con valores en S . Sea x_1, \dots, x_n una muestra independiente e idénticamente distribuida obtenida a partir de $h(\cdot)$. Sea $f(x)$ una función \mathcal{F} -medible, es decir, $f(X)$ es una variable aleatoria, entonces:

$$\frac{1}{N} \sum_{j=1}^N f(x_j) \tag{4.5}$$

converge por la Ley Fuerte de los Números Grandes a $E[f(x)]$, con probabilidad 1 cuando N converge a infinito.

Por lo tanto sólo se necesitaría obtener una muestra aleatoria a partir de h para estimar $E[f(X)]$, lo cual también en ocasiones se complica, así que se recurre a los métodos de Monte Carlo que proveen simulaciones del comportamiento de X con respecto a h , de los cuales se describirán dos a continuación.

(a) Muestreo por importancia

Definición 4.3.1. Sea X una variable aleatoria con función de densidad h . Sea g una función de probabilidad, tal que $g(x)$ es una aproximación para $c \cdot h(x)$, en el sentido que g tiene el comportamiento y dominio muy similar al de $c \cdot h(x)$, es decir, los puntos donde g y h asumen mayor probabilidad son los mismos. Adicionalmente suponga que es fácil de muestrear valores a partir de $g(x)$. Se define la **función de peso** como:

$$w(x) = \frac{h(x)}{g(x)}. \quad (4.6)$$

A la función $g(\cdot)$ se le llama **función de importancia**.

Proposición 4.3.1. Sea X una variable aleatoria con función de densidad $h(\cdot)$. Sea $g(\cdot)$ y $w(\cdot)$ la función de importancia y la función de peso, respectivamente, definidas como en la Definición 4.3.1 entonces

$$\lim_{N \rightarrow \infty} \frac{\frac{1}{N} \sum_{j=1}^N f(x_j) w(x_j)}{\frac{1}{N} \sum_{j=1}^N w(x_j)},$$

converge con probabilidad 1 a $E(f(X))$, donde x_j , $j = 1, 2, \dots, N$, es una muestra independientemente extraída a partir de $g(\cdot)$.

Demostración. Ver Carlin y Louis (1996). □

(b) Muestreo por rechazo

Sea X una variable aleatoria con función de densidad h y función de distribución H . Suponga que existe una M positiva y una densidad $g(x)$ tales que $h(x) < Mg(x)$, entonces los posibles valores de X serán obtenidos a través del siguiente algoritmo.

Algoritmo de método de muestreo por rechazo.

1. Generar x a partir de $g(x)$.
2. Generar u a partir de $Uniforme(0, 1)$.
3. Si $Mug(x) < h(x)$, se acepta x , en otro caso se rechaza.
4. Retornar al paso 1 y repetir hasta obtener la muestra deseada.

A $g(x)$ se le llama **función envolvente**.

Demostración. Ver Álvarez y Rodrigues (2003) y Casella y Robert (1999), entre otros. \square

Nota: Otro de los métodos de Monte Carlo es el que construye una cadena de Markov con espacio de estados igual al espacio de la variable aleatoria de la que se desea la muestra, tal que la distribución estacionaria de la cadena sea h , a éste método se le llama **método de Monte Carlo vía cadenas de Markov (MCMC)** por sus siglas en inglés). Para mayor información al respecto del método MCMC, ver por ejemplo, Casella y Robert (1999), Carlin y Louis(1996).

Capítulo 5

Inferencia ancestral

5.1 Introducción

En este capítulo final se describirá cómo utilizar el proceso de coalescencia de Kingman para realizar inferencia con respecto a los ancestros de una población. El caso presentado aquí es el estudiado por Griffiths y Tavaré (1994a, 1994b, 1995), en donde se utiliza un modelo llamado modelo de mutación con un número infinito de sitios.

El modelo de mutación con un número infinito de sitios describe el proceso de mutación en el árbol ancestral, bajo el supuesto de que una mutación en el árbol ancestral se localiza en una posición del gen donde no había ocurrido una antes.

El modelo con un número infinito de sitios es comúnmente utilizado para describir la variabilidad entre las secuencias del ADN que se observa en una muestra de genes, a través de la comparación de las secuencias del ADN en sus regiones homólogas (ver Griffiths y Tavaré (1994a, 1995)). Esto se debe a que las variaciones en las secuencias del ADN puede servir para el estudio de relaciones ancestrales entre especie. En los seres humanos la historia evolutiva se hace estudiando el ADN mitocondrial (ver Apéndice C), ya que se ha identificado en estudios genéticos que la molécula ADN mitocondrial puede proporcionar mucha información en la evolución del humano. También se ha observado que la variación entre secuencias del ADN mitocondrial es generada por los efectos de la mutación en el árbol ancestral de tales sucesiones. Se tiene además que la tasa a la cual ocurren las mutaciones en el ADN mitocondrial es casi 10 veces mayor que la tasa de ocurrencia de mutaciones en genes

que están en el núcleo de la células. De esta forma el ADN mitocondrial es muy útil para el estudio de la variación del ADN en un espacio de tiempo relativamente corto, dado que se puede encontrar diferencias en secuencias de individuos con grados de parentesco cortos. Adicionalmente el ADN mitocondrial en los humanos son casi exclusivamente heredados por la madre y por lo tanto este ADN es ideal para el estudio de árboles genealógicos del lado materno.

5.2 Los datos de la tribu india Nuu-Chah-Nult

Para ejemplificar la inferencia ancestral Griffiths y Tavaré (1994a) utilizaron unos datos muestreados de la tribu India Nuu-Chah-Nult de la isla de Vancouver, basados en registros arqueológicos, los cuales además indican que su población se ha mantenido más o menos constante por lo menos 8000 años.

Los datos de la muestra están compuestos de las secuencias del ADN obtenidos de 600 mujeres en la tribu. Las 600 secuencias produjeron 63 grupos, donde los pertenecientes a un mismo grupo eran idénticas. Cada secuencia en el ADN mitocondrial tiene 360 pares por base del segmento de la región control, la cual corresponde a las posiciones 16024-16383. La región que corresponde a los atributos comprende 201 sitios donde se tiene pirimidinas y 159 sitios donde se encuentran las purinas, 21 de los sitios pirimidina son variables (o segregados) y 5 son variables en la purina. Cada sitio del ADN es binario teniendo sólo dos posibles bases en cada sitio. Además como no hay recombinación cada sitio en la muestra tiene la misma historia ancestral.

Adicionalmente dado que se está considerando el modelo con un número infinito de sitios, entonces la muestra de 63 secuencias se reduce a 55 y de los 360 sitios quedan 352, y se tienen 18 sitios variables de los cuales 13 (del 6 al 18) le corresponde al pirimidine y 5 (del 1 al 5) a la purina.

En la Tabla 5.1, se representan los datos considerados para el análisis. En lugar de escribir los 360 pares de base, se considera solamente los 18 sitios donde por lo menos dos secuencias difieren.

Posición	1	1	2	2	3	1	1	1	1	1	2	2	2	2	2	2	Sitio		
	0	9	5	9	4	8	9	2	4	6	6	9	3	6	7	7		1	3
	6	0	1	6	4	8	1	4	9	2	6	4	3	7	1	5	9	9	
Sitio	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	
ld																			
1	A	G	G	A	A	T	C	C	T	C	T	T	C	T	C	T	T	C	2
2	A	G	G	A	A	T	C	C	T	T	T	T	C	C	T	C	T	T	2
3	G	A	G	G	A	C	C	C	T	C	T	T	C	C	C	T	T	T	1
4	G	G	A	G	A	C	C	C	C	C	T	T	C	C	C	T	T	C	3
5	G	G	G	A	A	T	C	C	T	C	T	T	C	T	C	T	T	C	19
6	G	G	G	A	G	T	C	C	T	C	T	T	C	T	C	T	T	C	1
7	G	G	G	G	A	C	C	C	T	C	C	C	C	C	C	T	T	T	1
8	G	G	G	G	A	C	C	C	T	C	C	C	C	C	C	T	T	T	1
9	G	G	G	G	A	C	C	C	T	C	T	T	C	C	C	C	C	T	4
10	G	G	G	G	A	C	C	C	T	C	T	T	C	C	C	C	T	T	8
11	G	G	G	G	A	C	C	C	T	C	T	T	C	C	C	T	T	C	5
12	G	G	G	G	A	C	C	C	T	C	T	T	C	C	C	T	T	T	4
13	G	G	G	G	A	C	C	T	T	C	T	T	C	C	C	T	T	C	3
14	G	G	G	G	A	C	C	C	C	C	T	T	C	C	C	T	T	C	1

Tabla 5.1: l.d. significa línea de descendencia y la última columna de la tabla indica el número de secuencias en la muestra en la línea de descendencia l.d.

5.3 Modelo con un número infinito de sitios

La estructura del árbol ancestral y la forma en que ocurre la mutación en el modelo de un número infinito de sitios serán descritos a continuación.

La estructura del árbol queda determinado como el coalescente de Kingman (descrito en el Capítulo 3), el cual como se puede recordar indica la relación ancestral de los individuos en una muestra. Por lo cual el número de distintos ancestros de una muestra de tamaño n al tiempo t queda representado por $D_n(t)$ y $D = \{D_n(t) : t \geq 0\}$ es un proceso de muerte pura que se mueve del estado k al $k - 1$ con tasa $k(k - 1)/2$.

Con respecto a las mutaciones se asume que ocurren como un proceso Poisson con tasa $\theta/2$ en forma independiente en cada rama del árbol, es decir, si la mutación ocurre con tasa μ por sucesión (por individuo) y por generación, entonces $\theta = 2N\mu$, donde N es el tamaño de la población donde la muestra fue obtenida.

Como ya se mencionó el modelo se basa en la comparación entre regiones homólogas, por lo cual bastará con tomar en cuenta sólo los sitios en que ocurren las mutaciones. Entonces los datos se pueden representar a través de sucesiones de los lugares donde ocurrirán las mutaciones. Como no se sabe cual es el ancestro entonces se considerará, por convención, que la base mutante es la que tiene el menor número de representantes cuando se compara el sitio considerando todas las secuencias. Por ejemplo para los datos de la Tabla 5.1 se tiene que la base mutante sería A para el sitio 1.

Se tiene entonces que un gen se describirá a partir de la secuencia de sus mutaciones, es decir, si se tienen N genes, entonces la secuencia i se describe por una secuencia $y_i = (y_{i0}, y_{i1}, \dots)$ de enteros positivos, donde cada coordenada indica el sitio donde ocurrió la mutación (ver Griffiths y Tavaré (1994a, 1995)). Como se pueden tener secuencias idénticas, entonces en lugar de expresar cada una en una secuencia, se representará la secuencia sólo una vez y se registrará por otro lado el número de veces que aparece en la muestra original, es decir, si por ejemplo se tienen d secuencias diferentes, entonces el número de veces que la secuencia i aparece en la muestra será indicado por n_i , $i = 1, 2, \dots, d$ y el vector que almacena estos valores es denotado por \bar{n} , es decir, $\bar{n} = (n_1, \dots, n_k, \dots, n_d)$.

A partir de lo anterior los datos pueden ser representados por un árbol con raíz llamado árbol condensado, el cual será denotado por T , donde cada

linaje es representado por una sola línea y las mutaciones están indicadas a lo largo de la trayectoria por sitios segregados. Entonces si en una muestra de tamaño n se tienen d distintas secuencias y_1, y_2, \dots, y_d con multiplicidades $\bar{n} = (n_1, n_2, \dots, n_d)$, se tiene que el árbol con raíz con multiplicidad de sus hojas \bar{n} , será representado por: $(T, \bar{n}) = (y_1, \dots, y_d, \bar{n})$.

Note que los datos se pueden representar a través de un árbol con raíz, siempre y cuando se conozca cual tipo es el ancestro, pero en la realidad en general no se sabe, lo cual hace que se considere un gráfico análogo al del árbol con raíz, en donde la base ancestral de cada sitio es desconocida. La representación que se estableció en este caso es llamada árbol sin raíz, en donde se representan las secuencias por vértices y el número de mutaciones por números a lo largo de los bordes entre secuencias. Al árbol sin raíz se le denota por Q . Análogamente al árbol con raíz, $(Q, \bar{n}) = (y_1, \dots, y_d, \bar{n})$ representará al árbol sin raíz con multiplicidad $\bar{n} = (n_1, n_2, \dots, n_d)$.

Se tiene entonces que cada conjunto de secuencias a partir del modelo con un número infinito de sitios le corresponde al menos un árbol con raíz (o sin raíz) y el número de sitios segregados es precisamente el número de mutaciones, que ocurrirán a lo largo del árbol. En los datos presentados en la Sección 5.2 se tiene que 18 sitios fueron segregados.

Definición 5.3.1. *A la representación de los datos en un árbol que tiene rotulados sus vértices se le llama árbol rotulado (o con rótulos), si además estos vértices tienen un particular orden al árbol se le llama árbol rotulado ordenado.*

Observación: Como en las secuencias y_1, y_2, \dots, y_d se indica donde ocurren las mutaciones, entonces $(T, \bar{n}) = (y_1, \dots, y_d, \bar{n})$ y $(Q, \bar{n}) = (y_1, \dots, y_d, \bar{n})$ son árboles rotulados y ordenados.

De la Sección 5.2 se tiene que los datos pueden ser representados de la siguiente forma:

$$\begin{array}{ll}
 y_1 = (1, 6, 4, 14, 0), & y_8 = (13, 12, 11, 18, 0) \\
 y_2 = (10, 1, 6, 4, 14, 0) & y_9 = (17, 16, 18, 0) \\
 y_3 = (2, 18, 0) & y_{10} = (16, 18, 0) \\
 y_4 = (9, 3, 0) & y_{11} = (0) \\
 y_5 = (6, 4, 14, 0) & y_{12} = (18, 0) \\
 y_6 = (5, 6, 4, 14, 0) & y_{13} = (8, 0) \\
 y_7 = (12, 11, 18, 0) & y_{14} = (7, 15, 0)
 \end{array}$$

con multiplicidad dada por $\bar{\pi} = (2, 2, 1, 3, 19, 1, 1, 1, 4, 8, 5, 4, 3, 1)$.

Observaciones:

1. Un posible árbol representando las secuencias y_1, y_2, \dots, y_{14} , es

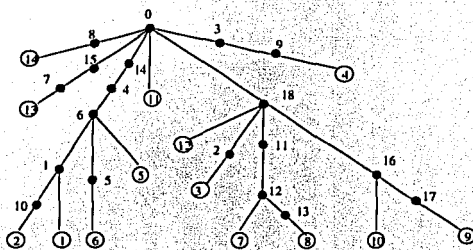


Figura 5.1: Note que en este árbol la secuencia con rótulo 11 es el ancestro de la muestra

2. Si se rotula al ancestro por 0 y al mutante por 1. Entonces los datos se pueden representar en una matriz de ceros y unos, donde cada renglón es una secuencia.

Tomando en cuenta el árbol presentado en la Figura 5.1 se tiene que en la representación de ceros y unos los datos de la Tabla 5.1 de la Sección 5.2, pueden ser representados por:

```

secuencia 1 10010100000010000
secuencia 2 100101000100010000
secuencia 3 010000000000000001
secuencia 4 001000001000000000
secuencia 5 000101000000010000
secuencia 6 000111000000010000
secuencia 7 000000000011000001
secuencia 8 000000000011100001
secuencia 9 000000000000000111
secuencia 10 0000000000000000101
secuencia 11 000000000000000000
secuencia 13 000000010000001000
secuencia 14 000000010000000000

```

3. Una observación importante que se debe contemplar es que un árbol sin raíz puede ser construido a partir de un árbol con raíz. En el caso que el linaje de la raíz no aparece en la muestra, éste no aparecerá en el árbol sin raíz. De esta forma un posible árbol sin raíz para los datos representados en la Tabla 5.1 de la Sección 5.2 es:

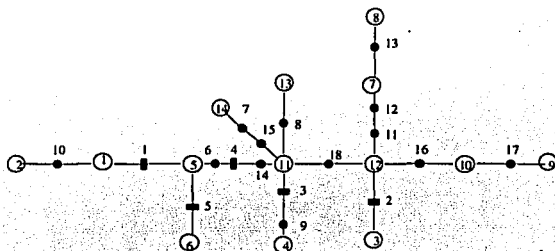


Figura 5.2: Donde • indica las mutaciones en las pirimidinas y ■ las mutaciones en el de las purinas.

Una forma de hacer las inferencias con respecto a los ancestros y tasa de mutación es a través de la probabilidad de que el árbol que describe los datos del problema tenga determinada estructura con multiplicidades \bar{n} .

Sea $P(T, \bar{n})$ la probabilidad de tener un árbol con raíz T y $P(Q, \bar{n})$ la probabilidad de tener un árbol sin raíz Q y $P^0(T, \bar{n})$ la probabilidad de tener un árbol con raíz T pero sin importar el orden en que aparecen las secuencias, donde \bar{n} es el vector de multiplicidades.

Note que si Q es el árbol sin raíz con d vértices, representando d distintas secuencias con multiplicidades $\bar{n} = (n_1, \dots, n_d)$, entonces se cumple:

$$P(Q, \bar{n}) = \sum_T P(T, \bar{n}). \quad (5.1)$$

Adicionalmente, se tiene que

$$P^0(T, n) = \frac{n!}{n_1! \dots n_d!} P(T, n). \quad (5.2)$$

Por lo tanto sólo bastará con conocer $P(T, \bar{n})$ para que se obtenga $P(Q, n)$ y $P^0(T, n)$.

Para obtener $P(T, \bar{n})$ se procederá de manera recursiva a través de observar sucesivamente los eventos más recientes (mutación o coalescencia) en el pasado. Como estos eventos podrán ser o una coalescencia o una mutación, se puede inferir cual es el tipo de historia ancestral que debe haber existido para que los eventos produzcan el árbol que representan los datos.

Teorema 5.3.1. Sea $(T, \bar{n}) = (y_1, \dots, y_d, \bar{n})$ el árbol con raíz con multiplicidad de sus hojas \bar{n} , y $P(T, \bar{n})$ la probabilidad de tener (T, \bar{n}) , entonces se cumple lo siguiente:

$$\begin{aligned}
 P(T, \bar{n}) = & \sum_{\{k:n_k \geq 2\}} \frac{n_k(n_k - 1)}{n(n - 1 + \theta)} P(T, \bar{n} - e_k) \\
 & + \sum_{\substack{\{k:n_k=1, \text{ con } y_{k_0} \neq y_{ij} \\ \text{para toda } y_1, y_2, \dots, y_d \\ \text{con } (i, j) \neq (k, 0) \text{ y} \\ S y_k \neq y_j \text{ para toda } j\}} \frac{\theta}{n(n - 1 + \theta)} P(S_k T, \bar{n}) \\
 & + \sum_{\substack{\{k:n_k=1, \text{ con } y_{k_0} \neq y_{ij} \\ \text{para toda } y_1, y_2, \dots, y_d \\ \text{con } (i, j) \neq (k, 0) \text{ y} \\ S y_k \neq y_j \text{ para toda } j\}} \sum_{\{j:S y_k = y_j\}} \frac{\theta}{n(n - 1 + \theta)} P(R_k T, R_k(\bar{n} + e_j)),
 \end{aligned}
 \tag{5.3}$$

donde e_k es el vector con todas las coordenadas cero, excepto en la k -ésima coordenada $k = 1, 2, \dots, d$, $S_k T$ indica la operación que borra la primera coordenada en el k -ésimo camino que lleva de uno de los extremos de T a la raíz y $R_k T$ es la operación que remueve el k -ésimo camino que lleva uno de los extremos de T a su raíz y $P(T_1, e_1) = 1$.

Demostración. Se tiene que una mutación o una coalescencia son los únicos eventos más recientes en el pasado posibles para que se produzcan el árbol que representan los datos.

Si se tienen n líneas ancestrales, entonces la coalescencia de dos de ellas ocurre a una tasa $n(n - 1)/2$ y la mutación en una de ellas ocurre con tasa $n\theta/2$ entonces debido a la independencia de la ocurrencia de los eventos:

- (a) La probabilidad de que ocurra una coalescencia dado que ocurrió un evento es:

$$\frac{n(n-1)/2}{n(n-1)/2 + n\theta/2} = \frac{(n-1)}{(n-1) + \theta}, \quad (5.4)$$

- (b) La probabilidad de que ocurra una mutación dado que ocurrió un evento es:

$$\frac{n\theta/2}{n(n-1)/2 + n\theta/2} = \frac{\theta}{(n-1) + \theta}. \quad (5.5)$$

Caso 1. Si el evento es la coalescencia, entonces ésta se da en las secuencias que provienen de un ancestro en común sin que ellas sean mutantes, es decir, secuencias idénticas. En este caso se debe coalescer donde la multiplicidad es mayor o igual que dos. Note que si la coalescencia ocurrió en la k -ésima clase ($k = 1, 2, \dots, d$), entonces la multilicidad de la secuencia se reduce en uno. Como e_k es el vector con todas las coordenadas cero, excepto en la k -ésima coordenada $k = 1, 2, \dots, d$. Entonces si una coalescencia ocurrió en la k -ésima clase, entonces el vector de multiplicidades \bar{n} cambia para $\bar{n} - e_k$. Sin embargo el árbol continua siendo el mismo, es decir, su estructura se preserva. Sea $(T, \bar{n} - e_k)$ el árbol T pero ahora con las multiplicidades dada por $\bar{n} - e_k$.

La probabilidad de que hubo coalescencia que está dada por (5.4). Por otro lado como se desea que haya coalescencia sin importar el orden en que coalescen se debe tomar en cuenta las distintas formas que se pueden elegir las secuencias que coalescen. Así que la probabilidad deseada queda determinada por:

$$\frac{n-1}{(n-1) + \theta} \frac{n_k(n_k - 1)}{2} \frac{2}{n(n-1)} P(T, \bar{n} - e_k), \quad \text{para } n_k \geq 2, \quad (5.6)$$

es decir, es el producto de la tasa de ocurrencia de una coalescencia por la probabilidad de que ocurra una coalescencia, por las posibles formas que la coalescencia ocurre, por la probabilidad de obtener $(T, \bar{n} - e_k)$.

Caso 2. En cuanto al evento de mutación note que se debe considerar en las secuencias donde $n_k = 1$ y que si hubo mutación es porque una nueva clase fue creada. Dado que los eventos de mutación no ocurren simultáneamente, entonces la nueva secuencia tiene multiplicidad 1.

Cuando ocurre un evento con mutación existen dos casos a considerarse; si el ancestro se encuentra en la muestra y si el ancestro no se encuentra en la muestra. Estos casos son analizados a continuación.

- (a) El ancestro está fuera de la muestra. Dado que ocurre una mutación, entonces el árbol antes que la mutación ocurra tiene una segregación menos en la secuencia donde ocurrió la mutación. Suponga que la mutación ocurrió en la secuencia k , dado que las mutaciones son listadas de acuerdo con el tiempo de ocurrencia de la más reciente hasta la más antigua, entonces se puede expresar el estado antes de que ocurra la mutación en la k -ésima secuencia a través de un operador que borra la primera coordenada de la trayectoria k . Se representará la probabilidad de este nuevo árbol por $P(S_k T, \bar{\pi})$. Como el ancestro no está en la muestra entonces esto se vale cuando $S y_k \neq y_j$ para toda j . Por lo tanto la probabilidad queda determinada por:

$$\frac{\theta}{(n-1+\theta)} \frac{1}{n} P(S_k T, \bar{\pi}), \quad (5.7)$$

para k tal que $n_k = 1$, con $y_{k_0} \neq y_{ij}$ para toda y_1, y_2, \dots, y_d con $(i, j) \neq (k, 0)$, $S y_k \neq y_j$ para toda j , es decir, la probabilidad de que haya mutación dada por (5.5) y se elija una secuencia para que mute.

- (b) Si el ancestro está en la muestra al tener una segregación menos en la k -ésima trayectoria ésta se torna igual al menos a alguna otra secuencia en la muestra por ejemplo la j -ésima. Entonces la multiplicidad de esta j -ésima secuencia es aumentada en 1 y la secuencia k es eliminada de la muestra. Como en el caso anterior esto también se puede expresar con el operador que remueve la trayectoria k de T denotado por $R_k T$, entonces la probabilidad de ocurrencia de este evento puede ser obtenida de forma análoga al caso donde el ancestro no está en la muestra. Por lo tanto la probabilidad está dada en forma análoga al caso anterior como sigue:

$$\frac{\theta}{(n-1+\theta)} \frac{1}{n} P(R_k T, R_k(\bar{\pi} + e_j)) \quad (5.8)$$

para k tal que $n_k = 1$, con $y_{k_0} \neq y_{ij}$ para toda y_1, y_2, \dots, y_d con $(i, j) \neq (k, 0)$ y para todo j , tal que $S y_k = y_j$.

Por lo tanto por (5.6), (5.7) y (5.8), se tiene:

$$\begin{aligned}
 P(T, \bar{n}) &= \sum_{\{k: n_k \geq 2\}} \frac{n_k(n_k - 1)}{n(n - 1 + \theta)} P(T, \bar{n} - e_k) \\
 &+ \sum_{\substack{\{k: n_k=1, \text{ con } y_{k_0} \neq y_{ij} \\ \text{para toda } y_1, y_2, \dots, y_d \} \\ \text{con } (i, j) \neq (k, 0) \text{ y} \\ S y_k \neq y_j \text{ para toda } j}} \frac{\theta}{n(n - 1 + \theta)} P(S_k T, \bar{n}) \\
 &+ \sum_{\substack{\{k: n_k=1, \text{ con } y_{k_0} \neq y_{ij} \\ \text{para toda } y_1, y_2, \dots, y_d \} \\ \text{con } (i, j) \neq (k, 0) \text{ y} \\ S y_k \neq y_j \text{ para toda } j}} \sum_{\{j: S y_k = y_j\}} \frac{\theta}{n(n - 1 + \theta)} P(R_k T, R_k(\bar{n} + e_j)).
 \end{aligned}$$

□

Para el caso de $P^0(T, \bar{n})$, se tiene un teorema análogo al anterior.

Teorema 5.3.2. *Sea $(T, \bar{n}) = (x_1, \dots, x_d, \bar{n})$ el árbol con raíz con multiplicidad de sus hojas \bar{n} , y $P^0(T, \bar{n})$ la probabilidad de (T, \bar{n}) sin importar el orden en que aparecen las secuencias, entonces se vale:*

$$\begin{aligned}
 P^0(T, \bar{n}) &= \sum_{\{k: n_k \geq 2\}} \frac{n_k(n_k - 1)}{n(n - 1 + \theta)} P^0(T, \bar{n} - e_k) \\
 &+ \sum_{\substack{\{k: n_k=1, \text{ con } y_{k_0} \neq y_{ij} \\ \text{para toda } y_1, y_2, \dots, y_d \} \\ \text{con } (i, j) \neq (k, 0) \text{ y} \\ S y_k \neq y_j \text{ para toda } j}} \frac{\theta}{n(n - 1 + \theta)} P^0(S_k T, \bar{n}) \\
 &+ \sum_{\substack{\{k: n_k=1, \text{ con } y_{k_0} \neq y_{ij} \\ \text{para toda } y_1, y_2, \dots, y_d \} \\ \text{con } (i, j) \neq (k, 0) \text{ y} \\ S y_k \neq y_j \text{ para toda } j}} \sum_{\{j: S y_k = y_j\}} \frac{\theta(n_j + 1)}{n(n - 1 + \theta)} P^0(R_k T, R_k(\bar{n} + e_j)),
 \end{aligned} \tag{5.9}$$

donde e_k es el vector con todas las coordenadas cero, excepto en la k -ésima coordenada $k = 1, 2, \dots, d$ y $S_k T$ indica la operación que borra la primera coordenada en el k -ésimo camino que lleva de uno de los extremos de T a

la raíz y $R_k T$ es la operación que remueve el k -ésimo camino que lleva uno de los extremos de T a su raíz y $P(T_1, e_1) = 1$.

Demostración. Como es válida (5.3), entonces se sustituirá en ésta la $P^0(T, \bar{n})$ utilizando la identidad dada por (5.2), para obtener lo que se desea.

Para facilitar los cálculos se factorizará el término en común y denotando por M_1, M_2 y M_3 al primer, segundo y tercer término de (5.3) respectivamente, se tiene que:

$$P(T, \bar{n}) = \frac{1}{n(n-1+\theta)}(M_1 + M_2 + M_3).$$

Por (5.2), se tiene:

$$P^0(T, \bar{n}) = \frac{1}{n(n-1+\theta)} \frac{n!}{n_1! \dots n_k! \dots n_d!} (M_1 + M_2 + M_3).$$

Para el primer término se tiene:

$$\begin{aligned} & \frac{n!}{n_1! \dots n_k! \dots n_d!} M_1 \\ &= \frac{n!}{n_1! \dots n_k! \dots n_d!} \sum_{\{k: n_k \geq 2\}} n_k(n_k - 1) P(T, \bar{n} - e_k) \\ &= \frac{n!}{n_1! \dots n_k! \dots n_d!} \sum_{\{k: n_k \geq 2\}} n_k(n_k - 1) \frac{n_1! \dots (n_k - 1)! \dots n_d!}{(n-1)!} P^0(T, \bar{n} - e_k) \\ &= \sum_{\{k: n_k \geq 2\}} \frac{n_k(n_k - 1)n!}{n_1! \dots n_k! \dots n_d!} \frac{n_1! \dots (n_k - 1)! \dots n_d!}{(n-1)!} P^0(T, \bar{n} - e_k) \\ &= \sum_{\{k: n_k \geq 2\}} n(n_k - 1) P^0(T, \bar{n} - e_k). \end{aligned} \tag{5.10}$$

La segunda igualdad es válida ya que las multiplicidades de $(T, \bar{n} - e_k)$ son las mismas que (T, \bar{n}) , excepto en la k -ésima secuencia, la cual es $n_k - 1$, dado que para la segunda es n_k .

En cuanto al segundo término, como el operador S_k no altera el número

de secuencias ni sus multiplicidades, entonces:

$$\begin{aligned}
 & \frac{n!}{n_1! \dots n_k! \dots n_d!} M_2 \\
 &= \frac{n!}{n_1! \dots n_k! \dots n_d!} \sum_{\substack{\{k:n_k=1, \text{ con } y_{k_0} \neq y_{ij} \\ \text{para toda } y_1, y_2, \dots, y_d\} \\ \text{con } (i,j) \neq (k,0) \text{ y} \\ S y_k \neq y_j \text{ para toda } j}} \theta P(S_k T, \bar{n}) \\
 &= \frac{n!}{n_1! \dots n_k! \dots n_d!} \sum_{\substack{\{k:n_k=1, \text{ con } y_{k_0} \neq y_{ij} \\ \text{para toda } y_1, y_2, \dots, y_d\} \\ \text{con } (i,j) \neq (k,0) \text{ y} \\ S y_k \neq y_j \text{ para toda } j}} \frac{n_1! \dots n_k! \dots n_d!}{n!} \theta P^0(S_k T, \bar{n}) \\
 &= \sum_{\substack{\{k:n_k=1, \text{ con } y_{k_0} \neq y_{ij} \\ \text{para toda } y_1, y_2, \dots, y_d\} \\ \text{con } (i,j) \neq (k,0) \text{ y} \\ S y_k \neq y_j \text{ para toda } j}} \theta P^0(S_k T, \bar{n}). \tag{5.11}
 \end{aligned}$$

En el tercer término se tiene que el operador R_k elimina la k -ésima secuencia, por lo cual la multiplicidad de ésta tendría que ser cero pero se define $0! = 1$ que sería lo mismo si $n_k! = 1$. En el caso de la j -ésima secuencia se tiene una adicional secuencia, entonces:

$$\begin{aligned}
 & \frac{n!}{n_1! \dots n_j! \dots n_d!} M_3 \\
 &= \frac{n!}{n_1! \dots n_d!} \sum_{\substack{\{k:n_k=1, \text{ con } y_{k_0} \neq y_{ij} \\ \text{para toda } y_1, y_2, \dots, y_d\} \\ \text{con } (i,j) \neq (k,0) \text{ y} \\ S y_k \neq y_j \text{ para toda } j}} \sum_{\{j:S x_k = x_j\}} \theta P(R_k T, R_k(\bar{n} + e_j)) \\
 &= \frac{n!}{n_1! \dots n_d!} \sum_{\substack{\{k:n_k=1, \text{ con } y_{k_0} \neq y_{ij} \\ \text{para toda } y_1, y_2, \dots, y_d\} \\ \text{con } (i,j) \neq (k,0) \text{ y} \\ S y_k \neq y_j \text{ para toda } j}} \sum_{\{j:S x_k = x_j\}} \theta \frac{n_1! \dots (n_j + 1)! \dots n_d!}{n!} P^0(R_k T, R_k(\bar{n} + e_j))
 \end{aligned}$$

$$\begin{aligned}
 &= \sum_{\substack{\{k: n_k=1, \text{ con } y_{k_0} \neq y_{ij} \\ \text{para toda } y_1, y_2, \dots, y_d \} \\ \text{con } (i,j) \neq (k,0) \text{ y} \\ S_{y_k} \neq y_j \text{ para toda } j}} \sum_{\{j: S_{x_k} = x_j\}} \theta \frac{n!(n_1! \dots (n_j+1)! \dots n_d!)}{(n_1! \dots n_j! \dots n_d!) n!} P^0(R_k T, R_k(\bar{n} + e_j)) \\
 &= \sum_{\substack{\{k: n_k=1, \text{ con } y_{k_0} \neq y_{ij} \\ \text{para toda } y_1, y_2, \dots, y_d \} \\ \text{con } (i,j) \neq (k,0) \text{ y} \\ S_{y_k} \neq y_j \text{ para toda } j}} \sum_{\{j: S_{x_k} = x_j\}} (n_j + 1) P^0(R_k T, R_k(\bar{n} + e_j)). \quad (5.12)
 \end{aligned}$$

Por lo tanto por (5.10), (5.11) y (5.12), se tiene lo siguiente:

$$\begin{aligned}
 P^0(T, \bar{n}) &= \sum_{\{k: n_k \geq 2\}} \frac{n(n_k - 1)}{n(n - 1 + \theta)} P^0(T, \bar{n} - e_k) \\
 &+ \sum_{\substack{\{k: n_k=1, \text{ con } y_{k_0} \neq y_{ij} \\ \text{para toda } y_1, y_2, \dots, y_d \} \\ \text{con } (i,j) \neq (k,0) \text{ y} \\ S_{y_k} \neq y_j \text{ para toda } j}} \frac{\theta}{n(n - 1 + \theta)} P^0(S_k T, \bar{n}) \\
 &+ \sum_{\substack{\{k: n_k=1, \text{ con } y_{k_0} \neq y_{ij} \\ \text{para toda } y_1, y_2, \dots, y_d \} \\ \text{con } (i,j) \neq (k,0) \text{ y} \\ S_{y_k} \neq y_j \text{ para toda } j}} \sum_{\{j: S_{y_k} = y_j\}} \frac{\theta(n_j + 1)}{n(n - 1 + \theta)} P^0(R_k T, R_k(\bar{n} + e_j)).
 \end{aligned}$$

□

Una observación importante que se debe considerar es que para muestras muy grandes (5.9) puede resultar muy costosa en cuanto al tiempo de cálculo, así que se recurre al método de Monte Carlo para obtener su aproximación. Antes de ello se presentará un resultado auxiliar.

Lema 5.3.3. Sea $q(\cdot)$ una función de probabilidad definida en $A \cup B$, tal que para $x \in A$, $q(x)$ es conocida y para $x \in B$, $q(x)$ es desconocida y sea $(r_{x,y})_{x,y \in A \cup B}$ una matriz de transición. Sea $X = \{X_n : n = 0, 1, \dots\}$ una cadena de Markov discreta homogénea en el tiempo, con espacio de estados $S = A \cup B$ y probabilidades de transición $(p_{x,y})_{x,y \in S}$, donde $p_{x,y} > 0$ siempre que $r_{x,y} > 0$. Suponga también que si X empieza en un estado del conjunto B , entonces X visita un estado en A con probabilidad 1.

Defina

$$h_{xy} = r_{xy}/p_{xy}, \quad x, y \in S.$$

Si además $q(\cdot)$ es tal que:

$$q(x) = \sum_{y \in S} r_{xy} q(y) = \sum_{y \in A} r_{xy} q(y) + \sum_{y \in B} r_{xy} q(y) \quad x \in B, \quad (5.13)$$

entonces

$$q(x) = E \left[q(x_\tau) \prod_{j=1}^{\tau} h_{x_{j-1} x_j} | X_0 = x \right] \quad x \in B, \quad (5.14)$$

donde τ es el tiempo para la primer visita en A de X .

Demostración. Ver Griffiths y Tavaré (1994a, 1994b). \square

Nota: Note que es conveniente tomar:

$$p_{xy} = \frac{r_{xy}}{f(x)}, \quad f(x) = \sum_{y \in A \cup B} r_{xy}, \quad h_{xy} = f(x). \quad (5.15)$$

ya que (5.14) sería de la siguiente forma:

$$q(x) = E \left[q(x_\tau) \prod_{j=1}^{\tau} f(x_j) | X_0 = x \right] \quad (5.16)$$

En el caso de (5.9) Griffiths y Tavaré (1994a, 1994b) muestran la apropiada cadena de Markov, la cual esta dada en el siguiente teorema.

Teorema 5.3.4. Sea $(T, \bar{n}) = (y_1, \dots, y_d, \bar{n})$ el árbol con raíz con multiplicidad de sus hojas \bar{n} , y $P^0(T, \bar{n})$ su probabilidad bajo el modelo de un número infinito de sitios. Sea $X = \{X(l) : l = 0, 1, 2, \dots\}$ la cadena de Markov con espacio de estados $S = \{x : x = (T, \bar{n})\}$ y probabilidades de transición:

$$P_{xy} = \begin{cases} \frac{n_k - 1}{J(T, \bar{n})n(n + \theta - 1)}, & \text{si } y = (T, \bar{n} - e_k) \\ \frac{\theta}{J(T, \bar{n})n(n + \theta - 1)}, & \text{si } y = (S_k T, \bar{n}) \\ \frac{\theta n_k + 1}{J(T, \bar{n})n(n + \theta - 1)}, & \text{si } y = (R_k T, R_k(\bar{n} + e_j)), \end{cases} \quad (5.17)$$

para $k=1, 2, \dots, d$ y donde

$$f(T, \bar{n}) = \sum_{\{k: n_k \geq 2\}} \frac{n_k - 1}{(n + \theta - 1)} + \frac{\theta m}{n(n + \theta - 1)},$$

y

$$\begin{aligned}
 m = & |\{k : n_k = 1, \text{ con } y_{k_0} \neq y_{ij} \text{ para toda } y_1, y_2, \dots, y_d \text{ con } (i, j) \neq (k, 0) \\
 & \text{y } S y_k \neq y_j \text{ para toda } j\}| \\
 + & \sum_{\substack{\{k: n_k=1, \text{ con } y_{k_0} \neq y_{ij} \\ \text{para toda } y_1, y_2, \dots, y_d \\ \text{con } (i, j) \neq (k, 0) \text{ y} \\ S y_k \neq y_j \text{ para toda } j\}}} \sum_{\{j: S y_k = y_j\}} (n_j + 1),
 \end{aligned}$$

entonces

$$P^0(T, \bar{n}) = E \left[\prod_{l=0}^{T-1} f(T(l), \bar{n}(l)) | X_0 = (T, \bar{n}) \right], \quad (5.18)$$

donde $X(l) = (T(l), \bar{n}(l))$ y $P(T_1, e_1) = 1$.*Demostración.* Por (5.9) se tiene que:

$$P^0(T, \bar{n}) = \sum_{(T', \bar{n}') \in S} r_{(T, \bar{n})(T', \bar{n}')} P^0(T', \bar{n}')$$

donde

$$r_{xy} = \begin{cases} \frac{n_k - 1}{(n + \theta - 1)}, & y = (T, \bar{n}) \\ \frac{\theta}{n(n + \theta - 1)}, & y = (S_k T, \bar{n}) \\ \frac{\theta n_j + 1}{n(n + \theta - 1)}, & y = (R_k T, R_k(\bar{n} + e_j)). \end{cases} \quad (5.19)$$

Tomando

$$\begin{aligned}
 f(T, \bar{n}) &= \sum_{(T^*, \bar{n}^*) \in S} r_{(T^*, \bar{n}^*)}(T^*, \bar{n}^*) \\
 &\stackrel{(1)}{=} \sum_{\{k: n_k \geq 2\}} \frac{(n_k - 1)}{(n - 1 + \theta)} \\
 &+ \sum_{\substack{\{k: n_k = 1, \text{ con } y_{k_0} \neq y_{ij} \\ \text{para toda } y_1, y_2, \dots, y_d \\ \text{con } (i, j) \neq (k, 0) \text{ y} \\ S y_k \neq y_j \text{ para toda } j\}} \frac{\theta}{n(n - 1 + \theta)} \\
 &+ \sum_{\substack{\{k: n_k = 1, \text{ con } y_{k_0} \neq y_{ij} \\ \text{para toda } y_1, y_2, \dots, y_d \\ \text{con } (i, j) \neq (k, 0) \text{ y} \\ S y_k \neq y_j \text{ para toda } j\}} \sum_{\{j: S y_k = y_j\}} \frac{\theta(n_j + 1)}{n(n - 1 + \theta)} \\
 &= \sum_{\{k: n_k \geq 2\}} \frac{n_k - 1}{(n + \theta - 1)} + \frac{\theta m}{n(n + \theta - 1)}.
 \end{aligned}$$

Entonces se cumplen las condiciones de Lema 5.3.3 dadas como en (5.15). Por lo tanto por (5.16), se tiene

$$P^0(T, \bar{n}) = E \left[\prod_{l=0}^{T-1} f(T(l), \bar{n}(l)) \mid X_0 = (T, \bar{n}) \right].$$

(1) El argumento es similar al dado para obtener (5.3). □

5.4 Estimaciones obtenidas

Como ahora se tiene $P^0(T, \bar{n})$ en términos de una esperanza, se tiene por lo visto en el Capítulo 4 que se puede aproximar esta probabilidad simulando g copias independientes del proceso $X = \{X(l) : l = 0, 1, 2, \dots\}$ y calculando para cada j , $j = 1, 2, \dots, g$, la función:

$$F_j = \prod_{l=0}^{T-1} f(T(l), \bar{n}(l)).$$

Griffiths y Tavaré (1994b) proveen un algoritmo de simulación para generar un árbol con raíz con multiplicidad \bar{n} , bajo el modelo de un número de sitios infinitos, empezando a partir de un ancestro en común como sigue:

1. Empezar con dos genes idénticos, elegidos a partir de la distribución inicial.
2. Cuando se obtiene r genes ($2 \leq r \leq n$), se elige un gen aleatoriamente y
 - (a) con probabilidad $(r-1)/(r-1+\theta)$ se añade un gen del tipo seleccionado, obteniendo $r+1$ genes,
 - (b) con probabilidad $\theta/(r-1+\theta)$ si el gen seleccionado fue el i se hace una transición al tipo j con probabilidad p_{ij} .
3. Se para hasta que se obtengan $n+1$ genes, y borrar el último gen que se duplicó, para tener una muestra de tamaño n

Para la estimación de $P(Q, \bar{n})$, se suma sobre todos los árboles con raíz para obtenerla, por (5.1).

Otro de los problemas es que tampoco se conoce el valor del parámetro θ , esto se puede resolver usando el método de Monte Carlo calculando $P^0(T, \bar{n})$ para (T, \bar{n}) fijo como una función de θ . Entonces se simula la cadena $X = \{X(t) : t = 0, 1, 2, \dots\}$ para un valor inicial θ_0 y se obtiene la probabilidad de los otros valores θ , usando:

$$P_{\theta}^0(T, \bar{n}) = E_{(T, \bar{n})}^{\theta_0} \left[\prod_{t=0}^{T-1} h_{(T(t), \bar{n}(t))}(T(t+1), \bar{n}(t+1)) \right] \quad (5.20)$$

donde h es definida por :

$$h_{(T, \bar{n})}(T^*, \bar{n}^*) = \begin{cases} f_{\theta_0}(T, \bar{n}) \frac{n+\theta_0-1}{(n+\theta-1)} & (T^*, \bar{n}^*) = (T, \bar{n} - e_k) \\ f_{\theta_0}(S_k T, \bar{n}) \frac{\theta(n+\theta_0-1)}{\theta_0(n+\theta-1)} & (T^*, \bar{n}^*) = (S_k T, \bar{n}) \\ f_{\theta_0}(R_k T, R_k(\bar{n} + e_j)) \frac{\theta(n+\theta_0-1)}{\theta_0(n+\theta-1)} & (T^*, \bar{n}^*) = (R_k T, R_k(\bar{n} + e_j)). \end{cases}$$

En lo que se refiere a la estimación de las tasa de mutación se analizará tanto la tasa total como las tasas por sitio.

Bajo el modelo de un número infinito de sitios, si se supone que las secuencias tienen m_Y sitios purine y m_R sitios pirimidine en una longitud

total de $m = m_Y + m_R$ y si se denota por μ_Y y μ_R sus respectivas tasas de mutación por base, se tiene que las tasas de mutación en la posición de A o G y en la de C o T son:

$$\theta_Y = m_Y \mu_Y, \quad \theta_R = m_R \mu_R.$$

La tasa total de mutación sería:

$$\theta = m_Y \mu_Y + m_R \mu_R$$

Las dos formas equivalentes para describir el proceso de mutación son:

1. Se supone que la mutación ocurre de acuerdo con el proceso Poisson con tasa $\theta/2$ en cada rama, entonces se tiene la mutación en el sitio Y con probabilidad $p_Y = \theta_Y/\theta$ y la mutación en el sitio R con probabilidad $p_R = \theta_R/\theta$
2. Las mutaciones en los dos tipos de sitios se comportan independientemente como un proceso Poisson con tasas $\theta_Y/2$ y $\theta_R/2$ respectivamente.

Como se tiene (5.18), entonces se puede estimar θ por máxima verosimilitud. Para los datos de la Sección 5.2 Griffiths y Tavaré (1994a) realizaron $g=200,000$ repeticiones del algoritmo de Monte Carlo para encontrar la probabilidad de los 19 árboles con raíz y obtener la del árbol sin raíz y se obtuvo por el método de máxima verosimilitud $\hat{\theta} = 4.8$. En la estimación de θ no se tiene información particular de cada sitio, sin embargo se puede modificar el algoritmo para obtener las tasas en cada sitio o también se puede analizar cada sitio por separado. De esta forma para estos datos Griffiths y Tavaré (1994a) obtuvieron seis árboles con raíz para los sitios purine y catorce árboles con raíz para los pirimidine, usando el algoritmo presentado arriba las tasas estimadas fueron:

$$\hat{\theta}_R = 1.22 \pm 0.61 \quad \hat{\theta}_Y = 3.31 \pm 1.14.$$

A partir de la estimación de θ (por MLE), se pudo por una parte comparar los árboles con raíz correspondientes al árbol sin raíz, en particular para los datos de la Sección 5.2 las probabilidades $P^0(T, \bar{\pi})$ resultaron variar entre 1.0×10^{-19} y 1.2×10^{-25} , de las cuales se obtuvo que los árboles más probables tienen la raíz entre las mutaciones 4 y 14, y entre la 4 y la 6. La probabilidad relativa que la raíz se encuentre en uno de estos sitios es de 97 por ciento.

Por otro lado se tiene que la estimación del tiempo para obtener el más reciente común antecesor (MRCA) también se puede obtener a través sólo del número de los datos como se vio en el Capítulo 3, de la forma siguiente (por (3.8)):

$$E[T_{MRCA}] = 2 \left(1 - \frac{1}{n} \right), \quad n \geq 1. \quad (5.22)$$

En cuanto al estimador de MRCA que involucra los datos se puede obtener de la siguiente manera:

$$P(T_{MRCA} \leq t | (Q, \bar{n})) = \frac{P(T_{MRCA} \leq t, (Q, \bar{n}))}{P((Q, \bar{n}))} \quad (5.23)$$

Para obtener el denominador se registra a cada movimiento de la cadena $X(t) = (T(t), n(t))$ el tiempo tomado para el movimiento. Si el número de secuencias es m cuando $X(t) = (T(t), n(t))$, entonces el tiempo de permanencia en éste estado tiene distribución exponencial con parámetro $m(m + \theta - 1)/2$.

Para $j = 1, 2, \dots, k$, se procede:

1. Se observa la cadena hasta que el estado (T_1, e_1) es alcanzado y se calcula

$$F_j(T) = \prod_{l=0}^{k-1} f(T(l), n(l)),$$

donde k es el número de pasos hasta que (T_1, e_1) es alcanzado.

2. Se registra el tiempo total $\tau_j(T)$ del anterior paso

Un estimador de $P(T_{MRCA} \leq t, (T, \bar{n}))$, si se tienen g simulaciones es entonces:

$$\frac{1}{g} \sum_{j=1}^g F_j(T) I(\tau_j(T) \leq t)$$

y se se suma ésta sobre todos los árboles con raíz se obtiene $P(T_{MRCA} \leq t, (Q, \bar{n}))$. Por lo tanto un estimador para (5.23) es:

$$\frac{\sum_T \sum_{j=1}^g F_j(T) I(\tau_j(T) \leq t)}{\sum_T \sum_{j=1}^g F_j(T)} \quad (5.24)$$

Para los datos de la Sección 5.2, se tiene que el tamaño de la muestra es $n = 55$, entonces la esperanza del T_{MRCA} dada por (5.22) es 1.96. Finalmente

note que esta estimación no esta en tiempo real. Sin embargo se puede obtener la estimación en tiempo real, a partir de conocer el tiempo para cada generación y el total de individuos. Griffiths y Tavaré (1994a) hacen la estimación en tiempo real, tomando cada generación de 20 años y como la población se estima entre 400 y 800 individuos toman $N = 600$, por lo tanto el tiempo real sería $1.96 \times 600 \times 20 \approx 23500$. Con respecto a la estimación del tiempo esperado para obtener el más reciente común ancestro utilizando los datos dado por (5.24) fue de 14400 años.

Conclusiones

Se tiene que la cantidad de investigación que se ha realizado en genética utilizando cadenas de Markov es de gran amplitud y sólo una pequeña parte es presentada en este trabajo, así que la importancia que tienen las cadenas de Markov no sólo en genética sino en múltiples áreas es trascendental, por todos los resultados que se tienen de ellas de los cuales se pudieron observar algunos en el Capítulo 1.

Cuando se desea modelar un fenómeno en el cual diversos factores actúan sobre él, se recurre a considerar primero un modelo con condiciones ideales, para analizarlo y después modificar estas condiciones hasta lograr el mayor apego a la realidad. Como los modelos de Wright-Fisher y Moran presentados es esta tesis, en los cuales las condiciones son básicas y el modelo n -coalescente hace uso del comportamiento de la población en estos modelos y a su vez el modelo con un número infinito de sitios se basa en el n -coalescente.

En este trabajo se muestra también como el uso de diversas áreas es primordial, ya que en este caso para obtener el ancestro en común se hizo uso de cadenas de Markov y de métodos de Montecarlo. En cuanto a estos métodos cabe destacar que también se han hecho aproximaciones con los métodos de Monte Carlo vía cadenas de Markov tanto en el algoritmo Metropolis-Hastings como en el de Gibbs, lo cual se podría desarrollar en otra tesis.

Si se desea hacer uso de programas presentados en la red que permiten inferir árboles ancestrales y tiempo para llegar al ancestro en común de una población, se presentan a continuación algunas direcciones donde algunos programas pueden ser encontrados.

- (a) <http://evolve.zoo.ox.ac.uk/beast/> En el programa BEAST se obtienen las simulaciones de los árboles por medio de los métodos de Monte Carlo vía cadenas de Markov.

- (b) <http://.no.embnet.org./Programs/EAL/phylip/coallike.php3> El programa COALLIKE aproxima la tasa de mutación, a partir de simulaciones que se tengan de los árboles con raíz.
- (c) <http://evolution.genetics.washington.edu/lamarc.html> En el paquete de programas LAMARC se obtienen las estimaciones de los parámetros usando también un método de Monte Carlo vía cadenas de Markov.

Apéndice A

Algunos resultados de la teoría de la probabilidad

Lema A.1. Sea (Ω, \mathcal{F}, P) un espacio de probabilidad, y sea $B \in \mathcal{F}$, tal que $P(B) \neq 0$, entonces la probabilidad condicional $P(\cdot|B)$ define una medida de probabilidad, es decir:

- a) $P(\Omega|B) = 1$.
- b) $P(A|B) \geq 0$, $A \in \mathcal{F}$.
- c) Sea $(A_n)_{n \geq 1}$ una sucesión numerable de elementos en \mathcal{F} mutuamente excluyentes, entonces:

$$P\left(\bigcup_{n=1}^{\infty} A_n|B\right) = \sum_{n=1}^{\infty} P(A_n|B).$$

Demostración. a) Es inmediata ya que

$$P(\Omega|B) = \frac{P(\Omega \cap B)}{P(B)} = 1.$$

b) Sea $A \in \mathcal{F}$, como $P(\cdot)$ es una medida, entonces

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \geq 0.$$

- c) Sea $(A_n)_{n \geq 1}$ una sucesión numerable de elementos en \mathcal{F} mutuamente excluyentes, entonces:

$$\begin{aligned} P\left(\bigcup_{n=1}^{\infty} A_n | B\right) &= \frac{P\left(\left(\bigcup_{n=1}^{\infty} A_n\right) \cap B\right)}{P(B)} = \frac{P\left(\bigcup_{n=1}^{\infty} (A_n \cap B)\right)}{P(B)} \\ &= \frac{\sum_{n=1}^{\infty} P(A_n \cap B)}{P(B)} = \sum_{n=1}^{\infty} P(A_n | B). \end{aligned}$$

Como $(A_n)_{n \geq 1}$ es una sucesión numerable de elementos en \mathcal{F} mutuamente excluyentes, entonces también $(A_n \cap B)_{n \geq 1}$ es una sucesión numerable de elementos en \mathcal{F} mutuamente excluyentes por lo tanto se cumple la cuarta igualdad, ya que $P(\cdot)$ es una medida. \square

Teorema A.1. (De la probabilidad total) Sea (Ω, \mathcal{F}, P) un espacio de probabilidad, y sean $B_1, B_2, \dots, B_n \in \mathcal{F}$ distintos del vacío, tales que forman una partición de Ω , es decir, $\bigcup_{i=1}^n B_i = \Omega$ y $B_i \cap B_j = \phi$, para toda i distinta de j . Sea $A \in \mathcal{F}$, entonces

$$P(A) = \sum_{i=1}^n P(A|B_i)P(B_i).$$

Demostración.

$$\begin{aligned} P(A) &= P(\Omega \cap A) \\ &= P\left(\left(\bigcup_{i=1}^n B_i\right) \cap A\right) = P\left(\bigcup_{i=1}^n (B_i \cap A)\right) \\ &= \sum_{i=1}^n \frac{P(B_i \cap A)P(B_i)}{P(B_i)} = \sum_{i=1}^n P(A|B_i)P(B_i). \end{aligned}$$

Como $(A_n)_{n \geq 1}$ es una sucesión numerable de elementos en \mathcal{F} mutuamente excluyentes, entonces también $(A_n \cap B)_{n \geq 1}$ es una sucesión numerable de elementos en \mathcal{F} mutuamente excluyentes por lo tanto se cumple la cuarta igualdad, ya que $P(\cdot)$ es una medida. \square

Lema A.2. Sea X una variable aleatoria con distribución geométrica con parámetro p , $p \in (0, 1)$, entonces:

$$E[X] = \frac{1}{p}.$$

Demostración. Como X tiene distribución geométrica, entonces

$$P(X = x) = q^{x-1}p \quad x = 1, 2, \dots$$

donde $0 < p < 1$. Entonces

$$\begin{aligned} E[X] &= \sum_{x=1}^{\infty} xq^{x-1}p = p \sum_{x=0}^{\infty} \frac{d}{dq} q^x \\ &= q \frac{d}{dq} \sum_{x=0}^{\infty} q^x = p \frac{d}{dq} \left(\frac{1}{1-q} \right) \\ &= \frac{p}{(1-q)^2} = \frac{1}{p}. \end{aligned}$$

□

Teorema A.2. Ley Fuerte de los Números Grandes. Sea X_1, X_2, \dots una sucesión de variables aleatorias independientes e idénticamente distribuidas, con media μ . Sea $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$, entonces

$$P\left(\lim_{n \rightarrow \infty} \bar{X}_n = \mu\right) = 1.$$

Demostración. Ver Ross (1996).

□

Lema A.3. Sean Y_1, Y_2, \dots, Y_n variables aleatorias idénticamente distribuidas con función de probabilidad $p(\cdot|\theta)$, donde θ es aleatorio. Sea $\hat{\theta}$ el estimador de θ obtenido de la distribución posteriori. Sean $E(\theta|y)$ y $\text{var}(\theta|y)$ la media y la varianza obtenidas a partir de la distribución $p_{\text{posteriori}}(\theta|y)$. Entonces

$$E[(\theta - \hat{\theta}(y))^2|y] = \text{var}(\theta|y) + E[(\theta|y) - \hat{\theta}]^2.$$

Demostración. Se tiene que:

$$\begin{aligned}
 & E_{\theta|y}\{(\theta - \hat{\theta}(y))^2\} \\
 &= E_{\theta|y}\{\theta - E(\theta|y) + E(\theta|y) - \hat{\theta}(y)\}^2 \\
 &= E_{\theta|y}\{(\theta - E(\theta|y))^2 + 2(E(\theta|y) - \hat{\theta}(y)(\theta - E(\theta|y))) + (E(\theta|y) - \hat{\theta}(y))^2\} \\
 &= E_{\theta|y}\{(\theta - E(\theta|y))^2\} \\
 &\quad + E_{\theta|y}\{2(E(\theta|y) - \hat{\theta}(y)(\theta - E(\theta|y)))\} + E_{\theta|y}\{(E(\theta|y) - \hat{\theta}(y))^2\} \\
 &= E_{\theta|y}\{(\theta - E(\theta|y))^2\} \\
 &\quad + 2(E(\theta|y) - \hat{\theta}(y)E_{\theta|y}\{(\theta - E(\theta|y))\}) + E_{\theta|y}\{(E(\theta|y) - \hat{\theta}(y))^2\} \\
 &= \text{var}_{\theta|y}(\theta) + 2(E(\theta|y) - \hat{\theta}(y)\{E_{\theta|y}\{\theta\} - E(\theta|y)\}) + (E(\theta|y) - \hat{\theta}(y))^2 \\
 &= \text{var}_{\theta|y}(\theta) + (E(\theta|y) - \hat{\theta}(y))^2.
 \end{aligned}$$

□

Apéndice B

Resultados auxiliares

Lema B.1. Sean x, y en los reales y sea M un entero positivo y considérese la siguiente suma:

$$S_M(x, y) = \binom{M}{0} v_1 v_2 \dots v_M + \binom{M}{1} u_0 v_2 \dots v_M + \dots + \binom{M}{M} u_0 u_1 \dots u_{M-1} \quad (\text{B.1})$$

donde

$$u_j = x + jy \quad y \quad v_j = 1 - x - jy,$$

entonces

$$S_M(x, y) = (1 - y)(1 - 2y) \dots (1 - My). \quad (\text{B.2})$$

Demostración. Se utilizará inducción matemática en M . Se demostrará para $M = 2$ que se vale (B.2).

Por hipótesis se tiene que

$$\begin{aligned} S_2(x, y) &= v_1 v_2 + \binom{2}{1} u_0 v_2 + u_0 u_1 \\ &= (1 - x - y)(1 - x - 2y) + 2x(1 - x - 2y) + x(x + y) \\ &= 1 - x - y - x + x^2 + xy - 2y + 2xy + 2y^2 + 2x - x^2 - 3xy \\ &= 2y^2 - 3y + 1 = (1 - y)(1 - 2y) \end{aligned}$$

Supóngase que se cumple para $n = M$, se demostrará que se cumple para

$$n = M + 1.$$

$$\begin{aligned}
 S_{M+1}(x, y) &= v_1 v_2 \dots v_{M+1} + \binom{M+1}{1} u_0 v_2 \dots v_{M+1} + \dots + \\
 &\quad + \binom{M+1}{r-1} u_0 u_1 \dots u_{r-1} v_r \dots v_{M+1} + \dots + u_0 u_1 \dots u_M \\
 &= v_1 v_2 \dots v_{M+1} + \left(\binom{M}{1} + \binom{M}{0} \right) u_0 v_2 \dots v_M + \\
 &\quad + \dots + \left(\binom{M}{r} + \binom{M}{r-1} \right) u_0 u_1 \dots u_{r-1} v_r \dots v_{M+1} + \\
 &\quad + \dots + \left(\binom{M}{M} + \binom{M}{M-1} \right) u_0 u_1 \dots u_{M-1} v_M + u_0 \dots u_M \\
 &= v_1 v_2 \dots v_{M+1} + \binom{M}{1} u_0 v_2 \dots v_M + \dots + \\
 &\quad + \binom{M}{r} + u_0 u_1 \dots u_{r-1} v_r \dots v_{M+1} + \dots + \binom{M}{M} u_0 u_1 \dots u_{M-1} v_M \\
 &\quad + \binom{M}{0} u_0 v_2 \dots v_M + \dots + \binom{M}{r-1} u_0 u_1 \dots u_{r-1} v_r \dots v_{M+1} + \\
 &\quad + \binom{M}{M-1} u_0 u_1 \dots u_{M-1} v_M + u_0 \dots u_M.
 \end{aligned}$$

Para simplificar los cálculos, la anterior suma se analizará en dos términos $S_{M+1}^{(1)}(x, y)$ y $S_{M+1}^{(2)}(x, y)$, designados de la siguiente forma:

$$\begin{aligned}
 S_{M+1}^{(1)}(x, y) &= v_1 \dots v_M v_{m+1} + \binom{M}{1} u_0 v_2 \dots v_M v_{m+1} + \dots + \\
 &\quad + \binom{M}{r} u_0 u_1 \dots u_{r-1} v_r \dots v_M v_{m+1} + \dots + \binom{M}{M} u_0 u_1 \dots u_{M-1} v_M. \\
 S_{M+1}^{(2)}(x, y) &= \binom{M}{0} u_0 v_2 \dots v_M + \dots + \binom{M}{r-1} u_0 u_1 \dots u_{r-1} v_r \dots v_{M+1} + \dots + \\
 &\quad + \binom{M}{M-1} u_0 u_1 \dots u_{M-1} v_M + u_0 \dots u_M.
 \end{aligned}$$

Primero se analizará $S_{M+1}^{(1)}(x, y)$:

$$\begin{aligned} S_{M+1}^{(1)}(x, y) &= v_1 \dots v_M v_{m+1} + \binom{M}{1} u_0 v_2 \dots v_M v_{m+1} + \dots + \\ &\quad \binom{M}{r} u_0 u_1 \dots u_{r-1} v_r \dots v_M v_{m+1} + \dots + \binom{M}{M} u_0 u_1 \dots u_{M-1} v_M \\ &= v_{m+1} \left(v_1 \dots v_M + \binom{M}{1} u_0 v_2 \dots v_M + \dots + \right. \\ &\quad \left. + \binom{M}{r} u_0 u_1 \dots u_{r-1} v_r \dots v_M + \dots + \binom{M}{M} u_0 u_1 \dots u_{M-1} \right) \\ &= v_{M+1} S_M(x, y) \end{aligned}$$

Antes de pasar a $S_{M+1}^{(2)}(x, y)$, observe que para $z = x + y$, se tiene que

$$S_M(z, y) = w_1 w_2 \dots w_M + \binom{M}{1} v_0 w_2 \dots w_M + \dots + v_0 v_1 \dots v_{M-1}.$$

donde

$$\begin{aligned} w_j &= 1 - z - jy = 1 - (x + y) - jy \\ &= 1 - x - (j + 1)y \\ &= v_{j+1} \end{aligned}$$

y

$$\begin{aligned} v_j &= z + jy = x + y + jy \\ &= x + (j + 1)y \\ &= u_{j+1}, \end{aligned}$$

entonces

$$\begin{aligned} S_M(z, y) &= \binom{M}{0} v_2 \dots v_{M+1} + \dots + \binom{M}{r-1} u_1 \dots u_{r-1} v_r \dots v_{M+1} \\ &\quad + \binom{M}{M-1} u_0 u_1 \dots u_{M-1} v_M + u_1 \dots u_M. \end{aligned}$$

Regresando con $S_{M+1}^{(2)}(x, y)$, se tiene lo siguiente:

$$\begin{aligned} S_{M+1}^{(2)}(x, y) &= \binom{M}{0} u_0 v_2 \dots v_{M+1} + \dots + \binom{M}{r-1} u_0 u_1 \dots u_{r-1} v_r \dots v_{M+1} \\ &\quad + \binom{M}{M-1} u_0 u_1 \dots u_{M-1} v_M + u_0 u_1 \dots u_M \\ &= u_0 \left(\binom{M}{0} v_2 \dots v_{M+1} + \dots + \binom{M}{r-1} u_1 \dots u_{r-1} v_r \dots v_{M+1} \right) \\ &\quad + \binom{M}{M-1} u_0 u_1 \dots u_{M-1} v_M + u_1 \dots u_M \\ &= u_0 S_M(x + y, y). \end{aligned}$$

Por lo tanto

$$\begin{aligned} S_{M+1}(x, y) &= S_{M+1}^{(1)}(x, y) + S_{M+1}^{(2)}(x, y) \\ &= v_{M+1} S_M(x, y) + u_0 S_M(x + y, y) \\ &= v_{M+1} [(1-y)(1-2y) \dots (1-My)] + u_0 [(1-y)(1-2y) \dots (1-My)] \\ &= (v_{M+1} + u_0) [(1-y)(1-2y) \dots (1-My)] \\ &\stackrel{(1)}{=} ((1-x - (M+1)y) + x) [(1-y)(1-2y) \dots (1-My)] \\ &= (1 - (M+1)y)(1-y)(1-2y) \dots (1-My), \end{aligned}$$

por lo cual el resultado se cumple.

(1) Hipótesis de inducción □

Lema B.2. Sea $x > 0$, $\ell \in \mathbb{N}$, tal que $1 \leq \ell \leq x$, entonces se vale lo siguiente

$$\Gamma(x) = \left[\prod_{k=1}^{\ell} (x-k) \right] \Gamma(x-\ell).$$

Demostración. La demostración se hará como la que se realiza por inducción. Se tiene que para $\ell = 1$ vale, ya que para $x > 0$, $\Gamma(x) = x\Gamma(x-1)$. Ahora

suponga válido para $\ell < x$, se demostrará que se vale para $\ell + 1 < x$.

$$\begin{aligned}\Gamma(x) &= \left[\prod_{k=1}^{\ell} (x-k) \right] \Gamma(x-\ell) \\ &= \left[\prod_{k=1}^{\ell} (x-k) (x-(\ell+1)) \right] \Gamma(x-(\ell+1)) \\ &= \left[\prod_{k=1}^{\ell+1} (x-k) \right] \Gamma(x-(\ell+1)).\end{aligned}$$

□

Lema B.3. Sea $N \geq 1$ y $\alpha_1, \alpha_2 \in \mathbb{R}$, tal que $0 < \alpha_1, \alpha_2 \leq 1$, entonces

(a) Para $j=1, 2, \dots, N$, se tiene

$$\Gamma\left(\frac{N(1-\alpha_2)}{1-\alpha_1-\alpha_2}\right) = \left[\prod_{k=1}^j \left(\frac{N(1-\alpha_2)}{1-\alpha_1-\alpha_2} - k\right) \right] \Gamma\left(\frac{N(1-\alpha_2)}{1-\alpha_1-\alpha_2} - j\right),$$

En particular se tiene que cuando $j = N$

$$\Gamma\left(\frac{N(1-\alpha_2)}{1-\alpha_1-\alpha_2}\right) = \left[\prod_{k=1}^N \left(\frac{N(1-\alpha_2)}{1-\alpha_1-\alpha_2} - k\right) \right] \Gamma\left(\frac{N\alpha_1}{1-\alpha_1-\alpha_2}\right).$$

(b) Se cumple también que

$$\Gamma\left(\frac{N}{1-\alpha_1-\alpha_2}\right) = \left[\prod_{k=1}^N \left(\frac{N}{1-\alpha_1-\alpha_2} - k\right) \right] \Gamma\left(\frac{N(\alpha_1 + \alpha_2)}{1-\alpha_1-\alpha_2}\right).$$

(c) Para $j=1, 2, \dots$

$$\begin{aligned}\Gamma\left(\frac{N\alpha_2}{1-\alpha_1-\alpha_2} + j\right) \\ = \left[\prod_{k=1}^j \left(\frac{N\alpha_2}{1-\alpha_1-\alpha_2} + j - k\right) \right] \Gamma\left(\frac{N\alpha_2}{1-\alpha_1-\alpha_2}\right).\end{aligned}$$

Demostración. (a) Se cumple de forma inmediata utilizando el Lema B.2, ya que $j < \frac{N(1-\alpha_2)}{1-\alpha_1-\alpha_2}$, $j = 1, \dots, N$, entonces

$$\Gamma\left(\frac{N(1-\alpha_2)}{1-\alpha_1-\alpha_2}\right) = \prod_{k=1}^j \left(\frac{N(1-\alpha_2)}{1-\alpha_1-\alpha_2} - k\right) \Gamma\left(\frac{N(1-\alpha_2)}{1-\alpha_1-\alpha_2} - j\right).$$

Por lo cual para $j = N$, se tiene que

$$\begin{aligned} & \Gamma\left(\frac{N(1-\alpha_2)}{1-\alpha_1-\alpha_2}\right) \\ &= \prod_{k=1}^N \left(\frac{N(1-\alpha_2)}{1-\alpha_1-\alpha_2} - k\right) \Gamma\left(\frac{N(1-\alpha_2)}{1-\alpha_1-\alpha_2} - N\right) \\ &= \prod_{k=1}^N \left(\frac{N(1-\alpha_2)}{1-\alpha_1-\alpha_2} - k\right) \Gamma\left(\frac{N(1-\alpha_2) - N(1-\alpha_1-\alpha_2)}{1-\alpha_1-\alpha_2}\right) \\ &= \prod_{k=1}^N \left(\frac{N(1-\alpha_2)}{1-\alpha_1-\alpha_2} - k\right) \Gamma\left(\frac{N\alpha_1}{1-\alpha_1-\alpha_2}\right). \end{aligned}$$

(b) De manera análoga por el Lema B.2, se vale

$$\begin{aligned} \Gamma\left(\frac{N}{1-\alpha_1-\alpha_2}\right) &= \prod_{k=1}^N \left(\frac{N}{1-\alpha_1-\alpha_2} - k\right) \Gamma\left(\frac{N}{1-\alpha_1-\alpha_2} - N\right) \\ &= \prod_{k=1}^N \left(\frac{N}{1-\alpha_1-\alpha_2} - k\right) \Gamma\left(\frac{N - N(1-\alpha_1-\alpha_2)}{1-\alpha_1-\alpha_2}\right) \\ &= \prod_{k=1}^N \left(\frac{N}{1-\alpha_1-\alpha_2} - k\right) \Gamma\left(\frac{N(\alpha_1 + \alpha_2)}{1-\alpha_1-\alpha_2}\right). \end{aligned}$$

(c) Este apartado también se cumple de manera inmediata por el Lema B.2,

$$\Gamma\left(\frac{N\alpha_2}{1-\alpha_1-\alpha_2+j}\right) = \left[\prod_{k=1}^j \left(\frac{N\alpha_2}{1-\alpha_1-\alpha_2} + j - k\right) \right] \Gamma\left(\frac{N\alpha_2}{1-\alpha_1-\alpha_2} - j\right).$$

□

Apéndice C

Algunos conceptos en genética

Para fines de esta tesis se describirán a continuación algunos conceptos en genética.

En los organismos vivos existen dos tipos de ácidos nucleicos, el ácido desoxirribonucleico (ADN) y el ácido ribonucleico (ARN). La función principal de los ácidos nucleicos es almacenar y transmitir información genética.

Se tiene que el DNA es el material genético de todos los organismos celulares. En las células se encuentra el ADN nuclear y el ADN mitocondrial, éste difiere del primero principalmente en lo siguiente:

- (a) Es circular en vez de lineal.
- (b) Consiste sólo de una secuencia, en vez de duplicidades.
- (c) Tiene un código ligeramente diferente.
- (d) Es exclusivamente transmitido por el lado materno.

Se tiene que el ADN mitocondrial es una molécula circular que tiene cerca de 16500 pares de base en su longitud. Una parte de esta molécula conocida como región control (D-loop), tiene 1100 pares de base y contiene promotores para la transcripción y la replicación de cada una de las ramas del ADN. En lo que se refiere a las mutaciones que pueden ocurrir en las secuencias del ADN se puede mencionar las que resultan de sustituciones de las bases A, C, G y T por otra. Los cambios que ocurren son sólo entre purinas (A, G) y pirimidinas (C, T), son conocidas como transversiones; y menos comúnmente

que las transversiones se tienen las transiciones que son cambios que ocurren solamente entre purinas o solamente entre pirimidinas.

Para mayores detalles ver Dill y Friedman (1992), Schleif (1993).

Bibliografía

- [1] Álvarez, L.J.; Rodrigues, E.R. (2003), *Bases Teóricas del Método de Monte Carlo*, Preimpreso.
- [2] Ash, R.B. (1972), *Real Analysis and Probability*, Academic Press, Inc, E.U.A.
- [3] Carlin, B.P.; Louis, T.A. (1996), *Bayes and empirical Bayes methods for data analysis*, Chapman and Hall, London.
- [4] Carlin, J.B.; Gelman, A.; Rudin, D.B. (1994), *Bayesian data analysis*, Preimpreso.
- [5] Casella, G.; Robert, C.P. (1999), *Monte Carlo statistical methods*, Springer, E.U.A.
- [6] Dill, F.J.; Friedman, J.M. (1992), *The national medical series for independent study genetics*, William & Wilkins.
- [7] Donnelly, P. (1984), The transient behaviour of the Moran model in population genetics, *Proc. Cambridge Phil. Soc.*, **95**, p. 349-350.
- [8] Feller, W. (1951), Diffusion processes in genetics, *Berkeley Symposium on Mathematical Statistics and Probability*.
- [9] Griffiths, R.C.; Tavaré, S. (1994a), Ancestral inference in population genetics, *Statistical Science*, **9**, p. 307-319.
- [10] Griffiths, R.C.; Tavaré, S. (1994b), Simulating probability distributions in the coalescent, *Theoretical Population Biology*, **46**, p. 131-59.

- [11] Griffiths, R.C.; Tavaré, S. (1995), Unrooted genealogical tree probabilities in the infinitely-many-sites model, *Mathematical Biosciences*, **127**, p. 77-98.
- [12] Karlin, S.; McGregor, J. (1964), "On some stochastic models in genetics" en *Stochastic models in Medicine and Biology*, The University of Wisconsin Press, E.U.A.
- [13] Karlin, S.; Taylor, H. (1975), *A First Course in Stochastic Processes*, Academic Press, Inc., E.U.A.
- [14] Karlin, S.; Taylor, H. (1981), *A Second Course in Stochastic Processes*, Academic Press, Inc., E.U.A.
- [15] Kingman, J. F. C. (1982a), On the genealogy of large populations, *Journal of Applied Probability*, **19A**, p. 27-43.
- [16] Kingman, J. F. C. (1982b), The coalescent, *Stochastic Processes and their Applications*, **13**, p. 235-248.
- [17] Moran, P.A.P. (1958), Random processes in genetics, *Proc. Cambridge Phil. Society*, **54**, p. 60-72.
- [18] Norris, J.R. (1997), *Markov chains*, Cambridge University Press, Reino Unido.
- [19] Rodrigues, E. (2001), *Modelos estocásticos en poblaciones genéticas*, Notas de curso-maestría.
- [20] Ross, S. (1996), *Stochastic processes*, 2a. ed., John Wiley & Sons, USA.
- [21] Ross, S. (1998), *A first course in probability*, 5a. ed., Prentice-Hall, Inc., USA.
- [22] Sánchez, R.P. (2002), *El proceso de líneas de descendencia en el estudio genético de poblaciones*, Tesis de licenciatura, Facultad de Ciencias-UNAM.
- [23] Schleich (1993), *Genetics and molecular Biology*, 2a. ed., Johns Hopkins, Baltimore, Maryland.