



UNIVERSIDAD
DE LA REPÚBLICA
URUGUAY

UNIVERSIDAD DE LA REPÚBLICA
FACULTAD DE CIENCIAS ECONÓMICAS Y DE ADMINISTRACIÓN

MONOGRAFÍA DE GRADO
LICENCIATURA EN ESTADÍSTICA

Procesos de Markov y su aplicación a la genética de poblaciones

Gerardo Martínez

Tutores:
María Inés FARIELLO y José Rafael LEÓN

23 de marzo de 2019

Agradecimientos

Admito que las secciones de agradecimientos en trabajos de fin de carrera son lo único que leo en la inmensa mayoría de ellos. Me divierte sobremanera leer ese montoncito de párrafos donde el científico deja su lenguaje técnico y sus ideas abstractas para mostrarse humano, en un acto tan íntimo como es el de agradecer. Y con esa humanidad, tan propia de cada pequeño científico, se obtienen como resultado un sinfín de posibles secciones de agradecimiento: las hay parcas en palabras, las hay elocuentes, las hay típicas, las hay idiosincráticas... Todas divertidos paréntesis dentro de trabajos que muchas veces soy incapaz de leer. Admito también que, aun antes de haber escrito menos de un par de páginas de este trabajo, ya pensaba en a quién quería agradecer y cómo. Ahora, en frente a esta tarea, no sé si terminaré agradeciendo como quería cuando recién empecé pero espero que todos los implicados lean estas líneas y reconozcan en ellas el papel que tuvieron en el desarrollo no sólo de este trabajo si no de mi vida en su totalidad.

No me imagino de qué habría tratado esta monografía si no fuera de algo relacionado a la genética de poblaciones y a la matemática. Parte de la culpa de que no me pueda imaginarme a mí mismo estudiando otra cosa la tiene Maine y, por eso, estoy muy agradecido. Gracias por haberme iniciado en este delirio que es la genética de poblaciones. Gracias, también, por haber apostado por mí en numerosas oportunidades. Un gustazo haber trabajado en esta monografía contigo. Mis disculpas por haber elegido el camino tortuoso de las difusiones: no puedo con mis ganas de llevar la contra.

Tampoco sé si podría haber escrito esta monografía sin el acompañamiento cercano de Chichi. Ha sido verdadero orgullo trabajar contigo. Fueron invaluables tus explicaciones en muchos momentos en donde sentía que cada vez sabía menos. Gracias por tenerme la paciencia de explicarme las cosas varias veces y de formas distintas: espero no haberte generado (muchos) dolores de cabeza.

Especial agradecimiento para mi familia que, a pesar de nunca entender lo que estudio (¿les cuento un secreto? Yo mucho tampoco entiendo...), me apoyaron de forma incondicional durante todos estos años. Todavía recuerdo aquella conversación en el auto:

— ¿De qué tenés clases? —me preguntó mi madre.

— De procesos estocásticos, le contesté.

Mi madre no me contestó pero su cara lo dijo todo. Espero que con este trabajo se hagan una idea de qué va esto de los procesos estocásticos.

Hay muchos amigos a los que tengo que agradecer pero por temor a importunar a alguno no voy a dar nombres: al que le quepa el sayo, que se lo ponga. Serán siempre recordadas las noches en vela, las birras compartidas, las largas charlas, los abrazos es-

trechos. Millones de gracias por todas las veces que me bancaron la cabeza en momentos en los que no podía más con el estrés de parciales y entregas. Por más noches de techno, deconstrucción de género y papelitos de locura: fueron necesarios en todo este proceso y seguirán siéndolos.

Por último, este trabajo se me hubiese hecho considerablemente cuesta arriba sin la compañía de una persona que ha ocupado un papel fundamental en este último año. Miles de gracias por las palabras de aliento, por mostrarme mis propias capacidades, por confiar en mí y por mostrarte orgulloso de que esté cerca tuyo. Espero que no te hayas agobiado con mis casi constantes tribulaciones durante el tiempo que hemos compartido. Gracias también por tus abrazos. Gracias, Diego.

Índice general

1. Introducción	6
2. Semigrupos y procesos de Markov	9
2.1. Introducción a los semigrupos	9
2.1.1. Conceptos esenciales de espacios de Banach	9
2.1.2. Introducción a la teoría de semigrupos	11
2.2. Conceptos esenciales de procesos estocásticos	17
2.2.1. Esperanza condicional	17
2.2.2. Martingalas	25
2.3. Procesos de Markov	27
3. Ecuaciones diferenciales estocásticas	31
3.1. Proceso de Wiener: Definición y propiedades básicas	31
3.2. Integral de Itô	33
3.2.1. Construcción de la integral	33
3.3. Ecuaciones diferenciales estocásticas	40
3.3.1. Aproximaciones numéricas	48
4. Modelos de genética de poblaciones	49
4.1. Conceptos de genética	49
4.2. Equilibrio de Hardy-Weinberg	51
4.3. Modelo de Wright-Fisher	54
4.3.1. Introducción y propiedades	54
4.3.2. Aproximación por difusión	58
4.3.3. Tiempo de absorción	63
4.3.4. Distribución de X_t y distribución límite	65
4.4. Modelo de Wright-Fisher con selección natural	67
4.4.1. Aproximación por difusión	69
4.4.2. Tiempo de absorción	74
4.4.3. Distribución de X_t y distribución límite	77
4.5. Modelo de Wright-Fisher con selección natural y mutación	78
4.5.1. Aproximación por difusión	80
4.5.2. Distribución estacionaria	81

5. Apéndice	84
5.1. Cadenas de Markov discretas	84
5.2. Espacio de Skorokhod	89
5.3. Integrabilidad uniforme	89
5.4. Distribución normal truncada	89
5.5. Bondad de ajuste de la difusión límite del proceso de Wright-Fisher a una normal truncada	90

Capítulo 1

Introducción

La genética de poblaciones tiene como objeto de estudio la estructura genética de las poblaciones naturales [Kimura, 1964]. Una manera de comprender esta estructura es a través del estudio de las diversas fuerzas que la determinan y que, en algunos casos, la modifican, generando cambios evolutivos. Estudiar toda la estructura genética de una población de cientos a millones de individuos desde un punto de vista teórico es una tarea si no imposible, bastante compleja. Desde la década del 1920 tras los trabajos seminales de Ronald Fisher, John Haldane y Sewall Wright, se ha analizado la composición genética de las poblaciones y su eventual evolución poniendo foco en las frecuencias de uno o unos pocos alelos (entendiendo alelo como variante de un gen) en la población. Desde este punto de vista el concepto de *evolución* será entendido como cambio en el tiempo de la frecuencia de un gen dentro de una población.

Tras el redescubrimiento (o más bien, la *resignificación* [Moore, 2001]) del trabajo de Mendel sobre plantas híbridas, gran parte del desarrollo de la genética de poblaciones ha estado enfocado en el estudio de caracteres poblacionales heredables. Los trabajos de Fisher, Haldane y Wright conformaron la forma de pensar en la genética de poblaciones teórica: el estudio de la frecuencia de un alelo heredable en una población es cuantitativo y es el resultado de una lista de supuestos sobre cómo funciona la herencia en la población estudiada. El modelo más simple desde el punto de vista matemático es el propuesto por Haldane en el cual la forma en que cambia la frecuencia de un alelo de una generación a la siguiente es determinística [Haldane, 1924]. En el conjunto de artículos *A Mathematical Theory of Natural and Artificial Selection*, Haldane estudió cómo es la evolución de un alelo cuando este está bajo una presión selectiva determinística. Este modelo sólo es aplicable cuando la población es de tamaño infinito y, por lo tanto, no es fácilmente aplicable a las poblaciones naturales. Sin embargo, un par de años atrás Fisher planteaba un modelo en donde existía la posibilidad de evolución sin presiones selectivas: se planteaba por primera vez la posibilidad de que existieran fluctuaciones aleatorias en las frecuencias de los alelos de una generación a la siguiente [Fisher, 1922].

En su trabajo *Evolution in Mendelian Populations* [Wright, 1931], Wright teoriza que debido al tamaño finito de las poblaciones naturales, sin presiones selectivas, mutación o migración, la frecuencia de los alelos puede fluctuar por puro azar. Aunque estudiar la frecuencia de un alelo con respecto al tiempo como un proceso estocástico no fue propuesto por Wright ni Fisher, la manera moderna de incluir el azar en la dinámica poblacional es

a través de procesos estocásticos. En particular, el cambio en las frecuencias génicas en poblaciones de tamaño reducido, con generaciones no solapadas y bajo neutralidad (sin mutación, migración ni selección) puede ser analizado de forma apropiada como una cadena de Márkov finita y de tiempo discreto. Este modelo llamado modelo de Wright-Fisher será el modelo central del presente trabajo por diversas razones. En primer lugar, es un modelo matemáticamente simple y con el cual se pueden ver las implicaciones que tiene el tamaño poblacional en la dinámica de la población. En segundo lugar, su simpleza permite, con poco esfuerzo, la generalización a situaciones más complejas, por ejemplo, añadiendo mutación y selección. Por último, este modelo sigue siendo relevante en estudios de genética de poblaciones modernos, a pesar del transcurso del tiempo.

Sin embargo, en general, los procesos de evolución son muy lentos y el número de individuos que componen las poblaciones es grande, por lo que será apropiado modelar la dinámica poblacional a través de un proceso de Markov continuo. Fue Feller el que presentó por primera vez un enfoque continuo del modelo de Wright-Fisher a través de un proceso de difusión [Feller, 1951]. No obstante, los mayores aportes a este análisis fueron realizados por Motoo Kimura, en una serie de artículos que culminaron en su artículo seminal *Diffusion Models in Population Genetics*. En este artículo, la frecuencia de un alelo p es observada en un intervalo de tiempo pequeño $(t, t + \delta)$ y se asume que el proceso que la gobierna es markoviano. Esta modelización, más rica pero considerablemente más complicada desde un punto de vista matemático, permite resolver preguntas que el modelo discreto no permite.

El objetivo del presente trabajo será el de estudiar el modelo de Wright-Fisher en tiempo discreto en tres casos distintos: bajo neutralidad, bajo selección y bajo mutación y selección. El desarrollo del modelo discreto resultará insuficiente para responder algunas preguntas; por lo tanto, se estudiará cómo el modelo discreto puede ser aproximado por un modelo continuo, más en particular, por un proceso de difusión. La presentación de esta aproximación diferirá de la presentada por Feller y Kimura y utilizará las herramientas propuestas por Ethier y Kurtz [Ethier y Kurtz, 2005]. El estudio del proceso de difusión límite se desarrollará mediante la teoría de semigrupos y la teoría de ecuaciones diferenciales estocásticas. Este planteo permitirá tanto un desarrollo teórico de los procesos markovianos como un desarrollo práctico y computacional a través de simulaciones.

El capítulo 2 estará dedicado al estudio de la teoría de semigrupos y a una introducción a la teoría de procesos markovianos. La primera sección estará dedicada a introducir los conceptos necesarios de teoría de operadores para poder definir el concepto de semigrupo de operadores y el concepto de generador infinitesimal de un semigrupo. Tras una breve introducción a conceptos básicos de probabilidad y procesos estocásticos, se definirán los procesos de Markov y se estudiará su vinculación a la teoría de semigrupos.

El capítulo 3 estará destinado a estudiar el concepto de ecuaciones diferenciales estocásticas. Se presentará al proceso de Wiener y algunas de sus propiedades fundamentales para luego introducir el concepto de integral de Itô. En una siguiente sección se construirá la integral de Itô en su versión más simple y ésta será utilizada para presentar el concepto de ecuación diferencial estocástica. El capítulo cierra con la vinculación entre las ecuaciones diferenciales estocásticas, los procesos markovianos y la teoría de semigrupos.

El capítulo 4 es el central del trabajo. En una primera sección presentará al modelo de Wright-Fisher neutral como una cadena de Markov discreta con espacio de estados finitos. Se estudiarán algunas de sus propiedades y se mostrará cómo las herramientas matemáticas dadas por el modelado como cadena de Markov serán insuficientes. Se probará que un reescalado en tiempo y espacio de la cadena de Markov discreta tiene como distribución límite una difusión. Con este último resultado se estudiarán el tiempo esperado a ciertos estados absorbentes y la distribución estacionaria del proceso. La siguiente sección estará dedicada a un estudio similar al realizado para el modelo de Wright-Fisher neutral pero se incluirá la selección natural al análisis. Se probará la distribución límite del reescalado en tiempo y espacio, lo que permitirá estudiar el impacto de distintos regímenes selectivos en la dinámica poblacional. El último modelo considerado consistirá en el agregado de probabilidades no nulas de mutación. Al igual que en los modelos anteriores, se mostrará que existe una difusión límite y se estudiará el impacto de las tasas de mutación en las frecuencias alélicas.

El código escrito en R que permite replicar los resultados obtenidos en esta monografía puede encontrarse en <https://github.com/WakeUpGary/Wright-Fisher>.

Capítulo 2

Semigrupos y procesos de Markov

La dinámica temporal de una cadena de Markov discreta puede ser estudiada a partir de su distribución inicial y su matriz de transiciones. Este estudio será más complejo para procesos estocásticos markovianos en tiempo continuo y espacio de estados no numerable. Para esto será necesario introducir el concepto de semigrupo y estudiar la vinculación entre semigrupos definidos en determinados espacios de funciones y su vinculación con las trayectorias de ciertos procesos.

2.1. Introducción a los semigrupos

2.1.1. Conceptos esenciales de espacios de Banach

Definición 2.1.1 (Espacio normado). *Un espacio vectorial X sobre un cuerpo $\mathbb{K}$ es un espacio normado si existe una función $\|\cdot\| : X \rightarrow \mathbb{R}$, llamada **norma**, tal que*

- 1) $\|\lambda x\| = |\lambda| \|x\|$, para todo $x \in X$ y $\lambda \in \mathbb{K}$.
- 2) $\|x + y\| \leq \|x\| + \|y\|$, para todos $x, y \in X$.
- 3) $\|x\| = 0 \Leftrightarrow x = \mathbf{0}_X$

Definición 2.1.2 (Operador lineal acotado). *Sean X, Y espacios normados sobre $\mathbb{K} = \mathbb{R}$ y $T : X \rightarrow Y$ un operador lineal. Se dice que T es un **operador acotado** si existe $K \in \mathbb{R}$ tal que para todo $x \in X$*

$$\|T(x)\| \leq K \|x\| \quad (2.1)$$

¿Cómo podemos hallar cuál es el mínimo valor de $K \in \mathbb{R}$ para el cuál se cumple (2.1)? Si $\|x\| = 0$, se tiene, por definición de norma, que $x = \mathbf{0}_X$ y $\|T(\mathbf{0}_X)\| = \|\mathbf{0}_Y\| = 0$. En este caso, cualquier valor K real haría cumplir (2.1). Podemos entonces restringir nuestra búsqueda a $X \setminus \{\mathbf{0}_X\}$. Por lo tanto, dividiendo (2.1) por $\|x\|$ se tiene que

$$\frac{\|T(x)\|}{\|x\|} \leq K$$

Luego, K debe ser por lo menos tan grande como el supremo de la expresión de la izquierda. Por lo tanto, el mínimo valor para el cual se cumple (2.1) es, efectivamente, este supremo:

$$K = \sup_{x \in X \setminus \{0\}_X} \frac{\|T(x)\|}{\|x\|}$$

Esta idea nos conduce a la siguiente definición:

Definición 2.1.3 (Norma de un operador). Sean X, Y espacios normados sobre $K = \mathbb{R}$ y $T : X \rightarrow Y$ un operador lineal. Se define la **norma del operador** T como

$$\|T\| = \sup_{x \in X \setminus \{0\}} \frac{\|T(x)\|}{\|x\|}$$

Proposición 2.1.1. Se consideran X, Y espacios normados y $T : X \rightarrow Y$ un operador lineal acotado. Entonces

- 1) Una fórmula equivalente para la norma de T es

$$\|T\| = \sup_{\substack{x \in X \\ \|x\|=1}} \|T(x)\|$$

- 2) La norma del operador T es efectivamente una norma.

Demostración. 1) Aplicando la linealidad de T en la definición de norma de un operador obtenemos que:

$$\|T\| = \sup_{x \in X \setminus \{0\}} \frac{\|T(x)\|}{\|x\|} = \sup_{x \in X \setminus \{0\}} \left\| T \left(\frac{x}{\|x\|} \right) \right\|$$

Para todo $x \in X \setminus \{0\}$ podemos considerar

$$y = \frac{x}{\|x\|}$$

De donde se concluye que

$$\|T\| = \sup_{\substack{y \in X \\ \|y\|=1}} \|T(y)\|,$$

como queríamos probar.

- 2) Probemos que $\|T\|$ es efectivamente una norma:

- a) Sea $\lambda \in \mathbb{R}$. Se tiene que

$$\|\lambda T\| = \sup_{\|x\|=1} \|\lambda T(x)\| = \sup_{\|x\|=1} |\lambda| \|T(x)\| = |\lambda| \sup_{\|x\|=1} \|T(x)\| = |\lambda| \|T\|$$

b) Sean T_1 y T_2 operadores lineales. Luego

$$\begin{aligned}\|T_1 + T_2\| &= \sup_{\|x\|=1} \|(T_1 + T_2)(x)\| \\ &\leq \sup_{\|x\|=1} \|T_1(x)\| + \sup_{\|x\|=1} \|T_2(x)\| = \|T_1\| + \|T_2\|\end{aligned}$$

c) $\|T\| = 0 \Leftrightarrow \sup_{x \in X \setminus \{0_X\}} \|T(x)\| = 0 \Leftrightarrow \forall x \in X \setminus \{0\} : T(x) = 0_Y \Leftrightarrow T = 0$

■

Observación 2.1.1. Toda norma $\|\cdot\|$ induce una métrica en X dada por

$$d(x, y) = \|x - y\|.$$

Al equipar a X con una distancia obtenemos también la topología métrica en X . Una distancia en X también permite hablar de la noción de completitud de X . Esta noción será clave en la siguiente definición:

Definición 2.1.4 (Espacio de Banach). *Un espacio normado X es un **espacio de Banach** si es un espacio completo con la métrica inducida por la norma.*

Definición 2.1.5 (Operador cerrado). *Dado X un espacio de Banach y $A : D(A) \subset X \rightarrow X$ un operador lineal, se dice que A es un **operador cerrado** si para toda sucesión $\{x_n\}_{n \in \mathbb{N}} \subset D(A)$ tal que*

$$\lim_{n \rightarrow \infty} x_n = x \quad \text{y} \quad \lim_{n \rightarrow \infty} A(x_n) = y$$

se cumple que $x \in D(A)$ y $A(x) = y$.

Definición 2.1.6 (Operador definido de forma densa). *Dado X un espacio de Banach y $A : D(A) \subset X \rightarrow X$ un operador lineal, se dice que A está definido **de forma densa** si $D(A)$ es denso en X .*

2.1.2. Introducción a la teoría de semigrupos

Si bien el concepto de semigrupo tiene sus orígenes en el álgebra, no será definido de esta manera en el presente en este trabajo. Por el contrario, este concepto será presentado desde un enfoque del análisis funcional. Para aquellos interesados en el enfoque puramente algebrista se puede consultar [Howie, 1976]. Una referencia en la que prima el enfoque funcional pero en la que no se evita el enfoque algebrista es la de [Hille y Phillips, 1957]. En la siguiente sección seguiremos el desarrollo de la teoría de semigrupos que sigue de [Engel y Nagel, 2006] y [Ethier y Kurtz, 2005].

Definición 2.1.7 (Semigrupo de operadores). *Sea X un espacio de Banach. Se define un **semigrupo de operadores** como una familia de operadores lineales acotados, $\mathcal{S} = \{S_t\}_{t \geq 0}$, definidos en X , tales que*

1. $S_0 = 1_X$, es decir, S_0 es el operador identidad en X .

$$2. S_{t_1+t_2} = S_{t_1} \circ S_{t_2}.$$

Ejemplo 2.1.1. Sea $X = C[0, \infty) = \{f : [0, \infty) \rightarrow \mathbb{R} : f \text{ es continua}\}$ con la norma del supremo; esto es

$$\|f\| = \sup_{x \in [0, +\infty)} |f(x)|$$

Definimos $\mathcal{S} = \{S_t\}_{t \geq 0}$ una familia de operadores lineales tales que

$$(S_t(f))(x) = f(x + t)$$

Veamos que $\mathcal{S}$ es un semigrupo en $C[0, +\infty)$. Probemos, en primer lugar, que para todo $t \geq 0$, $S_t : C[0, +\infty) \rightarrow C[0, +\infty)$ es un operador lineal acotado. Sea $f \in C[0, +\infty)$. Se tiene que

$$\|S_t(f)\| = \|f(x + t)\| = \sup_{(x+t) \in \mathbb{R}} |f(x + t)| = \sup_{y \in \mathbb{R}} |f(y)| = \|f\|$$

Lo que prueba que S_t es un operador acotado con constante $K = 1$.

Probemos ahora que S_0 es el operador identidad:

$$(S_0(f))(x) = f(x + 0) = f(x), \quad \forall x \in [0, +\infty)$$

Sean, a continuación, $t_1, t_2 \geq 0$. Queremos probar que $S_{t_1+t_2}(f) = (S_{t_1} \circ S_{t_2})(f)$. Luego, para todo $x \in [0, +\infty)$ se tiene que

$$(S_{t_1+t_2}(f))(x) = f(x + (t_1 + t_2))$$

y

$$(S_{t_1} \circ S_{t_2})(f)(x) = S_{t_1}(S_{t_2}(f))(x) = S_{t_1}(f)(x + t_2) = f(x + t_2 + t_1) = f(x + (t_1 + t_2))$$

Lo que prueba que $\mathcal{S}$ es un semigrupo en X .

Definición 2.1.8 (Semigrupo fuertemente continuo y contracción). Sean X un espacio de Banach y $\mathcal{S} = \{S_t\}_{t \geq 0}$ un semigrupo. Se dice que $\mathcal{S}$ es un semigrupo **fuertemente continuo** si

$$\lim_{t \rightarrow 0} S_t(f) = f$$

para toda $f \in X$. Se define una **contracción** o un **semigrupo contractivo** como un semigrupo, $\mathcal{S}$ que cumple que

$$\|S_t\| \leq 1$$

para todo $t \geq 0$.

Definición 2.1.9 (Generador infinitesimal). Sean X un espacio de Banach y $\mathcal{S} = \{S_t\}_{t \geq 0}$ un semigrupo. Se dice que $\mathcal{G} : D(\mathcal{G}) \subset X \rightarrow X$ es el generador infinitesimal de $\mathcal{S}$ si

$$\mathcal{G}(f) = \lim_{t \rightarrow 0} \frac{S_t(f) - f}{t}$$

El dominio $D(\mathcal{G})$ del operador $\mathcal{G}$ es el subconjunto de X para el cual este límite existe.

Si Δ es un intervalo real finito $[a, b]$ diremos que $u : \Delta \rightarrow X$ es Riemann integrable si dada una partición $a = t_0 \leq \dots \leq t_n = b$ y $\delta = \max_{i=1, \dots, n} (t_i - t_{i-1})$ existe

$$\lim_{\delta \rightarrow 0} \sum_{k=1}^n u(s_k)(t_k - t_{k-1})$$

con $s_k \in [t_k, t_{k+1}]$. Como es usual, si u es Riemann integrable notaremos con

$$\lim_{\delta \rightarrow 0} \sum_{k=1}^n u(s_k)(t_k - t_{k-1}) := \int_a^b u(t) dt$$

Lema 2.1.1. Sea $\mathcal{G} : D(\mathcal{G}) \subset X \rightarrow X$ el generador infinitesimal de un semigrupo $\{S_t\}_{t \geq 0}$ definido en un espacio de Banach X . Se cumplen las siguientes propiedades:

- 1) $\mathcal{G} : D(\mathcal{G}) \subset X \rightarrow X$ es un operador lineal.
- 2) Si $f \in D(\mathcal{G})$ entonces $S_t(f) \in D(\mathcal{G})$. Además

$$\frac{d}{dt} S_t(f) = S_t(\mathcal{G}(f)) = \mathcal{G}(S_t(f))$$

- 3) Para todo $t \geq 0$ y $f \in X$ se tiene que

$$\int_0^t S_u(f) du \in D(\mathcal{G})$$

- 4) Para todo $t \geq 0$ se tiene que

$$\begin{aligned} S_t(f) - f &= \mathcal{G} \left(\int_0^t S_u(f) du \right) \quad \text{si } f \in X \\ &= \int_0^t S_u(\mathcal{G}(f)) du \quad \text{si } f \in D(\mathcal{G}) \end{aligned}$$

Demostración. Ver [Engel y Nagel, 2006], p. 36. ■

A partir del lema (2.1.1) podemos dar una idea de por qué encontrar generadores infinitesimales es relevante para estudiar un semigrupo. Esto lo dará la siguiente proposición:

Proposición 2.1.2. El generador de un semigrupo fuertemente continuo es un operador lineal cerrado y definido de forma densa que determina al semigrupo de manera única.

Demostración. Sea $\mathcal{S} = \{S_t\}_{t \geq 0}$ un semigrupo fuertemente continuo definido en un espacio de Banach X . Por el lema 2.1.1 (1) el generador infinitesimal $\mathcal{G}$ es un operador lineal.

- $\mathcal{G}$ es cerrado

Sea $\{f_n\}_{n \in \mathbb{N}} \subset D(\mathcal{G})$ para la cual existen

$$\lim_{n \rightarrow \infty} f_n = f \quad \text{y} \quad \lim_{n \rightarrow \infty} \mathcal{G}(f_n) = g$$

Por definición de operador cerrado, queremos probar que $f \in D(\mathcal{G})$ y $\mathcal{G}(f) = g$. Por el lema 2.1.1 (4) tenemos que para todo $t > 0$

$$S_t(f_n) - f_n = \int_0^t S_u(\mathcal{G}(f_n)) du$$

Como $\mathcal{S}$ es un semigrupo fuertemente continuo tenemos convergencia uniforme de S_u en $[0, t]$. Luego tomando $n \rightarrow \infty$ se tiene que

$$S_t(f) - f = \int_0^t S_u(g) du$$

Dividiendo ambos lados por $1/t$ y tomando $t \rightarrow 0$ tenemos que

$$\mathcal{G}(g) = \lim_{t \rightarrow 0} \frac{S_t(f) - f}{t} = \lim_{t \rightarrow 0} \frac{\int_0^t S_u(g) du}{t} = S_0(g) = g$$

Se concluye que $\mathcal{G}(f) \in D(\mathcal{G})$ y $\mathcal{G}(f) = g$, como queríamos probar.

- $\mathcal{G}$ está definido de forma densa

Sea $f \in X$. Por el lema 2.1.1 (3) los elementos $\frac{1}{t} \int_0^t S_u(f) du$ pertenecen a $D(\mathcal{G})$ para todo $t \geq 0$. Como $\{S_u\}$ es fuertemente continuo se tiene que

$$\lim_{t \rightarrow 0} \frac{1}{t} \int_0^t S_u(f) du = f$$

Luego $D(\mathcal{G})$ es denso en X .

- $\mathcal{G}$ define $\{S_t\}_{t \geq 0}$ de forma única.

Sea $\mathcal{T} = \{T_t\}_{t \geq 0}$ otro semigrupo fuertemente continuo cuyo generador infinitesimal es $\mathcal{G}$. Para toda $f \in D(\mathcal{G})$ y $t > 0$ definimos $\eta_f : [0, t] \rightarrow X$ tal que

$$\eta_f(s) = S_{t-s}(T_s(f))$$

Para todo $s \in [0, t]$ fijo y $f \in D(\mathcal{G})$ definimos al conjunto A_s tal que

$$A_s = \left\{ \frac{T_{s+h}(f) - T_s(f)}{h} : h \in (0, 1] \right\} \cup \{\mathcal{G}(T_s(f))\}.$$

Consideramos la función $\phi : [0, 1] \rightarrow X$ tal que

$$\phi(h) = \begin{cases} \frac{T_{s+h}(f) - T_s(f)}{h} & \text{si } 0 < h \leq 1 \\ \mathcal{G}(T_s(f)) & \text{si } h = 0 \end{cases}.$$

Por definición de $\mathcal{G}$, ϕ es continua y como el intervalo $[0, 1]$ es compacto, se concluye que A_s es un subconjunto compacto de X para todo $s \in [0, t]$.

Se tiene que

$$\begin{aligned} \frac{1}{h}(\eta_f(s+h) - \eta_f(s)) &= \frac{1}{h}[S_{t-s-h}(T_{s+h}(f)) - S_{t-s}(T_s(f))] \\ &= \frac{1}{h}[S_{t-s-h}(T_{s+h}(f)) - S_{t-s-h}(T_s(f)) \\ &\quad + S_{t-s-h}(T_s(f)) - S_{t-s}(T_s(f))] \\ &= S_{t-s-h} \left(\frac{1}{h}(T_{s+h}(f) - T_s(f)) \right) \\ &\quad + \frac{1}{h}(S_{t-s-h} - S_{t-s})(T_s(f)) \end{aligned}$$

Por el lema (2.1.1 (2)) y utilizando el hecho de que A_s es un conjunto compacto, al tomar $h \rightarrow 0$ se obtiene que

$$\frac{d}{ds}\eta_f(s) = S_{t-s}(\mathcal{G}(T_s(f))) - \mathcal{G}(S_{t-s}(T_s(f))) = 0.$$

Como $\frac{d}{ds}\eta_f(s) = 0$, η_f es constante. Luego, por definición de η_f , $\eta_f(0) = S_t(f)$ y $\eta_f(t) = T_t(f)$. Se concluye que

$$T_t(f) = S_t(f)$$

para toda $f \in D(\mathcal{G})$. Luego, $T_t = S_t$ para todo $t \geq 0$, como se quería probar. ■

Ejemplo 2.1.2. Continuando con el ejemplo (2.1.1), sea $X = C[0, \infty)$ y $\mathcal{S} = \{S_t\}_{t \geq 0}$ tal que

$$(S_t f)(x) = f(x+t)$$

¿Es posible encontrar un generador infinitesimal para este semigrupo? Estamos interesados en encontrar un operador lineal $\mathcal{G} : D(\mathcal{G}) \subset C[0, \infty) \rightarrow C[0, \infty)$ tal que

$$\mathcal{G}(f) = \lim_{t \rightarrow 0} \frac{f(x+t) - f(x)}{t}$$

Se puede observar que el límite de la derecha es el *cociente incremental* y que si este límite existiese, obtendríamos la derivada de f en el punto t . Por lo tanto, el generador infinitesimal de $\mathcal{S}$ es el operador diferencial $\mathcal{G} : D(\mathcal{G}) \subset C[0, \infty) \rightarrow C[0, \infty)$ tal que

$$(\mathcal{G}f)(x) = \frac{d}{dx}f(x)$$

y luego $\mathcal{G}$ está definido para toda $f \in C^1[0, \infty)$.

Definición 2.1.10. Se consideran X un espacio de Banach y $A : D(A) \subset X$ un operador lineal cerrado en X . Un subespacio $C \subset D(A)$ se dice que es un **núcleo** de A si se cumple que la clausura de la restricción de A a D es A ; esto es

$$\overline{A|_D} = A$$

A continuación presentaremos, sin demostración, un teorema que será de suma utilidad para probar convergencia de procesos. Seguiremos la siguiente notación. Para cada $n \in \mathbb{N}$, X_n es un espacio de Banach así como también X lo es. Para cada $n \in \mathbb{N}$ definiremos como $\pi_n : X \rightarrow X_n$ una transformación lineal acotada. Notaremos con $f_n \rightarrow f$ si $f_n \in X_n$, $f \in X$ y $\lim_{n \rightarrow \infty} \|f_n - \pi_n(f)\| = 0$.

Teorema 2.1.1. Para cada $n \in \mathbb{N}$ se considera S_n una contracción en X_n , $\varepsilon_n > 0$ y $\mathcal{G}_n = \frac{1}{\varepsilon_n}(S_n - I)$ siendo I el operador identidad sobre X_n . Supongamos además que $\lim_{n \rightarrow \infty} \varepsilon_n = 0$. Definimos a $\mathcal{S} = \{S_t\}_{t \geq 0}$ un semigrupo contractivo fuertemente continuo en X con generador infinitesimal $\mathcal{G}$ y sea D un núcleo de $\mathcal{G}$. Entonces las siguientes afirmaciones son equivalentes:

- 1) Para cada $f \in X$ y todo $t \geq 0$

$$\lim_{n \rightarrow \infty} S_n^{(\lfloor 1/\varepsilon_n \rfloor)}(\pi_n(f)) = S_t(f)$$

uniformemente en todo intervalo acotado, donde $S_n^{(k)} = \underbrace{S_n \circ \cdots \circ S_n}_{k \text{ veces}}$.

- 2) Para cada $f \in X$ y todo $t \geq 0$

$$\lim_{n \rightarrow \infty} S_n^{(\lfloor 1/\varepsilon_n \rfloor)}(\pi_n(f)) = S_t(f)$$

- 3) Para toda $f \in D$ existe $f_n \in X_n$ para todo $n \in \mathbb{N}$ tal que $f_n \rightarrow f$ y $\mathcal{G}_n(f_n) \rightarrow \mathcal{G}(f)$.

Demostración. Ver [Ethier y Kurtz, 2005], p. 31. ■

2.2. Conceptos esenciales de procesos estocásticos

2.2.1. Esperanza condicional

Definición 2.2.1 (Espacio de probabilidad). Sea Ω un conjunto no vacío, $\mathcal{A}$ una σ -álgebra definida sobre Ω y $\mathbf{P} : \mathcal{A} \rightarrow [0, 1]$ una medida. Se define un **espacio de probabilidad** a la terna $(\Omega, \mathcal{A}, \mathbf{P})$.

Definición 2.2.2 (Variable aleatoria). Sea $(\Omega, \mathcal{A}, \mathbf{P})$ un espacio de probabilidad. Una función $X : \Omega \rightarrow \mathbb{R}$ es una **variable aleatoria** si

$$\{\omega \in \Omega : X(\omega) \in B\} \in \mathcal{A}$$

para todo B boreliano de $\mathbb{R}$. O, de forma equivalente, X es una variable aleatoria si

$$X^{-1}(B) \in \mathcal{A}$$

para todo B boreliano de $\mathbb{R}$.

Definición 2.2.3 (σ -álgebra generada). Dada X una variable aleatoria definida en un espacio de probabilidad $(\Omega, \mathcal{A}, \mathbf{P})$, se define $\sigma(X)$ la **σ -álgebra generada por X** al conjunto

$$\sigma(X) = \{X^{-1}(B) : B \in \mathcal{B}(\mathbb{R})\}$$

Definición 2.2.4 (esperanza condicional a un evento). Sea X una variable aleatoria con $\mathbb{E}|X| < \infty$ y $B \in \mathcal{A}$ tal que $\mathbf{P}(B) \neq 0$. Se define la **esperanza condicional de X dado B** como

$$\mathbb{E}(X|B) = \frac{1}{\mathbf{P}(B)} \int_B X d\mathbf{P}$$

Consideremos ahora una variable aleatoria X arbitraria e Y una variable aleatoria discreta, es decir, el conjunto de imágenes de Y es de la forma $\{y_1, y_2, \dots\}$. Pediremos también que $\mathbf{P}(Y = y_n) \neq 0$ para todo $n \in \mathbb{N}$. Condicionar sobre Y , entonces, será lo mismo que condicionar sobre los eventos de la forma $\{Y = y_n\}$. Como, a priori, no es posible determinar cuál de los eventos ocurrirá, consideraremos la sucesión de eventos

$$\mathbb{E}(X|\{Y = y_1\}), \mathbb{E}(X|\{Y = y_2\}), \dots$$

Una manera de hacer esto es construir una nueva variable aleatoria que sea constante e igual a $\mathbb{E}(X|\{Y = y_n\})$ en cada conjunto $\{Y = y_n\}$.

Definición 2.2.5 (esperanza condicional sobre una variable discreta). Sea X una variable aleatoria arbitraria con $\mathbb{E}|X| < \infty$ e Y una variable aleatoria que toma valores $\{y_1, y_2, \dots\}$. Definimos la **esperanza condicional de X dado Y** a la variable aleatoria $\mathbb{E}(X|Y)$ tal que

$$\mathbb{E}(X|Y)(\omega) = \mathbb{E}(X|\{Y = y_n\}) \text{ si } Y(\omega) = y_n$$

para todo $n \in \mathbb{N}$.

Proposición 2.2.1. Sea X una variable aleatoria arbitraria con $\mathbb{E}|X| < \infty$ e Y una variable aleatoria discreta, entonces

- 1) $\mathbb{E}(X|Y)$ es $\sigma(Y)$ -medible.
- 2) Para todo $A \in \sigma(Y)$

$$\int_A \mathbb{E}(X|Y) d\mathbf{P} = \int_A X d\mathbf{P}$$

Demostración. 1) Sea $\{y_1, y_2, \dots\}$ el recorrido de Y . Por definición $\sigma(Y)$ es la σ -álgebra generada por los conjuntos de la forma $\{Y = y_n\}$. Sea $B \in \sigma(Y)$. Existen, por lo tanto, un subconjunto $J \subset \mathbb{N}$ tal que

$$B = \bigcup_{j \in J} \{\omega \in \Omega : Y(\omega) = y_j\}.$$

Luego

$$\begin{aligned} \mathbb{E}(X|Y)^{-1}(B) &= \mathbb{E}(X|Y)^{-1} \left(\bigcup_{j \in J} \{\omega \in \Omega : Y(\omega) = y_j\} \right) \\ &= \bigcup_{j \in J} \mathbb{E}(X|Y)^{-1}(\{\omega \in \Omega : Y(\omega) = y_j\}) \\ &= \bigcup_{j \in J} \{\omega \in \Omega : Y(\omega) = y_j\} \in \sigma(Y) \end{aligned}$$

Lo que concluye que $\mathbb{E}(X|Y)$ es $\sigma(Y)$ -medible.

- 2) Para cada $n \in \mathbb{N}$ se tiene que

$$\begin{aligned} \int_{\{Y=y_n\}} \mathbb{E}(X|Y) d\mathbf{P} &= \int_{\{Y=y_n\}} \mathbb{E}(X|\{Y=y_n\}) d\mathbf{P} \\ &= \int_{\{Y=y_n\}} \left(\frac{1}{\mathbf{P}(Y=y_n)} \int_{\{Y=y_n\}} X d\mathbf{P} \right) d\mathbf{P} \\ &= \left(\frac{1}{\mathbf{P}(Y=y_n)} \int_{\{Y=y_n\}} X d\mathbf{P} \right) \underbrace{\int_{\{Y=y_n\}} d\mathbf{P}}_{\mathbf{P}(Y=y_n)} \\ &= \int_{\{Y=y_n\}} X d\mathbf{P} \end{aligned}$$

Como todo $A \in \sigma(Y)$ es la unión numerable y disjunta de conjuntos de la forma $\{Y = y_n\}$ se concluye que

$$\int_A \mathbb{E}(X|Y) d\mathbf{P} = \int_A X d\mathbf{P}$$

para todo $A \in \sigma(Y)$. ■

Motivados por el teorema anterior podemos definir el concepto de esperanza condicional dadas dos variables aleatorias X e Y arbitrarias:

Definición 2.2.6 (Esperanza condicional sobre una variable aleatoria). *Sea X una variable aleatoria tal que $\mathbb{E}|X| < \infty$ e Y una variable aleatoria arbitraria. Se define la esperanza condicional de X dado Y a una variable aleatoria, $\mathbb{E}(X|Y)$, que cumple que*

1. $\mathbb{E}(X|Y)$ es $\sigma(Y)$ -medible.
2. Para todo $A \in \sigma(Y)$

$$\int_A \mathbb{E}(X|Y) d\mathbf{P} = \int_A X d\mathbf{P}$$

Definición 2.2.7 (Sub- σ -álgebra). *Sea $(\Omega, \mathcal{A})$ un espacio medible. Se dice que $\mathcal{S}$ es una sub- σ -álgebra de $\mathcal{A}$ si*

1. $\mathcal{S} \subset \mathcal{A}$ y,
2. $\mathcal{S}$ es una σ -álgebra de Ω .

Definición 2.2.8 (Esperanza condicional sobre una sub- σ álgebra). *Sea X una variable aleatoria definida en un espacio de probabilidad $(\Omega, \mathcal{A}, P)$ tal que $\mathbb{E}|X| < \infty$. Se considera $\mathcal{S}$ una sub- σ -álgebra de $\mathcal{A}$. Se define la esperanza condicional de X dado $\mathcal{S}$ a una variable aleatoria, $\mathbb{E}(X|\mathcal{S})$, que cumple que*

1. $\mathbb{E}(X|\mathcal{S})$ es $\mathcal{S}$ -medible.
2. Para todo $A \in \mathcal{S}$

$$\int_A \mathbb{E}(X|\mathcal{S}) d\mathbf{P} = \int_A X d\mathbf{P}$$

Observación 2.2.1. La noción de esperanza condicional sobre una sub- σ -álgebra extiende la noción de esperanza condicional sobre una variable aleatoria arbitraria Y en el sentido de que

$$\mathbb{E}(X|\sigma(Y)) = \mathbb{E}(X|Y)$$

Observación 2.2.2. Dados X una variable aleatoria definida en un espacio de probabilidad $(\Omega, \mathcal{A}, P)$ y $\mathcal{S}$ una sub- σ -álgebra de $\mathcal{A}$, la esperanza $\mathbb{E}(X|\mathcal{S})$ es única. La prueba de este resultado puede verse en [Shiryayev, 1996].

Lema 2.2.1. Se consideran X, Y variables aleatorias definidas en un espacio de probabilidad $(\Omega, \mathcal{A}, \mathbf{P})$ y tales que $\mathbb{E}|X| < \infty$ y $\mathbb{E}|Y| < \infty$. Sea $\mathcal{S}$ una sub- σ -álgebra tal que X e Y son medibles en $\mathcal{S}$. Entonces, son equivalentes:

- 1) $X = Y$ casi seguramente.
- 2) Para todo $A \in \mathcal{S}$ se cumple que $\int_A X d\mathbf{P} = \int_A Y d\mathbf{P}$.

Demostración. $(\Rightarrow)$ Sea $B = \{\omega \in \Omega : X(\omega) \neq Y(\omega)\}$. Por hipótesis sabemos que $\mathbf{P}(B) = 0$. Luego

$$\int_A X d\mathbf{P} = \int_A Y d\mathbf{P}$$

para todo $A \in \mathcal{S}$.

$(\Leftarrow)$ Supongamos que $X \neq Y$ casi seguramente.

Para cada $n \in \mathbb{N}$ definimos

$$\begin{aligned} A_n &= \{\omega \in \Omega : (X - Y)(\omega) > 1/n\} \\ B_n &= \{\omega \in \Omega : (Y - X)(\omega) > 1/n\} \end{aligned}$$

Como $X \neq Y$ casi seguramente, existe $n \in \mathbb{N}$ tal que $\mathbf{P}(A_n) > 0$ o existe $m \in \mathbb{N}$ tal que $\mathbf{P}(B_m) > 0$. Supongamos que existe $n \in \mathbb{N}$ tal que $\mathbf{P}(A_n) > 0$. Entonces

$$\int_{A_n} X d\mathbf{P} - \int_{A_n} Y d\mathbf{P} = \int_{A_n} (X - Y) d\mathbf{P} \geq \frac{\mathbf{P}(A_n)}{n} > 0$$

Y, por lo tanto,

$$\int_{A_n} X d\mathbf{P} \neq \int_{A_n} Y d\mathbf{P}$$

Lo cual es una contradicción con el hecho de que $\int_A X d\mathbf{P} = \int_A Y d\mathbf{P}$. Luego

$$\mathbf{P}(X - Y > 0) = \mathbf{P}\left(\bigcup_{n \in \mathbb{N}} A_n\right) \leq \sum_{n \in \mathbb{N}} \mathbf{P}(A_n) = 0$$

Supongamos ahora que existe $m \in \mathbb{N}$ tal que $\mathbf{P}(B_m) > 0$. Se tiene que

$$\int_{B_m} Y d\mathbf{P} - \int_{B_m} X d\mathbf{P} = \int_{B_m} (Y - X) d\mathbf{P} \geq \frac{\mathbf{P}(B_m)}{m} > 0$$

Y, por lo tanto,

$$\int_{B_m} X d\mathbf{P} \neq \int_{B_m} Y d\mathbf{P}$$

Lo cual contradice nuestra hipótesis. Entonces

$$\mathbf{P}(Y - X > 0) = \mathbf{P}\left(\bigcup_{n \in \mathbb{N}} B_n\right) \leq \sum_{n \in \mathbb{N}} P(B_n) = 0$$

De donde se concluye que

$$\mathbf{P}(\{\omega \in \Omega : X(\omega) \neq Y(\omega)\}) = 0$$

■

Proposición 2.2.2. Sea X una variable aleatoria definida en un espacio de probabilidad $(\Omega, \mathcal{A}, \mathbf{P})$ tal que $\mathbb{E}|X| < \infty$ y $\mathcal{S}$ una sub- σ -álgebra tal que X es $\mathcal{S}$ -medible. Entonces

$$\mathbb{E}(X|\mathcal{S}) = X$$

casi seguramente.

Demostración. Por definición de esperanza condicional, $\mathbb{E}(X|\mathcal{S})$ es $\mathcal{S}$ -medible y para todo $A \in \mathcal{S}$

$$\int_A \mathbb{E}(X|\mathcal{S}) d\mathbf{P} = \int_A X d\mathbf{P}$$

Luego, por el lema (2.2.1), $\mathbb{E}(X|\mathcal{S}) = X$ casi seguramente. ■

Proposición 2.2.3 (Linealidad de la esperanza condicional). Sean X, Y variables aleatorias definidas en un espacio de probabilidad $(\Omega, \mathcal{A}, \mathbf{P})$ tales que $\mathbb{E}|X| < \infty$ y $\mathbb{E}|Y| < \infty$, y $\mathcal{S}$ una sub σ -álgebra de $\mathcal{A}$. Se cumple que

$$\mathbb{E}(aX + bY|\mathcal{S}) = a\mathbb{E}(X|\mathcal{S}) + b\mathbb{E}(Y|\mathcal{S})$$

Demostración. Sea $B \in \mathcal{S}$. Se tiene que

$$\begin{aligned} \int_B a\mathbb{E}(X|\mathcal{S}) + b\mathbb{E}(Y|\mathcal{S}) d\mathbf{P} &= a \int_B \mathbb{E}(X|\mathcal{S}) d\mathbf{P} + b \int_B \mathbb{E}(Y|\mathcal{S}) d\mathbf{P} \\ &= a \int_B X d\mathbf{P} + b \int_B Y d\mathbf{P} \\ &= \int_B aX + bY d\mathbf{P}, \end{aligned}$$

donde en la penúltima igualdad se utilizó la definición de $\mathbb{E}(X|Y)$. ■

Proposición 2.2.4 (Propiedad de la torre). Sea X una variable aleatoria tal que $\mathbb{E}|X| < \infty$ definida en $(\Omega, \mathcal{A}, \mathbf{P})$, $\mathcal{S}$ una sub- σ -álgebra de $\mathcal{A}$ y $\mathbb{E}(X|\mathcal{S})$ la esperanza condicional de X dado $\mathcal{S}$. Luego, para toda sub- σ -álgebra $\mathcal{B} \subset \mathcal{S}$ se cumple que

$$\mathbb{E}(\mathbb{E}(X|\mathcal{S})|\mathcal{B}) = \mathbb{E}(X|\mathcal{B})$$

Demostración. Por definición de esperanza condicional se tiene que para todo $A \in \mathcal{S}$

$$\int_A \mathbb{E}(X|\mathcal{S}) d\mathbf{P} = \int_A X d\mathbf{P}$$

y para todo $A \in \mathcal{B}$ se tiene

$$\int_A \mathbb{E}(X|\mathcal{B}) d\mathbf{P} = \int_A X d\mathbf{P}.$$

Como $\mathcal{B} \subset \mathcal{S}$ se tiene que

$$\int_A \mathbb{E}(X|\mathcal{S}) d\mathbf{P} = \int_A \mathbb{E}(X|\mathcal{B}) d\mathbf{P},$$

para todo $A \in \mathcal{B}$. Luego, por definición de esperanza condicional, $\mathbb{E}(\mathbb{E}(X|\mathcal{S})|\mathcal{B})$ cumple que

$$\int_A \mathbb{E}(\mathbb{E}(X|\mathcal{S})|\mathcal{B}) d\mathbf{P} = \int_A \mathbb{E}(X|\mathcal{S}) d\mathbf{P},$$

para todo $A \in \mathcal{B}$, de donde se concluye que

$$\mathbb{E}(\mathbb{E}(X|\mathcal{S})|\mathcal{B}) = \mathbb{E}(X|\mathcal{B})$$

■

Corolario 2.2.1 (Ley de la esperanza total). Sean X, Y variables aleatorias tales que $\mathbb{E}|X| < \infty$ y $\mathbb{E}|X| < \infty$ definidas en $(\Omega, \mathcal{A}, \mathbf{P})$. Entonces se cumple que

$$\mathbb{E}(\mathbb{E}(X|Y)) = \mathbb{E}(X)$$

Demostración. Sea $\mathcal{S} = \sigma(Y)$ y $\mathcal{B} = \{\emptyset, \Omega\} \subset \mathcal{S}$.

Por la proposición (2.2.4) se tiene que

$$\mathbb{E}(\mathbb{E}(X|\mathcal{S})|\mathcal{B}) = \mathbb{E}(X|\mathcal{B})$$

Por definición de esperanza condicional se tiene que

$$\mathbb{E}(X|Y) = \mathbb{E}(X|\sigma(Y)) = \mathbb{E}(X|\mathcal{S}).$$

Por otra parte, se observa que, como $\mathcal{B} = \{\emptyset, \Omega\}$,

$$\mathbb{E}(\mathbb{E}(X|Y)|\mathcal{B}) = \int_{\Omega} \mathbb{E}(X|Y) d\mathbf{P} = \mathbb{E}(\mathbb{E}(X|Y))$$

y

$$\mathbb{E}(X|\mathcal{B}) = \int_{\Omega} X d\mathbf{P} = \mathbb{E}(X)$$

Se concluye que

$$\mathbb{E}(\mathbb{E}(X|Y)) = \mathbb{E}(X)$$

como queríamos probar. ■

Proposición 2.2.5. Sea $\{Y_n\}_{n \in \mathbb{N}}$ una sucesión de variables aleatorias tales que $|Y_n| \leq X$, $\mathbb{E}|X| < \infty$ y $Y_n \xrightarrow{cs} Y$, entonces

$$\lim_{n \rightarrow \infty} \mathbb{E}(Y_n | \mathcal{S}) = \mathbb{E}(Y | \mathcal{S})$$

Demostración. Ver [Shiryaev, 1996], p. 218. ■

Proposición 2.2.6. Sean X, Y variables aleatorias definidas en un espacio de probabilidad con $\mathbb{E}|X| < \infty$ y $\mathbb{E}|Y| < \infty$, y $\mathcal{S}$ una sub σ -álgebra de $\mathcal{A}$. Supongamos además que $\mathbb{E}(XY) < \infty$ y que Y es $\mathcal{S}$ -medible. Entonces se cumple que

$$\mathbb{E}(XY | \mathcal{S}) = Y \mathbb{E}(X | \mathcal{S})$$

Demostración. Sea X una variable aleatoria arbitraria tal que $\mathbb{E}(|X|) < \infty$. Supongamos que $Y = \mathbb{1}_A$ con $A \in \mathcal{S}$. Probemos que

$$\mathbb{E}(X \mathbb{1}_A | \mathcal{S}) = \mathbb{1}_A \mathbb{E}(X | \mathcal{S}) \quad (2.2)$$

Sea $B \in \mathcal{S}$. Se tiene que

$$\begin{aligned} \int_B \mathbb{1}_A X \, d\mathbf{P} &= \int_{A \cap B} X \, d\mathbf{P} \\ &= \int_{A \cap B} \mathbb{E}(X | \mathcal{S}) \, d\mathbf{P} \\ &= \int_B \mathbb{1}_A \mathbb{E}(X | \mathcal{S}) \, d\mathbf{P}. \end{aligned}$$

donde en la penúltima igualdad se utilizó la definición de esperanza condicional. Luego,

$$\mathbb{E}(X \mathbb{1}_A | \mathcal{S}) = \mathbb{1}_A \mathbb{E}(X | \mathcal{S}) \quad (2.3)$$

Supongamos ahora que Y es una variable aleatoria de la forma

$$Y = \sum_{i=1}^n a_i \mathbb{1}_{A_i}$$

con $a_1, \dots, a_n \in \mathbb{R}$ y $A_1, \dots, A_n \in \mathcal{S}$ —a este tipo de variables aleatorias las llamamos *variables aleatorias simples*. Entonces

$$\begin{aligned} \mathbb{E}(XY | \mathcal{S}) &= \mathbb{E}\left(X \left(\sum_{i=1}^n a_i \mathbb{1}_{A_i}\right) \middle| \mathcal{S}\right) \\ &= \mathbb{E}\left(\sum_{i=1}^n a_i X \mathbb{1}_{A_i} \middle| \mathcal{S}\right) \\ &= \sum_{i=1}^n a_i \mathbb{E}(X \mathbb{1}_{A_i} | \mathcal{S}) \quad (\text{proposición 2.2.3}) \end{aligned}$$

$$\begin{aligned}
&= \sum_{i=1}^n a_i \mathbf{1}_{A_i} \mathbb{E}(X|\mathcal{S}) \\
&= Y \mathbb{E}(X|\mathcal{S})
\end{aligned} \tag{2.3}$$

Por lo tanto, para toda variable simple $\mathcal{S}$ -medible, Y , se cumple que

$$\mathbb{E}(XY|\mathcal{S}) = Y \mathbb{E}(X|\mathcal{S}) \tag{2.4}$$

Consideremos ahora a Y una variable aleatoria arbitraria con $\mathbb{E}(|Y|) < \infty$ y sea $\{Y_n\}_{n \in \mathbb{N}}$ una sucesión de variables aleatorias simples $\mathcal{S}$ -medibles y tales que

$$|Y_n| \leq Y \quad \text{y} \quad \lim_{n \in \mathbb{N}} Y_n = Y \quad \text{cs}$$

Por la ecuación (2.4) tenemos que, para todo $n \in \mathbb{N}$:

$$\mathbb{E}(XY_n|\mathcal{S}) = Y_n \mathbb{E}(X|\mathcal{S}).$$

Como $|XY_n| \leq |XY|$, por la proposición (2.2.5) se tiene que

$$\mathbb{E}(XY_n|\mathcal{S}) \xrightarrow{\text{cs}} \mathbb{E}(XY|\mathcal{S}).$$

Se concluye que

$$Y_n \mathbb{E}(X|\mathcal{S}) \xrightarrow{\text{cs}} Y \mathbb{E}(X|\mathcal{S}).$$

■

La esperanza condicional a una sub- σ -álgebra cumple otras propiedades. Aquí sólo se presentaron aquellas propiedades relevantes para el desarrollo del trabajo. Una presentación más completa de estas propiedades puede verse en [Shiryaev, 1996], capítulo 2, sección 7.

Definición 2.2.9 (Varianza condicional). *Sea X una variable aleatoria tal que $\mathbb{E}|X| < \infty$ e Y una variable aleatoria arbitraria. Se define la **varianza condicional de X dado Y** como*

$$\text{var}(X|Y) = \mathbb{E}[(X - \mathbb{E}(X|Y))^2|Y]$$

Observación 2.2.3. Podemos escribir la varianza condicional como

$$\text{var}(X|Y) = \mathbb{E}[(X^2 - 2X\mathbb{E}(X|Y) + \mathbb{E}(X|Y)^2)|Y]$$

Utilizando la proposición (2.2.3) se tiene que

$$\text{var}(X|Y) = \mathbb{E}[X^2|Y] - 2\mathbb{E}[X\mathbb{E}(X|Y)|Y] + \mathbb{E}[\mathbb{E}(X|Y)^2|Y]$$

Por definición de esperanza condicional, $\mathbb{E}(X|Y)$ es $\sigma(Y)$ -medible. Luego, por la proposición (2.2.6) se concluye que

$$\text{var}(X|Y) = \mathbb{E}(X^2|Y) - \mathbb{E}(X|Y)^2$$

Proposición 2.2.7 (Ley de varianza total). Dadas X una variable aleatoria tal que $\text{var}(X) < \infty$ e Y una variable aleatoria se cumple que

$$\text{var}(X) = \mathbb{E}(\text{var}(X|Y)) + \text{var}(\mathbb{E}(X|Y))$$

Demostración. Por definición de varianza se tiene que

$$\text{var}(X) = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2$$

Utilizando el corolario (2.2.1) podemos reformular esta ecuación como

$$\text{var}(X) = \mathbb{E}(\mathbb{E}(X^2|Y)) - [\mathbb{E}(\mathbb{E}(X|Y))]^2.$$

Por la observación (2.2.3) se tiene que

$$\text{var}(X) = \mathbb{E}(\text{var}(X|Y) + [\mathbb{E}(X|Y)]^2) - [\mathbb{E}(\mathbb{E}(X|Y))]^2.$$

Aplicando linealidad de la esperanza y reagrupando se concluye que

$$\text{var}(X) = \mathbb{E}(\text{var}(X|Y)) + \text{var}(\mathbb{E}(X|Y)).$$

■

2.2.2. Martingalas

Definición 2.2.10 (Proceso estocástico). Sea $(E, \mathcal{E})$ un espacio medible. Dado $(\Omega, \mathcal{A}, \mathbf{P})$ y un conjunto de índices I se define un **proceso estocástico** a una familia $X = \{X_i\}_{i \in I}$ de variables aleatorias definidas en $(\Omega, \mathcal{A}, \mathbf{P})$ que toman valores en E . Al conjunto $(E, \mathcal{E})$ le llamaremos **espacio de estados**.

Definición 2.2.11 (Filtración). Sea $(\Omega, \mathcal{A}, \mathbf{P})$ un espacio de probabilidad. Sea $\Sigma = \{t \geq 0\}$ con $t \in \mathbb{R}$ o $t \in \mathbb{Z}$. Se dice que $\mathcal{F} = \{\mathcal{F}_t\}_{t \in \Sigma}$ es una **filtración** si se cumple que

1. $\mathcal{F}_t$ es una sub- σ -álgebra de $\mathcal{A}$, para todo $t \in [0, +\infty)$, y
2. dados $t, s \in [0, +\infty)$ se cumple que $\mathcal{F}_t \subset \mathcal{F}_{t+s}$.

Un espacio $(\Omega, \mathcal{A}, \mathcal{F}, \mathbf{P})$ diremos que es un **espacio filtrado**.

Definición 2.2.12 (Filtración generada por un proceso estocástico). Sea $X = \{X_t\}_{t \in \Sigma}$ un proceso estocástico sobre un espacio de estados $(E, \mathcal{E})$. Definimos a la **filtración generada por X** o **filtración natural de X** y la notamos como $\mathcal{F}^X = \{\mathcal{F}_t\}_{t \in \Sigma}$ a la filtración tal que

$$\mathcal{F}_t = \sigma\{X_s^{-1}(B) : s \leq t, B \in \mathcal{E}\}, \quad (2.5)$$

Definición 2.2.13. Sea $\mathcal{F} = \{\mathcal{F}_t\}_{t \in \Sigma}$ una filtración. Sea $X = \{X_t\}_{t \in \Sigma}$ un proceso estocástico. Decimos que X está **adaptado** a $\mathcal{F}$ si para todo $t \in [0, +\infty)$, la variable X_t es $\mathcal{F}_t$ -medible.

Observación 2.2.4. Todo proceso estocástico está adaptado a su filtración natural.

Observación 2.2.5. Intuitivamente $\mathcal{F}_t$ representa nuestro conocimiento hasta el tiempo t . $\mathcal{F}_t$ contiene todos los eventos A tales que al tiempo t es posible decidir si ocurrieron o no. Como la familia $\{\mathcal{F}_t\}_{t \in \Sigma}$ es una familia creciente de conjuntos, nuestra información sobre la dinámica del proceso aumentará con el tiempo.

Definición 2.2.14 (Martingala, submartingala y supermartingala). Sea $\Sigma = \{t \geq 0\}$. Un proceso estocástico $X = \{X_t\}_{t \in \Sigma}$ es una **martingala** con respecto a una filtración $\mathcal{F} = \{\mathcal{F}_t\}_{t \in \Sigma}$ si

- 1) $\mathbb{E}|X_t| < \infty$ para todo $t \in \Sigma$.
- 2) X está adaptado a $\mathcal{F}$
- 3) $\mathbb{E}(X_t | \mathcal{F}_s) = X_s$ para todo $s, t \in \Sigma$ tales que $s < t$.

Si $\Sigma = \{t \geq 0 : t \in \mathbb{R}\}$ diremos que X es una **martingala en tiempo continuo**. Si $\Sigma = \{t \geq 0 : t \in \mathbb{Z}\}$ diremos que X es una **martingala en tiempo discreto**.

Si en lugar de la igualdad del punto (3) para todo $s, t \in \Sigma$ tales que $s < t$ se cumple que

$$\mathbb{E}(X_t | \mathcal{F}_s) \geq X_s$$

diremos que X es una **submartingala** mientras que si se cumple que

$$\mathbb{E}(X_t | \mathcal{F}_s) \leq X_s$$

diremos que X es una **supermartingala**.

Proposición 2.2.8. Sea $X = \{X_n\}_{n \in \mathbb{N}}$ una martingala con respecto a una filtración $\mathcal{F} = \{\mathcal{F}_n\}_{n \in \mathbb{N}}$. Entonces

$$\mathbb{E}(X_n) = \mathbb{E}(X_0), \quad \forall n \in \mathbb{N}$$

Demostración. Como X_n es una martingala se tiene que

$$\mathbb{E}(\mathbb{E}(X_{n+1} | \mathcal{F}_n)) = \mathbb{E}(X_n)$$

Por otra parte, como X_{n+1} es $\mathcal{F}_{n+1}$ -medible y $\mathcal{F}_n \subset \mathcal{F}_{n+1}$ se tiene, por la proposición (2.2.2) que

$$\mathbb{E}(X_{n+1}) = \mathbb{E}(\mathbb{E}(X_{n+1} | \mathcal{F}_n))$$

Luego, para todo $n \in \mathbb{N}$ se concluye que

$$\mathbb{E}(X_{n+1}) = \mathbb{E}(X_n)$$

como queríamos probar. ■

2.3. Procesos de Markov

En la siguiente sección definiremos el concepto de proceso de Markov intentando ilustrar cómo éste puede ser pensado como una extensión del concepto de *cadena de Markov* al tiempo continuo y a un espacio de estados arbitrario. Una vez presentado este concepto estudiaremos la vinculación entre la estructura probabilística de los procesos de Markov y el enfoque analítico de la teoría de semigrupos.

Definición 2.3.1 (Proceso markoviano). Sea $X = \{X_t\}_{t \geq 0}$ un proceso estocástico definido en un espacio de probabilidad $(\Omega, \mathcal{A}, \mathbf{P})$ y que toma valores en un espacio métrico (E, d) . Sea $\mathcal{F}^X = \{\mathcal{F}_t^X\}_{t \geq 0}$ la filtración natural de X . Se dice que X es un proceso **de Markov** o **markoviano** si

$$\mathbf{P}(X_{t+s} \in B | \mathcal{F}_t^X) = \mathbf{P}(X_{t+s} \in B | X_t),$$

para todos $s, t \geq 0$ y $B \in \mathcal{B}(E)$.

Un proceso de Markov es **homogéneo en el tiempo** si

$$\mathbf{P}(X_{t+h} \in B | X_t) = \mathbf{P}(X_t \in B | X_0)$$

para todo $B \in \mathcal{B}(E)$ y todo $t, h \geq 0$.

Observación 2.3.1. La definición de probabilidad condicionada implica calcular una probabilidad condicionada a un evento. En la definición de proceso de Markov escribimos $\mathbf{P}(X_{t+s} \in B | \mathcal{F}_t^X)$ donde $\mathcal{F}_t^X$ es una sub- σ -álgebra. Cuando escribimos $\mathbf{P}(X_{t+s} \in B | \mathcal{F}_t^X)$ formalmente estamos escribiendo

$$\mathbb{E}(\mathbb{1}_B(X_{t+s}) | \mathcal{F}_t^X).$$

Al igual que muchas referencias sobre procesos estocásticos, haremos este abuso de notación durante el resto del trabajo.

Observación 2.3.2. Equivalentemente, se puede definir un proceso $X = \{X_t\}_{t \geq 0}$ como markoviano si cumple que

$$\mathbb{E}(f(X_{t+s}) | \mathcal{F}_t^X) = \mathbb{E}(f(X_{t+s}) | X_t),$$

para todos $s, t \geq 0$ y $f : E \rightarrow \mathbb{R}$ $\mathcal{B}(E)$ -medible y acotada. Una prueba de que estas nociones son equivalentes puede verse en [Çinlar, 2011].

Recordemos que una cadena de Markov está caracterizada por su distribución inicial y su matriz de transición. La siguiente definición pretende extender esta noción para el caso en que X es un proceso estocástico markoviano.

Definición 2.3.2 (Función de transición). Sea E un espacio métrico y $\mathcal{B}(E)$ su σ -álgebra de Borel. Una función $\mathbf{P} : [0, +\infty) \times E \times \mathcal{B}(E) \rightarrow \mathbb{R}$ es una **función de transición homogénea en el tiempo** si

1. fijados $x \in E$ y $t \in [0, +\infty)$, $\mathbf{P}(t, x, \cdot)$ es una medida de probabilidad,

2. para todo $x \in E$, $\mathbf{P}(0, x, \cdot) = \delta_x$, con δ_x delta de Dirac en x ,
3. fijado $B \in \mathcal{B}(E)$, $\mathbf{P}(\cdot, \cdot, B)$ es medible en $\mathcal{B}([0, +\infty) \times E)$, y
4. para todos $s, t \geq 0$, $x \in E$, y $B \in \mathcal{B}(E)$ se cumple la llamada propiedad de Chapman-Kolmogorov

$$\mathbf{P}(t + s, x, B) = \int_E \mathbf{P}(s, y, B) \mathbf{P}(t, x, dy).$$

Definición 2.3.3 (Función de transición de un proceso markoviano). *Una función de transición $\mathbf{P}(t, x, B)$ es una **función de transición de un proceso estocástico markoviano homogéneo en el tiempo** X si*

$$\mathbf{P}(X_{t+s} \in B | \mathcal{F}_t^X) = \mathbf{P}(X_{t+s} \in B | X_t) = \mathbf{P}(s, X_t, B)$$

o, de forma equivalente, si

$$\mathbb{E}(f(X_{t+s}) | \mathcal{F}_t^X) = \int_E f(y) \mathbf{P}(s, X_t, dy)$$

Definición 2.3.4 (Distribución inicial). *Dado un proceso estocástico markoviano $X = \{X_t\}_{t \geq 0}$, definimos la **distribución inicial de X en $(E, \mathcal{B}(E))$** a la medida de probabilidad ν dada por*

$$\nu(B) = \mathbf{P}(X_0 \in B)$$

Observación 2.3.3. Al igual que obteníamos para cadenas de Markov en tiempo discreto y espacio de estados E la probabilidad de transición

$$\mathbf{P}(X_{n+1} = j | X_n = i) = p_{ij},$$

obtenemos ahora una función de transición $\mathbf{P}(s, X_t, B)$ que nos indica la probabilidad de que el proceso alcance en el tiempo $t + s$ el conjunto B sabiendo dónde se encontraba el proceso en tiempo t .

Observación 2.3.4. El ítem (4) de la definición (2.3.2) es la extensión de las ecuación de Chapman-Kolmogorov (5.1.2) para cadenas de Markov en tiempo discreto y espacio de estados finito para procesos estocásticos markovianos en tiempo continuo y espacio de estados E . Veamos que esta propiedad es razonable dado un proceso markoviano.

Sean $X = \{X_t\}_{t \geq 0}$ un proceso estocástico markoviano sobre un espacio de estados E y $\mathbf{P}(s, X_t, B)$ una función de transición. Luego

$$\begin{aligned} \mathbf{P}(t + s, X_u, B) &= \mathbf{P}(X_{u+t+s} \in B | \mathcal{F}_u^X) \\ &= \mathbb{E}(\mathbf{1}_B(X_{u+t+s}) | \mathcal{F}_u^X) \\ &= \mathbb{E}(\mathbb{E}(\mathbf{1}_B(X_{u+t+s}) | \mathcal{F}_{u+t}^X) | \mathcal{F}_u^X) && \text{(propiedad 2.2.4)} \\ &= \mathbb{E}(\mathbf{P}(X_{u+t+s} \in B | \mathcal{F}_{u+t}^X) | \mathcal{F}_u^X) \end{aligned}$$

$$\begin{aligned}
&= \mathbb{E}(\mathbf{P}(s, X_{u+t}, B) | \mathcal{F}_u^X) && \text{(definición de función de transición)} \\
&= \mathbb{E}(\mathbf{P}(s, X_{u+t}, B) | X_u) && \text{(definición de proceso markoviano)} \\
&= \int_E \mathbf{P}(s, y, B) \mathbf{P}(t, X_u, dy)
\end{aligned}$$

No es sencillo obtener fórmulas explícitas para las funciones de transición de procesos markovianos arbitrarios. Una excepción notable de esto es la función de transición de un movimiento browniano unidimensional (que definiremos en el capítulo 3):

$$\mathbf{P}(X_{t+s} \in B | X_t) = \mathbf{P}(s, X_t, B) = \int_B \frac{1}{\sqrt{2\pi s}} e^{-\frac{(y-x)^2}{2s}} dy$$

Por esta razón, no podremos en la gran mayoría de los casos, definir un proceso markoviano a través de su función de transición. Debemos utilizar otras técnicas de especificarlos. En particular, utilizaremos que la propiedad de Chapman-Kolmogorov nos define un semigrupo contractivo:

Proposición 2.3.1. Sean (E, d) un espacio métrico y $\mathfrak{B}$ el espacio de funciones

$$\mathfrak{B} = \{f : E \rightarrow \mathbb{R} : \mathcal{B}(E)\text{-medible y acotada}\}$$

equipado con la norma del supremo. Sea $\mathbf{P}(t, x, B)$ una función de transición definida en $[0, +\infty) \times E \times \mathcal{B}(E)$. Definimos $\mathcal{S} = \{S_t\}_{t \geq 0}$ tal que para todo $t \geq 0$, $S_t : \mathfrak{B} \rightarrow \mathfrak{B}$ y

$$(S_t f)(x) = \int_E f(y) \mathbf{P}(t, x, dy)$$

Luego, $\mathcal{S}$ es un semigrupo contractivo en $\mathfrak{B}$.

Demostración. En primer lugar, probemos que para todo $t \geq 0$, S_t es un operador lineal acotado:

$$\begin{aligned}
\|S_t(f)\| &= \sup_{x \in E} \left| \int_E f(y) \mathbf{P}(t, x, dy) \right| \\
&\leq \sup_{x \in E} \left| \int_E \left(\sup_{y \in E} |f(y)| \right) \mathbf{P}(t, x, dy) \right| \\
&= \|f\| \sup_{x \in E} \underbrace{\left| \int_E \mathbf{P}(t, x, dy) \right|}_{=1} = \|f\|
\end{aligned}$$

Observemos que para todo $t \geq 0$ se tiene entonces que

$$\|S_t(f)\| \leq \|f\|,$$

lo que implica que $\|S_t\| = 1$. En caso de que $\mathcal{S}$ sea un semigrupo, este hecho muestra que $\mathcal{S}$ es un semigrupo contractivo.

Probemos que, efectivamente, $\mathcal{S}$ es un semigrupo. Probemos que S_0 es el operador identidad en $\mathfrak{B}$. Sea $f \in \mathfrak{B}$. Por definición de $\mathcal{S}$ se tiene que

$$S_0(f) = \int_E f(y) \mathbf{P}(0, x, dy)$$

Como $P(t, x, B)$ es una función de transición (definición 2.3.2 (3)) se tiene que

$$\int_E f(y) \mathbf{P}(0, x, dy) = \int_E f(y) \delta_x = f(x)$$

Para concluir, probemos que $S_{t_1+t_2}(f) = (S_{t_1} \circ S_{t_2})(f)$. Se tiene que

$$\begin{aligned} S_{t_1+t_2}(f) &= \int_E f(y) \mathbf{P}(t_1 + t_2, x, dy) \\ &= \int_E f(y) \left(\int_E \mathbf{P}(t_2, z, dy) \mathbf{P}(t_1, x, dz) \right) \\ &= \int_E \underbrace{\left(\int_E f(y) \mathbf{P}(t_2, z, dy) \right)}_{S_{t_2}(f)} \mathbf{P}(t_1, x, dz) \\ &= S_{t_1}(S_{t_2}(f)) = (S_{t_1} \circ S_{t_2})(f) \end{aligned}$$

Lo que concluye que $\mathcal{S}$ es un semigrupo de operadores en $\mathfrak{B}$. ■

Dado un proceso de Markov $X = \{X_t\}_{t \geq 0}$ homogéneo en el tiempo y que toma valores en un espacio métrico E y una función de transición $\mathbf{P}(s, X_t, B)$ definimos el semigrupo asociado a X como

$$(S_t f)(x) = \mathbb{E}(f(X_t) | X_0 = x) = \int_E f(y) \mathbf{P}(t, x, dy) \quad (2.6)$$

Notaremos $\mathbb{E}_x(\cdot) := \mathbb{E}(\cdot | X_0 = x)$. De esta forma, el semigrupo asociado a X lo notaremos como

$$(S_t f)(x) = \mathbb{E}_x(f(X_t)) \quad (2.7)$$

Capítulo 3

Ecuaciones diferenciales estocásticas

3.1. Proceso de Wiener: Definición y propiedades básicas

De ahora en más los procesos estocásticos considerados tomarán valores en $E = \mathbb{R}$. Comenzaremos estudiando el proceso de Wiener, proceso estocástico fundamental para definir ecuaciones diferenciales estocásticas.

Definición 3.1.1 (Proceso de Wiener). *Un proceso estocástico $W = \{W_t\}_{t \geq 0}$ con valores en $\mathbb{R}$ es un **proceso de Wiener unidimensional** o un **movimiento browniano unidimensional** si*

1. $W_0 = 0$ casi seguramente.
2. $\mathbf{P}(\omega \in \Omega : W_t(\omega) \in C[0, +\infty)) = 1$.
3. El incremento $W_{t+h} - W_t$ es independiente de $\sigma(W_t)$ para todo $t, h \geq 0$.
4. $W_{t+h} - W_t \sim N(0, h)$ para todo $t, h \geq 0$.

Observación 3.1.1. Un proceso de Wiener unidimensional $W = \{W_t\}_{t \geq 0}$ es un proceso markoviano.

Sea $f : \mathbb{R} \rightarrow \mathbb{R}$ Borel-medible y acotada y $\mathcal{F} = \{\mathcal{F}_t\}_{t \geq 0}$ la filtración generada por W . Luego

$$\begin{aligned}\mathbb{E}(f(W_{t+s})|\mathcal{F}_t) &= \mathbb{E}(f(W_{t+s} - W_t + W_t)|\mathcal{F}_t) \\ &= \mathbb{E}(f(\sqrt{s}Z + W_t)|\mathcal{F}_t) \\ &= \int_{\mathbb{R}} \mathbb{E}(f(z + W_t)|\mathcal{F}_t) \frac{1}{\sqrt{t}} f_Z\left(\frac{z}{\sqrt{t}}\right) dz,\end{aligned}$$

donde f_Z es la función de densidad de una variable aleatoria normal estándar. Como W_s es medible en $\mathcal{F}_s$, por la proposición (2.2.2)

$$\mathbb{E}(f(z + W_s)|\mathcal{F}_s) = f(z + W_s).$$

Luego

$$\begin{aligned}\mathbb{E}(f(W_{t+s})|\mathcal{F}_t) &= \int_{\mathbb{R}} f(z + W_s) \frac{1}{\sqrt{t}} f_Z\left(\frac{z}{\sqrt{t}}\right) dz \\ &= \mathbb{E}(f(W_{t+s})|W_t)\end{aligned}$$

A continuación presentaremos una propiedad fundamental del proceso de Wiener que será de suma utilidad en el desarrollo de la teoría de las ecuaciones diferenciales estocásticas:

Proposición 3.1.1 (Variación cuadrática del proceso de Wiener). Sea $W = \{W_t\}_{t \geq 0}$ un proceso de Wiener. Para cada n y $T \geq 0$ se considera una partición del intervalo $[0, T]$, π_n , de la forma

$$\pi_n = \{0 = t_{n,0} < t_{n,1} < \dots < t_{n,n} = T\}.$$

Supondremos, además, que

$$\lim_{n \rightarrow \infty} \max_{0 \leq i \leq n-1} |t_{n,i+1} - t_{n,i}| = 0$$

Definimos

$$Q_n = \sum_{i=1}^n (W_{t_{n,i+1}} - W_{t_{n,i}})^2.$$

Entonces

$$Q_n \xrightarrow{\mathcal{L}^2} T, \quad \text{esto es,} \quad \mathbb{E}[(Q_n - T)^2] \xrightarrow{n \rightarrow \infty} 0$$

Demostración. Queremos probar que para todo $T > 0$

$$\mathbb{E}[(Q_n - T)^2] \xrightarrow{n \rightarrow \infty} 0.$$

Dado $n \in \mathbb{N}$, la idea será descomponer $Q_n - T$ de la siguiente forma:

$$Q_n - T = \left(\sum_{i=0}^{n-1} (W_{t_{n,i+1}} - W_{t_{n,i}})^2 \right) - T = \sum_{i=0}^{n-1} (W_{t_{n,i+1}} - W_{t_{n,i}})^2 - (t_{n,i+1} - t_{n,i})$$

y denominar $Y_i = (W_{t_{n,i+1}} - W_{t_{n,i}})^2 - (t_{n,i+1} - t_{n,i})$, para todo $i = 0, \dots, n-1$. Luego

$$\mathbb{E}[(Q_n - T)^2] = \mathbb{E} \left(\sum_{i=0}^{n-1} Y_i \right)^2 = \left(\mathbb{E} \left(\sum_{i=0}^{n-1} Y_i \right) \right)^2 + \text{var} \left(\sum_{i=0}^{n-1} Y_i \right)$$

Calculemos cada uno de estos sumandos

$$\mathbb{E} \left(\sum_{i=0}^{n-1} Y_i \right) = \sum_{i=0}^{n-1} \mathbb{E}(Y_i)$$

$$\begin{aligned}
&= \sum_{i=0}^{n-1} \underbrace{\mathbb{E}(W_{t_{n,i+1}} - W_{t_{n,i}})^2}_{t_i - t_{i-1}} - (t_{n,i+1} - t_{n,i}) \\
&= 0
\end{aligned}$$

Antes de calcular el término $\text{var}(\sum_{i=0}^{n-1} Y_i)$ notemos que las variables Y_i son independientes por definición de proceso de Wiener. Se tiene entonces que

$$\begin{aligned}
\text{var}\left(\sum_{i=0}^{n-1} Y_i\right) &= \sum_{i=0}^{n-1} \text{var}(Y_i) \\
&= \sum_{i=0}^{n-1} \text{var}[(W_{t_{n,i+1}} - W_{t_{n,i}})^2 - (t_{n,i+1} - t_{n,i})] \\
&= \sum_{i=0}^{n-1} \text{var}[(W_{t_{n,i+1}} - W_{t_{n,i}})^2] \\
&= \sum_{i=0}^{n-1} \text{var}[(\sqrt{t_{n,i+1} - t_{n,i}} Z)^2]
\end{aligned}$$

donde Z es una variable aleatoria normal estándar. Luego

$$\begin{aligned}
\text{var}\left(\sum_{i=0}^{n-1} Y_i\right) &= \text{var}(Z^2) \sum_{i=0}^{n-1} (t_{n,i+1} - t_{n,i})^2 \\
&\leq \text{var}(Z^2) \left[\max_{0 \leq i \leq n-1} (t_{n,i+1} - t_{n,i}) \right] \sum_{i=0}^{n-1} t_{n,i+1} - t_{n,i} \\
&= \text{var}(Z^2) \left[\max_{0 \leq i \leq n-1} (t_{n,i+1} - t_{n,i}) \right] T
\end{aligned}$$

Como $Z^2 \sim \chi_1^2$, $\text{var}(Z^2) < \infty$. Luego, por la condición (3.1.1):

$$\lim_{n \rightarrow \infty} \mathbb{E}[(Q_n - T)^2] \leq \lim_{n \rightarrow \infty} \text{var}(Z^2) \left[\max_{0 \leq i \leq n-1} (t_{n,i+1} - t_{n,i}) \right] T = 0$$

■

3.2. Integral de Itô

3.2.1. Construcción de la integral

Sea $(\Omega, \mathcal{A}, \{\mathcal{F}_t\}, \mathbf{P})$ un espacio de probabilidad filtrado. Sea $\{W_t\}_{t \geq 0}$ un proceso de Wiener $\mathcal{F}_t$ -medible —podemos tomar, por ejemplo, $\mathcal{F}$ como la filtración natural dada por W . Estamos interesados en definir la integral estocástica

$$\int_0^t X(s) dW(s)$$

Si bien las trayectorias de W son continuas casi seguramente, no podemos definir esta integral en el sentido de Riemann-Stieljes porque las trayectorias del proceso de Wiener no tienen variación acotada (ver [Petrov y Mordecki, 2008]). La construcción de la integral estocástica será, en cierta medida, similar al de la integral de Lebesgue pero utilizará características propias del proceso de Wiener.

Definición 3.2.1 (Proceso estocástico progresivamente medible). Sea $X = \{X_t\}_{t \geq 0}$, un proceso estocástico definido en un espacio filtrado $(\Omega, \mathcal{A}, \mathcal{F}_t, \mathbf{P})$ y que toma valores en $\mathbb{R}$. Se dice que X es un proceso estocástico **progresivamente medible** si para todo $(s, \omega) \in [0, t] \times \Omega$, $X_t(\omega)$ es medible en $\mathcal{B}[0, t] \times \mathcal{F}_t$.

Sea $T > 0$. Consideramos la clase $\mathcal{H}_2[0, T]$ de procesos estocásticos $\{\{H_t\} : t \in [0, T]\}$ progresivamente medibles y que satisfacen la condición

$$\int_0^T \mathbb{E}(H_s^2) ds < \infty$$

Sea $0 = t_0 < t_1 < \dots < t_n = T$ una partición del intervalo $[0, T]$. Definimos un *proceso simple* a un proceso estocástico de la forma $H = \{H_t\}_{0 \leq t \leq T}$ tal que

$$H_t = \sum_{i=1}^{n-1} h_i \mathbb{1}_{[t_i, t_{i+1})}(t),$$

donde, para todo $i = 0, \dots, n-1$, h_i es una variable aleatoria $\mathcal{F}_{t_i}$ -medible y tal que $\mathbb{E}(h_i^2) < \infty$.

Observemos que todo proceso simple pertenece a $\mathcal{H}_2[0, T]$. Para todo $t \in [0, T]$ tenemos que

$$\begin{aligned} \mathbb{E}(H_t^2) &= \mathbb{E} \left[\left(\sum_{i=1}^{n-1} h_i \mathbb{1}_{[t_i, t_{i+1})}(t) \right)^2 \right] \\ &= \mathbb{E} \left(\sum_{i=1}^{n-1} h_i^2 \mathbb{1}_{[t_i, t_{i+1})}(t) + 2 \sum_{i < j} h_i h_j \mathbb{1}_{[t_i, t_{i+1})}(t) \mathbb{1}_{[t_j, t_{j+1})}(t) \right) \\ &= \mathbb{E}(h_{t_k}^2) < \infty \end{aligned}$$

donde la última igualdad se debe a que los intervalos $[t_i, t_{i+1})$ son disjuntos dos a dos y que para todo $t \in [0, T]$ existe un $t_k \in [0, T]$ tal que $t \in [t_k, t_{k+1})$.

Definimos la integral de un proceso simple H con respecto a un proceso de Wiener W como

$$\int_0^t H_t dW_t := \sum_{i=0}^{n-1} h_i (W_{t_{i+1}} - W_{t_i}) \quad (3.1)$$

La siguiente proposición nos mostrará que la integral de un proceso simple con respecto a un proceso de Wiener cumple una propiedad deseable: la integral será un operador lineal en $\mathcal{H}^2[0, T]$.

Proposición 3.2.1. Sean $H^{(1)}$ y $H^{(2)}$ procesos simples en $\mathcal{H}_2[0, T]$. Para todos $\alpha, \beta \in \mathbb{R}$ se cumple que

- 1) $\alpha H_t^{(1)} + \beta H_t^{(2)}$ es un proceso simple en $\mathcal{H}^2[0, T]$ y,
- 2) $\int_0^T (\alpha H_t^{(1)} + \beta H_t^{(2)}) dW_t = \int_0^T \alpha H_t^{(1)} dW_t + \int_0^T \beta H_t^{(2)} dW_t$

Demostración. 1) Sean $H^{(1)}$ y $H^{(2)}$ procesos simples. Por definición de proceso simple, existen particiones $\pi^{(1)} = \{A_i\}_{i=0, \dots, n-1}$ y $\pi^{(2)} = \{B_j\}_{j=0, \dots, m-1}$ del intervalo $[0, T]$ tales que

$$H_t^{(1)} = \sum_{i=0}^{n-1} h_i^{(1)} \mathbf{1}_{[t_i, t_{i+1})}(t), \quad H_t^{(2)} = \sum_{j=0}^{m-1} h_j^{(2)} \mathbf{1}_{[s_j, s_{j+1})}(t)$$

Notemos que, como $[t_i, t_{i+1}) \subset [0, T]$ y $[0, T] = \bigcup_{j=0}^{m-1} [s_j, s_{j+1})$, $[t_i, t_{i+1})$ puede ser escrito como

$$[t_i, t_{i+1}) = [t_i, t_{i+1}) \cap \left(\bigcup_{j=0}^{m-1} B_j \right)$$

para todo $i = 0, \dots, n-1$. De la misma forma, podemos escribir $[s_j, s_{j+1})$ como

$$[s_j, s_{j+1}) = [s_j, s_{j+1}) \cap \left(\bigcup_{i=0}^{n-1} [t_i, t_{i+1}) \right).$$

para todo $j = 0, \dots, m-1$. Como la intersección de intervalos semiabierto es a su vez un intervalo semiabierto se tiene que los conjuntos $[t_i, t_{i+1}) \cap [s_j, s_{j+1})$ son intervalos semiabierto. Notaremos

$$[u_k, u_{k+1}) := [t_i, t_{i+1}) \cap [s_j, s_{j+1})$$

con $k \in \{1, \dots, n+m-1\}$.

Definimos

$$g_k^{(1)} = h_i^{(1)} \quad \text{si } t_i \leq u_k < t_{i+1}$$

y

$$g_k^{(2)} = h_j^{(2)} \quad \text{si } s_j \leq u_k < s_{j+1}$$

Notemos que

$$H_t^{(1)} = \sum_{i=0}^{n-1} h_i^{(1)} \mathbf{1}_{[t_i, t_{i+1})}(t) = \sum_k g_k^{(1)} \mathbf{1}_{[u_k, u_{k+1})}(t)$$

y

$$H_t^{(2)} = \sum_{j=0}^{m-1} h_j^{(2)} \mathbf{1}_{[s_j, s_{j+1})}(t) = \sum_k g_k^{(2)} \mathbf{1}_{[u_k, u_{k+1})}(t).$$

Se concluye que

$$\alpha H_t^{(1)} + \beta H_t^{(2)} = \sum_k (\alpha g_k^{(1)} + \beta g_k^{(2)}) \mathbb{1}_{[u_k, u_{k+1})}(t)$$

es un proceso simple.

2) Por definición de integral estocástica para procesos simples se tiene que

$$\begin{aligned} \int_0^T (\alpha H_t^{(1)} + \beta H_t^{(2)}) dW_t &= \sum_k (\alpha g_k^{(1)} + \beta g_k^{(2)}) (W_{u_{k+1}} - W_{u_k}) \\ &= \alpha \sum_k g_k^{(1)} (W_{u_{k+1}} - W_{u_k}) + \beta \sum_k g_k^{(2)} (W_{u_{k+1}} - W_{u_k}) \\ &= \alpha \sum_i^{n-1} h_i^{(1)} (W_{t_{i+1}} - W_{t_i}) + \beta \sum_j^{m-1} h_j^{(2)} (W_{s_{j+1}} - W_{s_j}) \\ &= \alpha \int_0^T H_t^{(1)} dW_t + \beta \int_0^T H_t^{(2)} dW_t \end{aligned}$$

■

Proposición 3.2.2. Para todo proceso simple $\{H_t\}_{t \in [0, T]} \in \mathcal{H}^2[0, T]$ se cumple que

1. $\mathbb{E} \left(\int_0^T H_t dW_t \right) = 0$
2. $\mathbb{E} \left[\left(\int_0^T H_t dW_t \right)^2 \right] = \int_0^T \mathbb{E}(H_t^2) dt$

Demostración. 1. Sea $\{H_t\}_{t \in [0, T]} \in \mathcal{H}^2[0, T]$ un proceso simple. Luego

$$\begin{aligned} \mathbb{E} \left(\int_0^T H_t dW_t \right) &= \mathbb{E} \left[\sum_{i=0}^{n-1} h_i (W_{t_{i+1}} - W_{t_i}) \right] \\ &= \sum_{i=0}^{n-1} \mathbb{E} [h_i (W_{t_{i+1}} - W_{t_i})] \\ &= \sum_{i=0}^{n-1} \mathbb{E}(h_i) \underbrace{\mathbb{E}(W_{t_{i+1}} - W_{t_i})}_0 = 0, \end{aligned}$$

donde en la penúltima igualdad se utilizó que h_i es $\mathcal{F}_{t_i}$ -medible y $W_{t_{i+1}} - W_{t_i}$ es independiente de $\mathcal{F}_{t_i}$.

2. Calculemos, en una primera instancia, el cuadrado:

$$\left(\int_0^T H_t dW_t \right)^2 = \left(\sum_{i=0}^{n-1} h_i (W_{t_{i+1}} - W_{t_i}) \right)^2$$

$$\begin{aligned}
&= \sum_{i=0}^{n-1} h_i^2 (W_{t_{i+1}} - W_{t_i})^2 \\
&\quad + 2 \sum_{0 \leq i < j \leq n-1} h_i h_j (W_{t_{i+1}} - W_{t_i})(W_{t_{j+1}} - W_{t_j})
\end{aligned}$$

Tomando esperanza se tiene que

$$\begin{aligned}
\mathbb{E} \left[\left(\int_0^T H_t dW_t \right)^2 \right] &= \sum_{i=1}^{n-1} \mathbb{E}(h_i^2 (W_{t_{i+1}} - W_{t_i})^2) \\
&\quad + 2 \sum_{1 \leq i < j \leq n} \mathbb{E}(h_i h_j (W_{t_{i+1}} - W_{t_i})(W_{t_{j+1}} - W_{t_j}))
\end{aligned}$$

Por independencia se tiene que, para todo $i = 1, \dots, n-1$

$$\mathbb{E}(h_i^2 (W_{t_{i+1}} - W_{t_i})^2) = \mathbb{E}(h_i^2) \underbrace{\mathbb{E}[W_{t_{i+1}} - W_{t_i}]^2}_{t_{i+1} - t_i} = \int_{t_i}^{t_{i+1}} \mathbb{E}(h_i^2) ds$$

Por otra parte, por independencia:

$$\mathbb{E}(h_i h_j (W_{t_{i+1}} - W_{t_i})(W_{t_{j+1}} - W_{t_j})) = \mathbb{E}[h_i h_j (W_{t_{i+1}} - W_{t_i})] \underbrace{\mathbb{E}[W_{t_{j+1}} - W_{t_j}]}_0 = 0$$

Se concluye que

$$\mathbb{E} \left[\left(\int_0^T H_t dW_t \right)^2 \right] = \sum_{i=1}^{n-1} \int_{t_i}^{t_{i+1}} \mathbb{E}(h_i^2) ds = \int_0^T \mathbb{E}(H_t^2) dt$$

■

Sea $L^2(\mathbf{P})$ el espacio de variables aleatorias tales que

$$\mathbb{E}(X^2) = \int_{\mathbb{R}} X^2 d\mathbf{P} < \infty$$

Podemos considerar a $L^2(\mathbf{P})$ como un espacio de Banach con la norma $\|\cdot\| : L^2(\mathbf{P}) \rightarrow \mathbb{R}$ dada por

$$\|X\| = \sqrt{\mathbb{E}(X^2)},$$

para toda $X \in L^2(\mathbf{P})$. Por otra parte, consideramos la norma $\|\cdot\| : \mathcal{H}_2[0, T] \rightarrow \mathbb{R}$ tal que

$$\|H\| = \sqrt{\int_0^T \mathbb{E}(H_t^2) dt}$$

Sea $\overline{\mathcal{H}}_2[0, T]$ el subconjunto de procesos simples de $\mathcal{H}_2[0, T]$. Se considera la función $f : \overline{\mathcal{H}}_2[0, T] \rightarrow L^2(\mathbf{P})$ que a cada proceso H le asigna su integral con respecto a W ; esto es:

$$f(H) = \int_0^T H_t dW_t \quad (3.2)$$

Por la proposición (3.2.2) se tiene que f es una isometría entre un subconjunto de $\mathcal{H}_2[0, T]$ y $L^2(\mathbf{P})$.

Proposición 3.2.3. El subconjunto de procesos simples es un subconjunto denso en el espacio $\mathcal{H}_2[0, T]$. Esto es, para todo proceso $H \in \mathcal{H}_2[0, T]$ existe una sucesión de procesos simples $\{H_t^{(n)}\}_{n \in \mathbb{N}} \subset \mathcal{H}_2[0, T]$ tal que

$$\lim_{n \rightarrow \infty} \|H - H_n\| = 0, \quad \text{o, equivalentemente,} \quad \lim_{n \rightarrow \infty} \int_0^T \mathbb{E}(H_t^{(n)} - H_t)^2 dt = 0$$

Demostración. La prueba será dividida en partes.

- Todo proceso $H \in \mathcal{H}_2[0, T]$ no acotado puede ser aproximado por procesos acotados.

Para probar esto se considera $H = \{H_t\}_{t \in [0, T]} \in \mathcal{H}_2[0, T]$ no acotado. Definimos la sucesión de procesos simples $\{H_t^{(N)}\}_{N \in \mathbb{N}}$ tal que

$$H_t^{(N)} = H_t \mathbf{1}_{[-N, N]}(H_t).$$

- Todo proceso continuo y acotado puede ser aproximado por procesos simples.

Sea $H = \{H_t\}_{t \in [0, T]} \in \mathcal{H}_2[0, T]$ un proceso continuo y acotado. Definimos la sucesión de procesos simples $\{H_t^{(N)}\}_{N \in \mathbb{N}}$ tales que

$$H_t^{(N)} = H_{\lfloor Nt \rfloor / N}$$

Es claro que $\lim_{N \rightarrow \infty} \|H - H_N\| = 0$

Como H está acotado existe $K \in \mathbb{R}$ tal que

$$|H_t| \leq K,$$

para todo $t \in [0, T]$. Para todo $N \in \mathbb{N}$ se obtiene, por lo tanto,

$$|H_t^{(N)}| \leq |H_t| \leq K.$$

Luego, como

$$\lim_{N \rightarrow \infty} H_t^{(N)} = H_t,$$

para todo $t \in [0, T]$, se concluye, por el teorema de convergencia dominada que

$$\lim_{N \rightarrow \infty} \int_0^T \mathbb{E}(H_t^{(N)} - H_t)^2 dt = 0$$

- Todo proceso acotado puede ser aproximado por procesos continuos.

Probar esto probará la proposición ya que dado un proceso $H \in \mathcal{H}_2[0, T]$ acotado, podrá ser aproximado por procesos continuos que a su vez pueden ser aproximados por procesos simples.

Se considera, entonces, $H = \{H_t\}_{t \in [0, T]} \in \mathcal{H}_2[0, T]$ acotado. Definimos la sucesión de procesos $\{H_t^{(N)}\}_{N \in \mathbb{N}}$ tales que

$$H_t^{(N)} = N \int_{(t-1/N)^+}^t H_s ds$$

donde $(t - 1/N)^+ = \max\{0, t - 1/N\}$. Los procesos $H^{(N)}$ son continuos y además se cumple que

$$\|H^{(N)}\| \leq N(t - t + 1/N) \|H\| \leq K$$

Por lo tanto $H^{(N)}$ será uniformemente acotado y progresivamente medible. Definimos $\{G_s\}$ tal que

$$G_s = \int_0^s H_t dt$$

Por el teorema de diferenciación de Lebesgue, para todo $s \in [0, T]$ existe G' y se cumple que

$$H_t = G'_t = \lim_{N \rightarrow \infty} H_t^{(N)}$$

Aplicando el teorema de convergencia dominada se concluye que

$$\lim_{N \rightarrow \infty} \int_0^T \mathbb{E}(H_t^{(N)} - H_t)^2 dt = 0$$

■

En virtud de la proposición (3.2.3) podemos extender la isometría lineal definida en (3.2) a todo el espacio $\mathcal{H}_2[0, T]$ de forma única y, de esta forma, definir la integral para todo proceso $H \in \mathcal{H}_2[0, T]$.

Definición 3.2.2 (Integral estocástica o de Itô). Sea $H = \{H_t\}_{t \in [0, T]} \in \mathcal{H}_2[0, T]$ un proceso arbitrario y $\{H_t^{(N)}\}_{N \in \mathbb{N}} \subset \mathcal{H}_2[0, T]$ una sucesión de procesos simples tales que

$$\lim_{n \rightarrow \infty} \int_0^T \mathbb{E}(H_t^{(n)} - H_t)^2 dt = 0.$$

Se define la **integral estocástica** o **integral de Itô**, $\int_0^T H_t dW_t$ como

$$\int_0^T H_t dW_t = \lim_{N \rightarrow \infty} \int_0^T H_t^{(N)} dW_t$$

Las propiedades obtenidas en (3.2.1) y (3.2.2) pueden ser extendidas a la integral de Itô. Esto será resumido en la siguiente proposición

Proposición 3.2.4 (Propiedades de la integral de Itô). Sean $H^{(1)}, H^{(2)}, H \in \mathcal{H}_{(2)}[0, T]$. Se cumple entonces que

- 1) $\int_0^T (\alpha H_t^{(1)} + \beta H_t^{(2)}) dW_t = \alpha \int_0^T H_t^{(1)} dW_t + \beta \int_0^T H_t^{(2)} dW_t$
- 2) $\mathbb{E} \left(\int_0^T H_t dW_t \right) = 0$
- 3) $\mathbb{E} \left[\left(\int_0^T H_t dW_t \right)^2 \right] = \int_0^T \mathbb{E}(H_t^2) dt$

3.3. Ecuaciones diferenciales estocásticas

A continuación presentaremos un resultado fundamental del cálculo estocástico: la fórmula de Itô. Presentaremos esta fórmula en tres versiones: una versión débil para funciones tres veces diferenciables que será demostrada, la versión clásica para funciones de clase C^2 y una versión más general para ecuaciones diferenciales estocásticas.

Teorema 3.3.1 (Fórmula de Itô: versión débil). Sea $\{W_t\}_{t \geq 0}$ un proceso de Wiener adaptado a la filtración $\{\mathcal{F}_t\}_{t \geq 0}$. Sea $f : \mathbb{R} \rightarrow \mathbb{R}$ con derivadas f', f'' y f''' acotadas. Entonces:

$$f(W_t) - f(W_0) = \int_0^t f'(W_s) dW_s + \frac{1}{2} \int_0^t f''(W_s) ds$$

Demostración. Sea $\pi_n = \{0 = t_{n,0} < t_{n,1} < \dots < t_{n,n} = t\}$ una sucesión de particiones encajadas del intervalo $[0, t]$ tales que

$$\lim_{n \rightarrow \infty} \max_{0 \leq k \leq n-1} |t_{n,k+1} - t_{n,k}| = 0 \quad (3.3)$$

Para cada $n \in \mathbb{N}$, podemos escribir entonces la siguiente igualdad:

$$f(W_t) - f(W_0) = \sum_{i=0}^{n-1} [f(W_{t_{n,i+1}}) - f(W_{t_{n,i}})]$$

Consideremos el desarrollo de Taylor de orden 2 de f centrado en $x = W_{t_{n,i}}$ para cada $i = 0, \dots, n-1$:

$$\begin{aligned} f(W_{t_{n,i+1}}) - f(W_{t_{n,i}}) &= f'(W_{t_{n,i}})(W_{t_{n,i+1}} - W_{t_{n,i}}) + \frac{1}{2} f''(W_{t_{n,i}})(W_{t_{n,i+1}} - W_{t_{n,i}})^2 \\ &\quad + \frac{1}{3!} f'''(\tilde{t}_{n,i})(W_{t_{n,i+1}} - W_{t_{n,i}})^3, \end{aligned}$$

donde, por la expresión del resto de Lagrange, $\tilde{t}_{n,i} \in [t_{n,i}, t_{n,i+1}]$.

Sumando y restando el término $\frac{1}{2} f''(W_{t_{n,i}})(t_{n,i+1} - t_{n,i})$ se obtiene

$$f(W_{t_{n,i+1}}) - f(W_{t_{n,i}}) = f'(W_{t_{n,i}})(W_{t_{n,i+1}} - W_{t_{n,i}}) + \frac{1}{2} f''(W_{t_{n,i}})(t_{n,i+1} - t_{n,i})$$

$$\begin{aligned}
& + \frac{1}{3!} f'''(W_{t_{n,i}})(W_{t_{n,i+1}} - W_{t_{n,i}})^3 \\
& + \frac{1}{2} f''(W_{t_{n,i}})[(W_{t_{n,i+1}} - W_{t_{n,i}})^2 - (t_{n,i+1} - t_{n,i})]
\end{aligned}$$

Sumando ahora sobre i obtenemos

$$\begin{aligned}
f(W_t) - f(W_0) &= \underbrace{\sum_{i=0}^{n-1} f'(W_{t_{n,i}})(W_{t_{n,i+1}} - W_{t_{n,i}})}_{I_{n,1}} + \underbrace{\frac{1}{2} \sum_{i=0}^{n-1} f''(W_{t_{n,i}})(t_{n,i+1} - t_{n,i})}_{I_{n,2}} \\
&+ \underbrace{\frac{1}{3!} \sum_{i=0}^{n-1} f'''(W_{t_{n,i}})(W_{t_{n,i+1}} - W_{t_{n,i}})^3}_{I_{n,3}} \\
&+ \underbrace{\frac{1}{2} \sum_{i=0}^{n-1} f''(W_{t_{n,i}})[(W_{t_{n,i+1}} - W_{t_{n,i}})^2 - (t_{n,i+1} - t_{n,i})]}_{I_{n,4}}
\end{aligned}$$

Para cada $s \in [0, T]$ definimos

$$t_n(s) = \sum_{i=0}^{n-1} t_{n,i} \mathbb{1}_{[t_{n,i}, t_{n,i+1})}(s).$$

Tenemos que $t_n(s) \xrightarrow{n \rightarrow \infty} s$ uniformemente.

Por la definición de integral estocástica tenemos que

$$I_{n,1} = \sum_{i=0}^{n-1} f'(W_{t_{n,i}})(W_{t_{n,i+1}} - W_{t_{n,i}}) = \int_0^t f'(W_{t_n(s)}) dW_s \xrightarrow{n \rightarrow \infty} \int_0^t f'(W_s) dW_s$$

en probabilidad.

Por otra parte, como f'' es continua tenemos que

$$I_{n,2} = \frac{1}{2} \sum_{i=0}^{n-1} f''(W_{t_{n,i}})(t_{n,i+1} - t_{n,i}) = \frac{1}{2} \int_0^t f''(W_{t_n(s)}) ds \xrightarrow{n \rightarrow \infty} \frac{1}{2} \int_0^t f''(W_s) ds$$

Utilizando que $|f'''(x)| \leq K$ para todo $x \in \mathbb{R}$ se obtiene que

$$\begin{aligned}
|I_{n,3}| &= \frac{1}{3!} \left| \sum_{i=0}^{n-1} f'''(W_{t_{n,i}})(W_{t_{n,i+1}} - W_{t_{n,i}})^3 \right| \\
&\leq \frac{1}{3!} \sum_{i=0}^{n-1} |f'''(W_{t_{n,i}})(W_{t_{n,i+1}} - W_{t_{n,i}})^3| \\
&\leq \frac{K}{3!} \sum_{i=0}^{n-1} |(W_{t_{n,i+1}} - W_{t_{n,i}})^3|
\end{aligned}$$

$$\leq \frac{K}{3!} \max_{0 \leq i \leq n-1} |W_{t_{n,i+1}} - W_{t_{n,i}}| \sum_{i=0}^{n-1} (W_{t_{n,i+1}} - W_{t_{n,i}})^2$$

Debido a que estamos considerando una sucesión de particiones que cumplen (3.3) y que, por definición, las trayectorias de un proceso de Wiener son continuas, se tiene que

$$\max_{0 \leq i \leq n-1} |W_{t_{n,i+1}} - W_{t_{n,i}}| \xrightarrow{n \rightarrow \infty} 0$$

casi seguramente.

Por otra parte, por la proposición (3.1.1),

$$\sum_{i=0}^{n-1} (W_{t_{n,i+1}} - W_{t_{n,i}})^2 \xrightarrow{n \rightarrow \infty} t,$$

en media cuadrática. Por lo tanto, $I_{n,3} \xrightarrow{n \rightarrow \infty} 0$ en probabilidad.

Resta probar que $I_{n,4} \xrightarrow{n \rightarrow \infty} 0$ en probabilidad. Para probar esto calculamos $\mathbb{E}(I_{n,4}^2)$.

$$\begin{aligned} \mathbb{E}(I_{n,4}^2) &= \mathbb{E} \left(\frac{1}{2} \sum_{i=0}^{n-1} f''(W_{t_{n,i}}) [(W_{t_{n,i+1}} - W_{t_{n,i}})^2 - (t_{n,i+1} - t_{n,i})] \right)^2 \\ &= \frac{1}{4} \sum_{i=0}^{n-1} \mathbb{E} [f''(W_{t_{n,i}})^2 [(W_{t_{n,i+1}} - W_{t_{n,i}})^2 - (t_{n,i+1} - t_{n,i})]^2] \\ &\quad + \frac{1}{2} \sum_{0 \leq i < j \leq n-1} \mathbb{E} (f''(W_{t_{n,i}}) [(W_{t_{n,i+1}} - W_{t_{n,i}})^2 - (t_{n,i+1} - t_{n,i})] \\ &\quad \times f''(W_{t_{n,j}}) [(W_{t_{n,j+1}} - W_{t_{n,j}})^2 - (t_{n,j+1} - t_{n,j})]) \end{aligned}$$

Para $i < j$ las variables

$$f''(W_{t_{n,i}}) [(W_{t_{n,i+1}} - W_{t_{n,i}})^2 - (t_{n,i+1} - t_{n,i})] f''(W_{t_{n,j}})$$

son $\mathcal{F}_{t_{n,j}}$ -medibles y, por definición de proceso de Wiener, $W_{t_{n,j+1}} - W_{t_{n,j}}$ es independiente de $\mathcal{F}_{t_{n,j}}$. Luego

$$\begin{aligned} &\mathbb{E} (f''(W_{t_{n,i}}) [(W_{t_{n,i+1}} - W_{t_{n,i}})^2 - (t_{n,i+1} - t_{n,i})] f''(W_{t_{n,j}}) \\ &\quad \times [(W_{t_{n,j+1}} - W_{t_{n,j}})^2 - (t_{n,j+1} - t_{n,j})]) = \\ &\mathbb{E} (f''(W_{t_{n,i}}) [(W_{t_{n,i+1}} - W_{t_{n,i}})^2 - (t_{n,i+1} - t_{n,i})] f''(W_{t_{n,j}})) \\ &\quad \times \underbrace{\mathbb{E} [(W_{t_{n,j+1}} - W_{t_{n,j}})^2 - (t_{n,j+1} - t_{n,j})]}_0 = 0 \end{aligned}$$

Como, por hipótesis, existe $C \in \mathbb{R}$ tal que $|f''(x)| \leq C$ obtenemos que

$$\mathbb{E}(I_{n,4}^2) = \frac{1}{4} \sum_{i=0}^{n-1} \mathbb{E} [f''(W_{t_{n,i}})^2 [(W_{t_{n,i+1}} - W_{t_{n,i}})^2 - (t_{n,i+1} - t_{n,i})]^2]$$

$$\begin{aligned}
&\leq \frac{C^2}{4} \sum_{i=0}^{n-1} \mathbb{E} [(W_{t_{n,i+1}} - W_{t_{n,i}})^2 - (t_{n,i+1} - t_{n,i})]^2 \\
&= \frac{C^2}{4} \sum_{i=0}^{n-1} \text{var} [(W_{t_{n,i+1}} - W_{t_{n,i}})^2] \\
&= \frac{C^2}{4} \sum_{i=0}^{n-1} \text{var} [(\sqrt{t_{n,i+1} - t_{n,i}} Z)^2] \\
&= \frac{C^2}{4} \text{var}(Z^2) \frac{C^2}{4} \sum_{i=0}^{n-1} (t_{n,i+1} - t_{n,i})^2 \\
&\leq \frac{C^2}{4} \text{var}(Z^2) \max_{0 \leq k \leq n-1} |t_{n,k+1} - t_{n-k}| t \xrightarrow{n \rightarrow \infty} 0,
\end{aligned}$$

donde $Z \sim N(0, 1)$. ■

Teorema 3.3.2 (Fórmula de Itô). Sea $\{W_t\}_{t \geq 0}$ un proceso de Wiener adaptado a la filtración $\{\mathcal{F}_t\}_{t \geq 0}$. Supongamos que, además, el incremento $W_{t+h} - W_t$ es independiente de la σ -álgebra $\mathcal{F}_t$. Sea $f : \mathbb{R} \rightarrow \mathbb{R}$ de clase C^2 . Entonces:

$$f(W_t) - f(0) = \int_0^t f'(W_s) dW_s + \frac{1}{2} \int_0^t f''(W_s) ds$$

Demostración. Ver [Borodin, 2017], p. 100. ■

Ejemplo 3.3.1. Sea $f : \mathbb{R} \rightarrow \mathbb{R}$ tal que $f(x) = x^2$. Como

$$f'(x) = 2x, \quad f''(x) = 2$$

por (3.3.2) tenemos que

$$\begin{aligned}
W_t^2 &= 2 \int_0^t W_s dW_s + \int_0^t ds \\
&= 2 \int_0^t W_s dW_s + t
\end{aligned}$$

Reescribiendo esto:

$$\int_0^t W_s dW_s = \frac{W_t^2}{2} - \frac{t}{2}$$

Definición 3.3.1. Sea $(\Omega, \mathcal{A}, \mathbf{P})$ un espacio de probabilidad. Sean $W = \{W_t\}_{t \in [0, T]}$ un proceso de Wiener y ξ una variable aleatoria independiente de W . Sea $\mathcal{F}_t = \sigma\{\xi, W_s : 0 \leq s \leq t\}$ la σ -álgebra generada por ξ y el proceso de Wiener en el intervalo $[0, t]$.

Se consideran $a, b : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}$ funciones medibles.

Un proceso $X = \{X_t\}_{t \in [0, T]}$ se dice que es la **solución fuerte** de la ecuación diferencial estocástica

$$\begin{cases} dX_t = a(t, X_t)dt + b(t, X_t)dW_t \\ X_0 = \xi \end{cases}$$

si se cumple que

1. X es un proceso con trayectorias continuas.

2. X es adaptado a $\mathcal{F}_t$.

3. Para todo $t \in [0, T]$ se cumple que

$$\int_0^t |a(s, X_s)| + |b(s, X_s)|^2 ds < \infty$$

4. Para todo $t \in [0, T]$ se cumple que

$$X_t = \xi + \int_0^t a(s, X_s) ds + \int_0^t b(s, X_s) dW_s$$

Ejemplo 3.3.2. A partir del ejemplo (3.3.1) tenemos que el proceso $X_t = W_t^2$ es solución de la ecuación diferencial estocástica

$$\begin{cases} dX_t = dt + X_t dW_t \\ X_0 = 0 \end{cases}$$

Proposición 3.3.1. Sean $a, b : [0, T] \rightarrow \mathbb{R}$. Se asume que existe $C \in \mathbb{R}$ tal que para todo $t \in [0, T]$ y $x, y \in \mathbb{R}$

$$|a(t, x) - a(t, y)| + |b(t, x) - b(t, y)| \leq C|x - y|$$

y

$$|a(t, x)| + |b(t, y)| \leq C(1 + |x|)$$

Sea ξ una variable aleatoria tal que $\mathbb{E}(\xi^2) < \infty$. Entonces, existe una única solución fuerte de la ecuación diferencial estocástica

$$\begin{cases} dX_t = a(t, X_t)dt + b(t, X_t)dW_t \\ X_0 = \xi \end{cases}$$

Demostración. Ver [Borodin, 2017], p. 130. ■

Definición 3.3.2 (Difusión). Sea $X = \{X_t\}_{t \in [0, T]}$ un proceso markoviano que toma valores en $\mathbb{R}$. El proceso X es una **difusión** si se cumplen las siguientes condiciones:

1) Para todo $\varepsilon > 0$, todo $t \in [0, T]$ y todo $x \in \mathbb{R}$ se cumple que

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

2) Existen funciones $a, b : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}$ tales que existe $\varepsilon > 0$ tal que para todo $t \in [0, T]$ y todo $x \in \mathbb{R}$ se cumple que

$$\begin{aligned} \lim_{h \rightarrow 0} \frac{\mathbb{E}_x [(X_{t+h} - X_t) \mathbf{1}_{[0, \varepsilon]}(|X_{t+h} - X_t|)]}{h} &= a(t, x) \\ \lim_{h \rightarrow 0} \frac{\mathbb{E}_x [(X_{t+h} - X_t)^2 \mathbf{1}_{[0, \varepsilon]}(|X_{t+h} - X_t|)]}{h} &= b^2(t, x) \end{aligned}$$

Las funciones $a(t, x)$ y $b^2(t, x)$ se denominarán *coeficiente de deriva* y *coeficiente de difusión*, respectivamente.

El proceso estocástico $X = \{X_t\}$ es una **difusión homogénea en el tiempo** si

1) Para todo $\varepsilon > 0$, todo $t \in [0, T]$ y todo $x \in \mathbb{R}$ se cumple que

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

2) Existen funciones $a, b : \mathbb{R} \rightarrow \mathbb{R}$ tales que existe $\varepsilon > 0$ tal que para todo $t \in [0, T]$ y todo $x \in \mathbb{R}$ se cumple que

$$\begin{aligned} \lim_{h \rightarrow 0} \frac{\mathbb{E}[(X_{t+h} - X_t) \mathbf{1}_{[0, \varepsilon]}(|X_{t+h} - X_t|) | X_t = x]}{h} &= a(x) \\ \lim_{h \rightarrow 0} \frac{\mathbb{E}[(X_{t+h} - X_t)^2 \mathbf{1}_{[0, \varepsilon]}(|X_{t+h} - X_t|) | X_t = x]}{h} &= b^2(x) \end{aligned}$$

La primera condición para que un proceso sea una difusión indica que las trayectorias del proceso no presentan cambios grandes en tiempos arbitrariamente chicos. Esto es, el proceso tendrá trayectorias continuas.

La segunda condición puede ser interpretada de la siguiente forma. La esperanza de los incrementos truncados condicionado a que se parte en un tiempo t del punto x sobre un periodo infinitamente pequeño está dada por $a(x)$ mientras que la varianza condicional está dada por $b^2(x)$.

Proposición 3.3.2. Sean $a, b : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}$ funciones continuas. Sea X solución fuerte de

$$\begin{cases} dX_t = a(t, X_t)dt + b(t, X_t)dW_t \\ X_0 = \xi \end{cases}$$

Entonces X es una difusión con coeficiente de deriva $a(t, x)$ y coeficiente de difusión $b^2(t, x)$.

Demostración. Ver [Borodin, 2017], p. 274. ■

Proposición 3.3.3 (Generalización de la fórmula de Itô). Sea $X = \{X_t\}_{t \in [0, T]}$ la solución fuerte de la ecuación diferencial estocástica

$$\begin{cases} dX_t = a(X_t)dt + b(X_t)dW_t \\ X_0 = x \in \mathbb{R} \end{cases}$$

con a, b funciones càdlàg en $\mathbb{R}$. Se considera una función $f : \mathbb{R} \rightarrow \mathbb{R} \in C^2$. Entonces se cumple que

$$f(X_t) - f(X_0) = \int_0^t f'(X_s)a(X_s)ds + \int_0^t f'(X_s)b(X_s)dW_s + \frac{1}{2} \int_0^t f''(X_s)(b(s))^2 ds$$

Demostración. Ver [Borodin, 2017], p. 104. ■

A continuación, utilizaremos la proposición (3.3.3) para hallar el generador infinitesimal del semigrupo asociado a la solución fuerte de una ecuación diferencial estocástica.

Sea $X = \{X_t\}_{t \in [0, T]}$ un proceso que toma valores reales y es la solución de la ecuación diferencial estocástica

$$\begin{cases} dX_t = a(X_t)dt + b(X_t)dW_t \\ X_0 = x \in \mathbb{R} \end{cases}$$

Supongamos que $b(x) > 0$ para todo $x \in \mathbb{R}$ y que a, b son funciones Lipschitz.

Por la proposición (3.3.2) el proceso X es markoviano y tiene una función de transición $P(s, x, B)$. Recordemos el semigrupo de operadores definidos en $\mathfrak{B}$

$$(S_t f)(x) = \mathbb{E}_x(f(X_t))$$

Sea $C_a^2(\mathbb{R})$ el espacio de funciones de clase C^2 cuyas derivadas son acotadas. A partir de la proposición (3.3.3) podemos obtener el generador infinitesimal de este semigrupo:

$$\begin{aligned} (\mathcal{G}f)(x) &= \lim_{t \rightarrow 0} \frac{(S_t f)(x) - f(x)}{t} = \lim_{t \rightarrow 0} \frac{\mathbb{E}_x(f(X_t)) - f(x)}{t} \\ &= \lim_{t \rightarrow 0} \frac{\mathbb{E}_x(f(X_t) - f(x))}{t} \\ &= \lim_{t \rightarrow 0} \frac{\mathbb{E}_x(f(X_t) - f(X_0))}{t} \\ &= \underbrace{\lim_{t \rightarrow 0} \frac{\mathbb{E}_x(\int_0^t f'(X_s)b(X_s) dW_s)}{t}}_{(1)} + \underbrace{\lim_{t \rightarrow 0} \frac{\mathbb{E}_x(\int_0^t f'(X_s)a(X_s) ds)}{t}}_{(2)} \\ &\quad + \underbrace{\lim_{t \rightarrow 0} \frac{\mathbb{E}_x(\frac{1}{2} \int_0^t f''(X_s)(b(X_s))^2 ds)}{t}}_{(3)} \end{aligned}$$

Estudiemos la expresión en (1). Por la proposición (3.2.4) tenemos que

$$\mathbb{E}_x \left(\int_0^t f'(X_s)b(X_s) dW_s \right) = 0$$

para todo $t \in [0, T]$. Luego

$$\lim_{t \rightarrow 0} \frac{\mathbb{E} \left(\int_0^t f'(X_s)b(X_s) dW_s | X_0 = x \right)}{t} = 0$$

Para la expresión en (2) se tiene que

$$\begin{aligned} \lim_{t \rightarrow 0} \frac{\mathbb{E}_x(\int_0^t f'(X_s)a(X_s) ds)}{t} &= \mathbb{E}_x \left(\lim_{t \rightarrow 0} \frac{\int_0^t f'(X_s)a(X_s) ds}{t} \right) \\ &= f'(x)a(x) \end{aligned}$$

De forma análoga, para la expresión en (3) se tiene que

$$\begin{aligned} \lim_{t \rightarrow 0} \frac{\mathbb{E}_x \left(\frac{1}{2} \int_0^t f''(W_s) (b(X_s))^2 ds \right)}{t} &= \mathbb{E}_x \left(\frac{1}{2} \lim_{t \rightarrow 0} \frac{\int_0^t f''(W_s) (b(X_s))^2 ds}{t} \right) \\ &= \frac{1}{2} f''(x) b^2(x) \end{aligned}$$

En conclusión, el generador infinitesimal $\mathcal{G} : C_a^2(\mathbb{R}) \rightarrow C^2(\mathbb{R})$ de la solución de la ecuación diferencial estocástica

$$\begin{cases} dX_t = a(X_t)dt + b(X_t)dW_t \\ X_0 = x \in \mathbb{R} \end{cases}$$

es

$$(\mathcal{G}f)(x) = f'(x)a(x) + \frac{1}{2}f''(x)b^2(x) \quad (3.4)$$

A continuación enunciaremos, sin demostrar, una proposición que será útil para dar fórmulas explícitas a las esperanzas de ciertos tiempos de parada.

Proposición 3.3.4. Sea $X = \{X_t\}_{t \geq 0}$ solución de la ecuación diferencial estocástica

$$\begin{cases} dX_t = a(X_t)dt + b(X_t)dW_t \\ X_0 = x \in \mathbb{R} \end{cases}$$

Sea el tiempo de parada $H_{\alpha, \beta} = \min\{t \geq 0 : X_t \notin (\alpha, \beta)\}$. Se consideran $f, F : [a, b] \rightarrow \mathbb{R}$ continuas a trozos y f no negativa. Sea $\phi : \{a, b\} \rightarrow \mathbb{R}$. Entonces

$$\begin{aligned} Q(x) &= \mathbb{E}_x \left[\phi(X_{H_{\alpha, \beta}}) \exp \left(- \int_0^{H_{\alpha, \beta}} f(X_s) ds \right) \right. \\ &\quad \left. + \int_0^{H_{\alpha, \beta}} F(X_s) \exp \left(- \int_0^s f(X_u) du \right) ds \right] \end{aligned}$$

es la única solución a la ecuación diferencial

$$\begin{cases} \frac{1}{2}b^2(x)Q''(x) + a(x)Q'(x) - f(x)Q(x) = -F(x) \\ Q(a) = \phi(a) \\ Q(b) = \phi(b) \end{cases}$$

Demostración. Ver [Borodin, 2017], p. 291. ■

Observación 3.3.1. En particular, tomando en la proposición (3.3.4) $f(x) = 0$, $F(x) = 1$ y $\phi(a) = \phi(b) = 0$ se obtiene que $Q(x) = \mathbb{E}_x(H_{a,b})$ es la solución de la ecuación diferencial

$$(\mathcal{G}Q)(x) = -1, \quad Q(a) = Q(b) = 0,$$

siendo $\mathcal{G}$ el generador infinitesimal del semigrupo asociado a la difusión $X = \{X_t\}_{t \geq 0}$.

3.3.1. Aproximaciones numéricas

Esquema de Euler

Sea $X = \{X_t\}_{0 \leq t \leq T}$ solución fuerte de la ecuación diferencial estocástica

$$\begin{cases} dX_t = a(t, X_t)dt + b(t, X_t)dW_t \\ X_0 = \xi \end{cases}$$

El *esquema de Euler-Maruyama* es un proceso iterativo para obtener trayectorias aproximadas de X :

Dada una partición del intervalo $[0, T]$, $t_0 = 0 < t_1 < \dots < t_n = T$ definimos como aproximación de Euler al proceso estocástico de tiempo continuo $Y = \{Y_t\}_{0 \leq t \leq T}$ que satisface que

$$Y_{i+1} = Y_i + a(t_i, Y_i)(t_{i+1} - t_i) + b(t_i, Y_i)(W_{t_{i+1}} - W_{t_i}),$$

con $W = \{W_t\}_{0 \leq t \leq T}$ un proceso de Wiener.

Este esquema determina valores aproximados sólo en los puntos de la partición $t_0, t_1, \dots, t_n$. Para obtener valores de los puntos intermedios utilizaremos interpolación lineal. Para cada $t \in [0, T]$ definimos $n_t = \max_{i=0,1,\dots,n} \{t_i \leq t\}$. Luego la interpolación que realizaremos es

$$Y_t = Y_{n_t} + \frac{t - t_{n_t}}{t_{n_t+1} - t_{n_t}}(Y_{n_t+1} - Y_{n_t})$$

Una prueba de que la aproximación Y es consistente puede verse en [Kloeden y Platen, 1995].

Esquema de Milstein

Sea $X = \{X_t\}_{0 \leq t \leq T}$ un proceso de Ito que satisface la ecuación diferencial estocástica

$$\begin{cases} dX_t = a(t, X_t)dt + b(t, X_t)dW_t \\ X_0 = \xi \end{cases}$$

Sea una partición del intervalo $[0, T]$, $t_0 = 0 < t_1 < \dots < t_n = T$. En el esquema de Milstein, nos referiremos como aproximación de Milstein al proceso estocástico de tiempo continuo $Y = \{Y_t\}_{0 \leq t \leq T}$ que surge de añadir un término adicional al esquema de Euler:

$$\begin{aligned} Y_{i+1} = & Y_i + a(t_i, Y_i)(t_{i+1} - t_i) + b(t_i, Y_i)(W_{t_{i+1}} - W_{t_i}), \\ & + \frac{1}{2}b(t_i, Y_i)\frac{\partial}{\partial x}b(t_i, Y_i) [(W_{t_{i+1}} - W_{t_i})^2 - (t_{i+1} - t_i)] \end{aligned}$$

con $W = \{W_t\}_{0 \leq t \leq T}$ un proceso de Wiener.

Al igual que en el esquema de Euler, utilizaremos interpolación lineal para obtener valores intermedios entre $[t_i, t_{i+1}]$.

Una prueba de que la aproximación Y es consistente puede verse en [Kloeden y Platen, 1995].

Capítulo 4

Modelos de genética de poblaciones

4.1. Conceptos de genética

En esta sección se introducirán algunos términos básicos de genética para entender posteriormente los modelos de genética de poblaciones estudiados. Esta sección no pretende ser exhaustiva sino familiarizar al lector no versado en genética con algunos conceptos importantes para este trabajo. Guías más desarrolladas que vinculan la genética con la evolución pueden verse en los libros [[Griffiths et al., 2004](#)], [[Herron y Freeman, 2014](#)], [[Futuyma, 2005](#)].

Los individuos de una misma especie no son iguales: presentan variaciones. Estas variaciones pueden tener diversos orígenes: los individuos pueden diferir entre ellos debido a las variaciones en los genes que poseen, o debido a las interacciones que tienen con el ambiente, o debido a cómo el ambiente afecta la expresión de ciertos genes. Estamos interesados en este trabajo en las variaciones en los genes, llamadas de ahora en más *variaciones genéticas*. Pero ¿a qué nos referimos con gen?

Gran parte de las estructuras de las células que forman a los organismos están formadas por proteínas. Las proteínas son cadenas formadas por aminoácidos. Los veinte aminoácidos que existen le dan a la proteína propiedades químicas diversas. En particular, a pesar del término *cadena*, las proteínas no son estructuras lineales si no que tienden a plegarse y formar estructuras tridimensionales. Qué forma tendrán estas estructuras dependerá de qué aminoácidos conformen la proteína. La información sobre cómo construir las cadenas de aminoácidos está en el llamado *material genético*.

En algunos virus, el material genético será el ARN; en el resto de los organismos será el ADN. Tanto el ARN como el ADN son ácidos nucleicos: cadenas de nucleótidos unidos entre sí mediante enlaces fosfodiéster. Los nucleótidos son moléculas formadas por un grupo fosfato, un azúcar y una base nitrogenada. En el ARN el azúcar es una ribosa mientras que en el ADN el azúcar es una desoxirribosa. Las bases nitrogenadas que forman parte de los nucleótidos del ADN son adenina (A), citocina (C), guanina (G) y uracilo (U); los nucleótidos del ADN utilizarán también A, C y G, pero reemplazarán U por timina (T). Podemos pensar entonces a los ácidos nucleicos como ‘palabras’ formadas por el alfabeto $\{A, C, G, U\}$ en el caso del ARN y $\{A, C, G, T\}$ en el caso del ADN —aunque esta comparación no es del todo correcta porque, al igual que las proteínas, los ácidos

nucleicos formarán también estructuras tridimensionales.

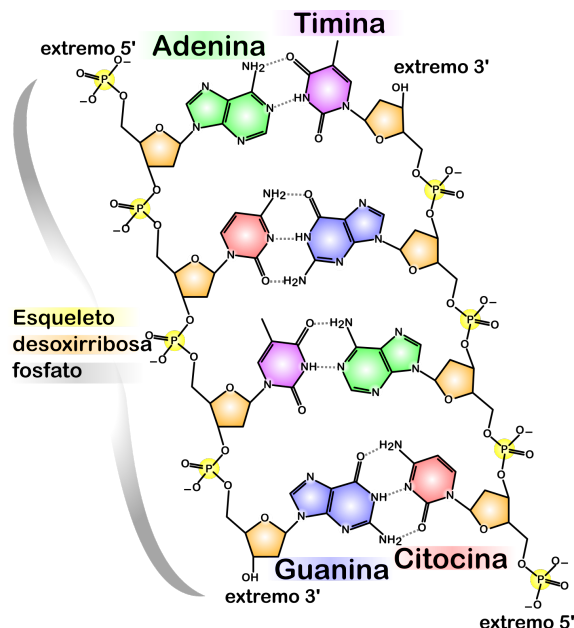


Figura 4.1: Estructura del ADN.

En porciones de la cadena llamadas **regiones codificantes** las cadenas de ADN o ARN son ‘leídas’ de a ternas de nucleótidos y estas ternas son codificadas en un aminoácido en particular —a este conjunto de tres letras que codifican para un aminoácido lo denominaremos **codón**. Aquellas regiones del ADN o ARN (aquellas ‘palabras’) que una vez leídas dan como resultado una proteína serán denominados **genes**¹. Los genes de los organismos están integrados dentro de largas cadenas de ADN (o ARN) denominados cromosomas. Al conjunto de cromosomas de un organismo (que puede incluir desde un solo cromosoma como es el caso de las bacterias a centenas de cromosomas como en algunas especies de plantas) se le denomina **genoma**. Al lugar físico en donde se encuentra un gen dentro de un genoma lo denominaremos el **locus** del gen; denominaremos también locus a posiciones arbitrarias dentro del genoma, aún cuando estas no sean codificantes.

Dentro de una población, existen variaciones en los genomas de los individuos, que llamaremos variación genética. Pensemos en una posición en particular en el ADN (o ARN) de los organismos de una población. Supongamos que la mitad de la población tiene en esta posición una A y la otra mitad tiene una G. Diremos en este caso que existen dos variantes o **alelos** en esta población para esta posición. Eventualmente, podrían existir cuatro alelos en esta población (cada alelo correspondiente a cada una de las bases nitrogenadas, A, G, C, T). De la misma forma, si en vez de considerar una posición en particular dentro del genoma, consideramos un codón en particular tendremos 64 alelos posibles —en la primera, segunda y tercera posición tenemos 4 alelos y, por lo tanto, tenemos $4 \times 4 \times 4 = 64$ alelos.

Desde el punto de vista de la genética de poblaciones una **población** es un grupo de

¹Esta definición, desde un punto de vista bioquímico moderno, no es del todo correcta. El concepto de gen es altamente discutido. Una revisión histórica de este concepto puede verse en [Portin y Wilkins, 2017]

organismos de una misma especie que habitan en un área geográfica restringida que les permite aparearse entre sí. Como en esta definición hablamos de *apareamiento*, en esta definición está el supuesto de que los organismos considerados presentan reproducción sexual.

En individuos sexuales, los genotipos no son transmitidos de una generación a la siguiente. Los genotipos son divididos en gametos a través de los procesos de segregación de gametos y de recombinación; luego estos gametos forman genotipos en el proceso de fecundación:

$$\text{genotipo} \rightarrow \text{gametos} \rightarrow \text{genotipo} \quad (4.1)$$

La cuestión central de la genética de poblaciones es estudiar cómo las frecuencias de ciertos alelos evolucionan de una generación a la siguiente. Para esto, construiremos modelos explicitando ciertos parámetros de las poblaciones. Por ejemplo: cuál es el tamaño poblacional, cuál es la regla con la que se aparean los gametos, cómo se introducen (si es que esto ocurre) nuevos alelos a la población y cuál es la supervivencia relativa de los individuos portadores de cada alelo.

4.2. Equilibrio de Hardy-Weinberg

El modelo más sencillo de apareamiento es el de *apareamiento al azar*: en este modelo, la probabilidad de que un individuo se aparee con otro de determinado genotipo es igual a la frecuencia de este genotipo en la población. Una manera de pensar este modelo de apareamiento es la siguiente: Tomamos todos los gametos (óvulos y espermatozoides) en una gran bolsa, seleccionamos un espermatozoide y un óvulo con los ojos cerrados; este par formará un nuevo cigoto. Este modelo es claramente una simplificación de cómo ocurre el apareamiento en poblaciones reales, por ejemplo, de mamíferos, pero puede no ser tan descabellado para aquellas especies marinas con fecundación externa.

Supongamos que esta población entra en etapa reproductiva de forma sincronizada y que segrega infinitos gametos A y a . Siguiendo la analogía de la bolsa, supongamos que la probabilidad de sacar un gameto A (sea óvulo o espermatozoide) de la bolsa es p y la probabilidad de sacar un alelo a es q . Como estamos hablando de probabilidades, deberá cumplirse que

$$p + q = 1$$

Hagamos el supuesto de que la población se reproduce de forma aleatoria. De nuestra gran bolsa saquemos un espermatozoide A y un óvulo A . Por el supuesto del apareamiento al azar tenemos que

$$\begin{aligned} \mathbf{P}(\text{gameto } A \text{ y gameto } A) &= \mathbf{P}(\text{gameto } A)\mathbf{P}(\text{gameto } A) \\ &= p \times p = p^2 \end{aligned}$$

De manera análoga tendremos que

$$\begin{aligned} P(\text{gameto } a \text{ y gameto } a) &= P(\text{gameto } a) \times P(\text{gameto } a) \\ &= q \times q = q^2 \end{aligned}$$

y

$$\begin{aligned} P(\text{gameto } A \text{ y gameto } a) &= P(\text{gameto } a)P(\text{gameto } A) \\ &= p \times q \end{aligned}$$

En la siguiente población, por lo tanto, tendremos las siguientes proporciones:

AA	Aa	aa
p^2	$2pq$	q^2

Dejemos que los cigotos se desarrollen y se conviertan en adultos. Estamos interesados ahora en calcular cómo serán las frecuencias de los alelos A y a en este nuevo *pool* génico. Como los adultos A_1A_1 constituyen una proporción de p^2 de la población, brindarán una proporción p^2 de gametos A_1 . De la misma forma, los adultos A_1A_2 constituyen una proporción de $2pq$ de la población y por tanto brindarán una proporción de $2pq$ gametos; la mitad de estos gametos serán A_1 y la otra mitad A_2 . Por lo tanto, la frecuencia del alelo A_1 en este nuevo *pool* génico será

$$p^2 + \frac{1}{2}2pq = p^2 + pq = p^2 + p(1 - p) = p$$

Realizando operaciones análogas se obtiene que la frecuencia del alelo a en el nuevo *pool* génico será q . Este hecho arroja una consecuencia fundamental de este modelo y que es la razón de su nombre, **equilibrio de Hardy-Weinberg**: *las frecuencias alélicas de una generación a la siguiente no cambian*.

Supuestos del equilibrio de Hardy-Weinberg

Durante del desarrollo del modelo se realizaron, de forma implícita, varios supuestos que serán explicitados a continuación. El equilibrio de Hardy-Weinberg muestra que las frecuencias alélicas de dos alelos A y a para un locus A no cambiarán de generación a generación si la población y el locus considerados están bajo los siguientes supuestos:

1. Población infinita
2. Ausencia de selección natural
3. Ausencia de mutación
4. Ausencia de migración desde otras poblaciones
5. Organismos diploides.
6. Reproducción sexual

7. Apareamiento al azar
8. Generaciones no solapadas
9. Locus autosómico; es decir, el locus en cuestión no está en un cromosoma sexual de la población.
10. Frecuencias alélicas no dependientes del sexo.

Debido a la cantidad enorme de supuestos que presenta este modelo, es de esperarse que la aplicación a poblaciones reales esté restringida. Mientras que algunos de estos supuestos pueden ser fácilmente modificados (el de organismos diploides, por ejemplo), otros supuestos serán imprescindibles para que las frecuencias alélicas de generación a generación permanezcan incambiadas. En particular, uno de los supuestos que caerán inmediatamente al estudiar poblaciones reales es el de población infinita: ¿qué ocurre con las frecuencias alélicas cuando consideramos poblaciones finitas pero dejamos el resto de los supuestos fijos? La respuesta rápida: la población evoluciona. Esto será estudiado en el siguiente apartado.

Generalización a k alelos

El razonamiento realizado anteriormente puede ser fácilmente extendido al caso donde existan k alelos, $A_1, \dots, A_k$ para un locus particular. Supongamos que las frecuencias de los alelos $A_1, \dots, A_k$ son $p_1, \dots, p_k$ respectivamente. Al igual que en el caso de dos alelos, se cumplirá que

$$p_1 + p_2 + \dots + p_k = 1$$

Permitamos que los adultos segreguen infinitos gametos y que estos gametos se reproduzcan al azar. La probabilidad de formar un cigoto de tipo $A_i A_i$, con $i = 1, \dots, k$ es p_i^2 . Por otra parte, para calcular la probabilidad de formar un cigoto $A_i A_j$, con $i \neq j$, observemos que esta conformación puede darse o bien por la unión de un óvulo A_i con un espermatozoide A_j o bien por la unión de un espermatozoide A_i y un óvulo A_j : ambas uniones tienen probabilidad $p_i p_j$ de ocurrir. Luego, la probabilidad de obtener un cigoto $A_i A_j$ es $2p_i p_j$. Podemos escribir la probabilidad de un cigoto cualquiera, entonces, como

$$\mathbf{P}(A_i A_j) = (2 - \delta_{ij}) p_i p_j, \quad \delta_{ij} = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases}$$

Calculemos, a continuación, la frecuencia del alelo A_1 a partir de la frecuencia de los genotipos $A_i A_j$ luego de la reproducción —la frecuencia del resto de los alelos se podrá calcular de manera análoga:

$$\begin{aligned} \mathbf{P}(A_1) &= \mathbf{P}(A_1 A_1) + \frac{1}{2} \mathbf{P}(A_1 A_2) + \dots + \frac{1}{2} \mathbf{P}(A_1 A_k) \\ &= p_1^2 + p_1 p_2 + \dots + p_1 p_k = p_1 \underbrace{(p_1 + p_2 + \dots + p_k)}_1 = p_1 \end{aligned}$$

Por lo tanto, al igual que en el caso de dos alelos, en el caso de k alelos, si la población sigue los supuestos del equilibrio de Hardy-Weinberg, las frecuencias alélicas se mantienen de generación en generación.

4.3. Modelo de Wright-Fisher

4.3.1. Introducción y propiedades

El siguiente modelo planteado consiste en considerar una población que está bajo todos los supuestos del equilibrio de Hardy-Weinberg excepto el considerado sobre el tamaño poblacional. En este caso, el tamaño poblacional dejará de ser infinito y pasará a tener un tamaño fijo N :

Sean una población diploide de tamaño N y un locus en particular en el genoma de esta población; existen, por ser una población diploide, $2N$ copias de este locus. Supongamos que para este locus existen dos alelos, A y a , en esta población.

Comencemos con una población inicial para el cual la frecuencia del alelo A es $i/2N$ y la del alelo a es $(2N - i)/2N$. Por ejemplo, supongamos que nuestra población inicial es

$$AA \quad aA \quad AA \quad AA \quad aa,$$

en este caso, la frecuencia del alelo A es $7/10$ y la del tipo a es $3/10$. Queremos construir, a partir de esta población, la población de la siguiente generación. Asumiremos que el tamaño de la población es constante; por lo tanto, la población en la siguiente generación será también de tamaño $2N$. Supondremos que los individuos segregarán infinitos alelos y que se reproducirán al azar. Esto nos permitirá construir la nueva población de la siguiente forma: muestrearemos $2N$ alelos con reposición a partir de la generación inicial; muestrearemos alelos A con probabilidad $7/10$ y alelos a con probabilidad $3/10$; cada alelo será muestreado con independencia del otro. En la siguiente generación, entonces, podríamos obtener la población

$$aa \quad aA \quad Aa \quad AA \quad AA.$$

La frecuencia del alelo A en esta nueva generación es $6/10$ y la del alelo a es $4/10$. ¿Cuál es la probabilidad de obtener estas frecuencias dadas mis frecuencias en la generación anterior?

Observemos que se obtienen 6 A con probabilidad $\left(\frac{7}{10}\right)^6$ y que se obtienen 4 a con probabilidad $\left(\frac{3}{10}\right)^4$. Como la probabilidad de obtener a es independiente de la de obtener A y, además, no nos interesará cómo están conformados los pares, la frecuencia del alelo A en la siguiente generación, dada mi generación inicial, es $\binom{10}{6} \left(\frac{7}{10}\right)^6 \left(\frac{3}{10}\right)^4 \approx 0,2$.

Sea $Z_n : \Omega \rightarrow \{0, 1, \dots, 2N\}$ el número de alelos de tipo A en la generación n . Como el número de alelos A de la generación $n + 1$ sólo depende de la cantidad de alelos de la generación n , $\{Z_n\}_{n \in \mathbb{N}}$ es una cadena de Márkov con espacio de estados $E = \{0, 1, \dots, 2N\}$. Siguiendo el razonamiento visto en el ejemplo anterior, obtenemos probabilidades de transición que siguen una distribución binomial:

$$p_{ij} = \mathbf{P}(Z_{n+1} = j | Z_n = i) = \binom{2N}{j} \left(\frac{i}{2N}\right)^j \left(1 - \frac{i}{2N}\right)^{2N-j} \quad (4.2)$$

Definición 4.3.1 (Proceso de Wright-Fisher). *Diremos que una cadena de Markov en tiempo discreto $Z = \{Z_n\}_{n \in \mathbb{N}}$ es un **proceso de Wright-Fisher** de parámetros $N \in \mathbb{N}$ y $p_0 \in \{0, 1/2N, \dots, 1\}$ sobre $E = \{0, 1, \dots, 2N\}$ Si*

- i) *la distribución inicial es $\mathbf{P}(Z_0 = 2Np_0) = 1$ y,*
- ii) *la matriz de transición $\mathbb{P} = ((p_{ij}))_{i,j=0,1,\dots,2N}$ es tal que*

$$p_{ij} = \mathbf{P}(Z_{n+1} = j | Z_n = i) = \binom{2N}{j} \left(\frac{i}{2N}\right)^j \left(1 - \frac{i}{2N}\right)^{2N-j}$$

Observación 4.3.1. El proceso presenta dos estados absorbentes, a saber, 0 y $2N$. Esto se obtiene debido a la forma en que la generación en tiempo $n + 1$ a partir de la generación n . A modo de ejemplo, si nuestra población en tiempo n es

$$AA \quad AA \quad AA \quad AA \quad AA,$$

en tiempo $n + 1$ sólo podremos obtener una población similar; cuando se obtiene una población de este tipo diremos que el *alelo A se fijó*. El caso análogo ocurre cuándo el alelo a se fija.

En la figura (4.2) se pueden observar trayectorias simuladas de un proceso de Wright-Fisher. Podemos notar que a medida que el número de generaciones aumenta, la varianza de Z_n aumenta. Probaremos este resultado a continuación. Para esto será necesario probar otra característica de un proceso de Wright-Fisher: este proceso es una martingala en tiempo discreto y espacio de estados finito.

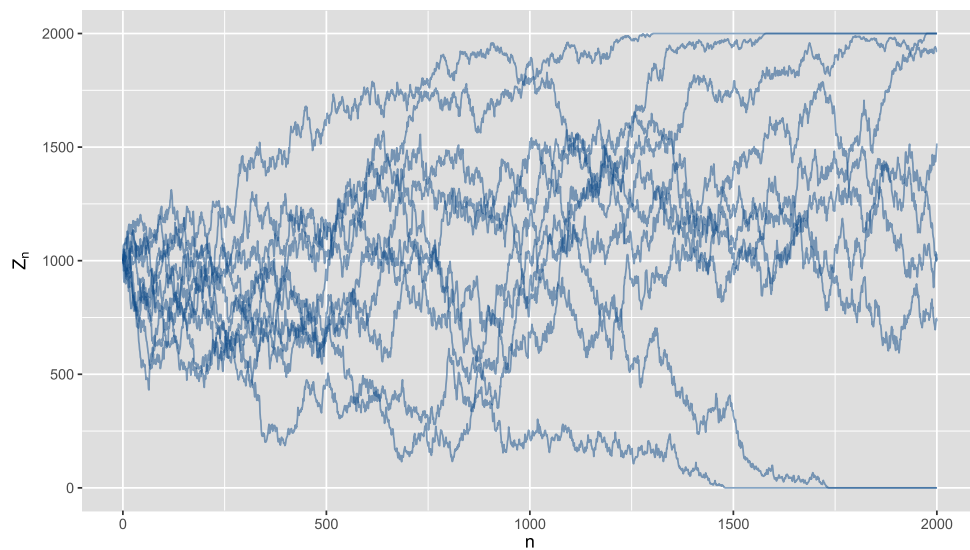


Figura 4.2: Trayectorias de un proceso de Wright-Fisher de parámetros $N = 1000$ y $z_0 = 1000$.

Proposición 4.3.1. Todo proceso de Wright-Fisher de parámetros N y p_0 , $Z = \{Z_n\}_{n \in \mathbb{N}}$ es una martingala con respecto a $\mathcal{F}$ la filtración generada por Z .

Demostración. Debemos probar que Z cumple las propiedades que definen una martingala.

- Para todo $n \in \mathbb{N}$ se cumple que $\mathbb{E}(Z_n) = \sum_{i=0}^{2N} i \mathbf{P}(Z_n = i) < \infty$
- Por la observación (2.2.4), Z está adaptado a $\mathcal{F}$.
- Por la estructura markoviana del proceso se obtiene que

$$\mathbb{E}(Z_{n+1} | \mathcal{F}_n) = \mathbb{E}(Z_{n+1} | Z_n)$$

Luego, como $Z_{n+1} | Z_n \sim \text{Bin}(2N, Z_n/2N)$ se tiene que

$$\mathbb{E}(Z_{n+1} | Z_n) = 2N \frac{Z_n}{2N} = Z_n$$

■

Corolario 4.3.1. Sea $Z = \{Z_n\}_{n \in \mathbb{N}}$ un proceso de Wright-Fisher de parámetros N y z_0 . Entonces, para todo $n \in \mathbb{N}$ se cumple que $\mathbb{E}(Z_n) = z_0$.

Demostración. Por la proposición (2.2.8), $\mathbb{E}(Z_n) = \mathbb{E}(Z_0) = z_0$ para todo $n \in \mathbb{N}$. ■

Proposición 4.3.2. Sea $Z = \{Z_n\}_{n \in \mathbb{N}}$ un proceso de Wright-Fisher de parámetros N y p_0 . Entonces, para todo $n \in \mathbb{N}$ se cumple que

$$\text{var}(Z_n) = z_0(1 - z_0) \left(1 - \left(1 - \frac{1}{2N}\right)^n\right)$$

Demostración. Sea $n \in \mathbb{N}$. Por definición de varianza se tiene que

$$\text{var}(Z_n) = \mathbb{E}(Z_n^2) - [\mathbb{E}(Z_n)]^2$$

Analicemos el segundo término:

$$\begin{aligned} \mathbb{E}(Z_n^2) &= \mathbb{E}[\mathbb{E}(Z_n^2 | Z_{n-1})] \\ &= \mathbb{E}(\text{var}(Z_n | Z_{n-1}) + \mathbb{E}(Z_n | Z_{n-1})^2) \\ &= \mathbb{E}\left(2N \left(\frac{Z_{n-1}}{2N}\right) \left(1 - \frac{Z_{n-1}}{2N}\right) + (Z_{n-1})^2\right) \\ &= \mathbb{E}\left(Z_{n-1} - \frac{Z_{n-1}^2}{2N} + Z_{n-1}^2\right) \\ &= \mathbb{E}(Z_{n-1}) + \left(1 - \frac{1}{2N}\right) \mathbb{E}(Z_{n-1}^2) \\ &= \mathbb{E}(Z_{n-1}) + \left(1 - \frac{1}{2N}\right) (\text{var}(Z_{n-1}) + [\mathbb{E}(Z_{n-1})]^2) \end{aligned}$$

Por la proposición (4.3.1) podemos reescribir la ecuación (4.3.2) como

$$\text{var}(Z_n) = \mathbb{E}(Z_n^2) - [\mathbb{E}(Z_{n-1})]^2$$

Luego

$$\begin{aligned} \text{var}(Z_n) &= \mathbb{E}(Z_{n-1}) + \left(1 - \frac{1}{2N}\right) (\text{var}(Z_{n-1}) + [\mathbb{E}(Z_{n-1})]^2) - [\mathbb{E}(Z_{n-1})]^2 \\ &= \text{var}(Z_{n-1}) \left(1 - \frac{1}{2N}\right) + \mathbb{E}(Z_{n-1}) \left(1 - \frac{\mathbb{E}(Z_{n-1})}{2N}\right) \end{aligned}$$

Utilizando el corolario (4.3.1) se obtiene la siguiente relación de recurrencia:

$$\text{var}(Z_n) = \text{var}(Z_{n-1}) \left(1 - \frac{1}{2N}\right) + z_0 \left(1 - \frac{z_0}{2N}\right).$$

Resolviendo la ecuación de recurrencia se obtiene

$$\text{var}(Z_n) = z_0(1 - z_0) \left(1 - \left(1 - \frac{1}{2N}\right)^n\right)$$

como queríamos probar. ■

En la figura (4.2) se observa que algunas poblaciones alcanzan el estado 0 y otras alcanzan el estado $2N$ antes de $n = 2N$ mientras que otras no lo hacen. Partiendo de la observación (4.3.1), estamos interesados en calcular cuál es la probabilidad de que se alcance un estado absorbente y en caso de que se alcance, cuál es el tiempo esperado para que esto ocurra.

Para calcular la probabilidad de absorción y el tiempo esperado a la misma, definimos el tiempo de parada

$$\tau = \inf_{n \in \mathbb{N}} \{Z_n = 0 \vee Z_n = 2N\} \quad (4.3)$$

Proposición 4.3.3. Para todo $i \in E = \{0, 1, \dots, 2N\}$ se cumple que

$$\mathbf{P}(Z_\tau = 2N | Z_0 = i) = \frac{i}{2N}$$

Demostración. Como el espacio de estados $E = \{0, \dots, 2N\}$ es finito se tiene que $\mathbf{P}(\tau < \infty) = 1$. Por otra parte, como 0 y $2N$ son estados absorbentes se tiene que

$$\lim_{n \rightarrow \infty} Z_n = Z_\tau \quad (4.4)$$

Por la proposición (4.3.1) se obtiene que

$$i = \mathbb{E}(Z_n | Z_0 = i).$$

Podemos descomponer esta esperanza del siguiente modo

$$\mathbb{E}_i(Z_n) = \mathbb{E}(Z_n \mathbf{1}_{\{\tau \leq n\}} | Z_0 = i) + \mathbb{E}(Z_n \mathbf{1}_{\{\tau > n\}} | Z_0 = i)$$

Por (4.4), tomando $n \rightarrow \infty$ se obtiene

$$i = \mathbb{E}(Z_\tau | Z_0 = i)$$

Por otra parte, calculando la esperanza condicional directamente

$$i = \mathbb{E}(Z_\tau | Z_0 = i) = 0 \times \mathbf{P}(Z_\tau = 0 | Z_0 = i) + 2N \times \mathbf{P}(Z_\tau = 2N | Z_0 = i)$$

De donde se concluye que

$$\mathbf{P}(Z_\tau = 2N | Z_0 = i) = \frac{i}{2N}$$

■

Estamos interesados ahora en calcular cuál es el tiempo medio necesario para alcanzar los estados absorbentes 0 y $2N$. Notemos con k_i^A a $\mathbb{E}_i(\tau_A)$ con

$$\tau_A = \inf_{n \in \mathbb{N}} \{Z_n \in \{0, 2N\}\}$$

Por la proposición (5.1.4) se tiene que

$$k_i^A = \begin{cases} 0 & \text{si } i \in A \\ 1 + \sum_{j \notin A} p_{ij} k_j^A & \text{si } i \notin A \end{cases}$$

Este resultado, si bien correcto, no nos brinda demasiada información sobre el tiempo esperado. Para resolver este problema y dar una fórmula cerrada para este la esperanza del tiempo de absorción realizaremos una aproximación por difusión.

4.3.2. Aproximación por difusión

Para cada $N \in \mathbb{N}$ se considera $Z^{(N)} = \{Z_n^{(N)}\}_{n \in \mathbb{N}}$ el proceso de Wright-Fisher de parámetros N y p_0 que toma valores en $E_N = \{0, 1, \dots, 2N\}$. Sea

$$X_t^{(N)} = \frac{1}{2N} Z_{\lfloor 2Nt \rfloor}^{(N)}$$

un reescalado de tiempo y espacio del proceso original. Como puede observarse en la figura (4.3), para cada $N \in \mathbb{N}$ se obtiene un proceso estocástico con trayectorias en $D_{\tilde{E}_N}[0, +\infty)$, con $\tilde{E}_N = \{0, 1/2N, \dots, 1\}$.

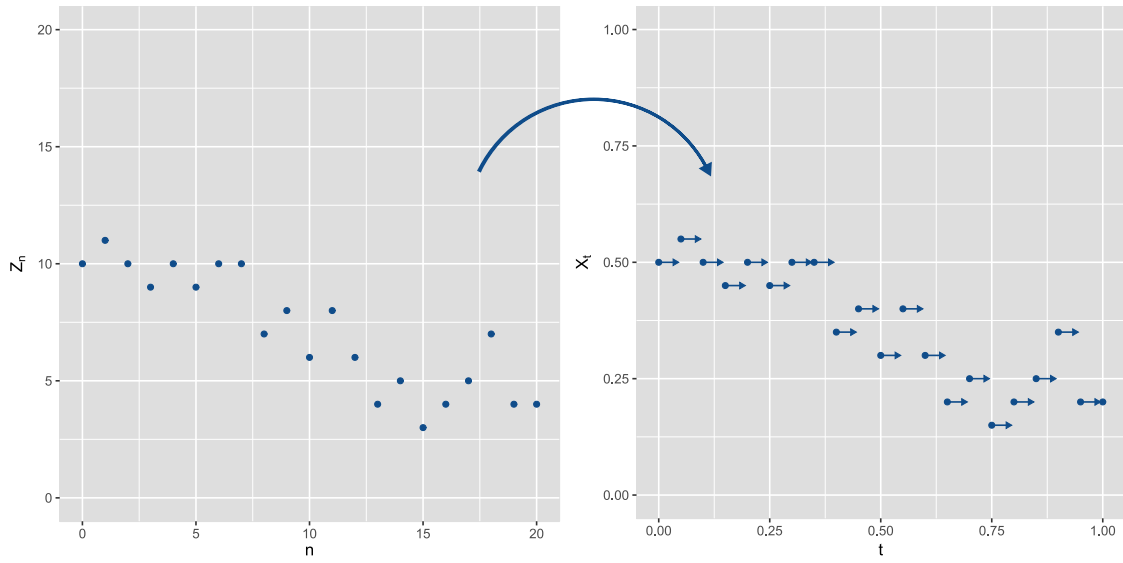


Figura 4.3: Reescalado de un proceso de Wright-Fisher en tiempo y espacio. A la izquierda, un proceso de Wright-Fisher $Z = \{Z_n\}$ de parámetros $N = z_0 = 10$. A la derecha, el proceso continuo $X = \{X_t\}_{t \geq 0}$ tal que $X_t = Z_{\lfloor 2Nt \rfloor}$.

Para cada $N \in \mathbb{N}$ tenemos un semigrupo asociado $S_N = \{S_{N,t}\}_{t \geq 0}$ tal que

$$(S_{N,t}(f))(x) = \mathbb{E}(f(X_t^{(N)}) | X_0^{(N)} = x)$$

para toda $f : \tilde{E}_N \rightarrow \mathbb{R}$ continua y toda $x \in \tilde{E}_N$. Observemos cómo actúa este semigrupo. Supongamos que $t < \frac{1}{2N}$. Entonces

$$(S_{N,t}(f)) = \mathbb{E}(f(X_t^{(N)}) | X_0^{(N)} = x) = \mathbb{E}(f(X_0^{(N)}) | X_0^{(N)} = x) = f(x).$$

Es decir, $S_{N,t} = I$, el operador identidad para todo $t < 1/2N$. Supongamos ahora que $t = \frac{1}{2N}$. Entonces

$$\begin{aligned} (S_{N,1/2N}f)(x) &= \mathbb{E}(f(X_{1/2N}^{(N)}) | X_0^{(N)} = x) \\ &= \mathbb{E}\left(f\left(\frac{1}{2N}Z_1^{(N)}\right) \middle| Z_0^{(N)} = 2Nx\right) \end{aligned}$$

Supongamos ahora que $\frac{1}{2N} \leq t < \frac{2}{2N}$. Podemos escribir entonces $t = t' + \frac{1}{2N}$, con $t' < \frac{1}{2N}$. Se tiene entonces que

$$\begin{aligned} (S_{N,t}f)(x) &= S_{N,t'+1/2N}(f)(x) \\ &= [S_{N,t'} \circ (S_{N,1/2N}(f))](x) && \text{(definición de semigrupo)} \\ &= (S_{N,1/2N}f)(x) && (S_{N,t'} \text{ es el operador identidad}) \end{aligned}$$

Supongamos que $t = 2/2N$. Luego

$$(S_{N,2/2N}(f))(x) = (S_{N,1/2N+1/2N}(f))(x)$$

$$= (S_{N,1/2N}^{(2)}(f))(x)$$

Ahora, para cualquier $\frac{k}{2N} \leq t < \frac{k+1}{2N}$ podemos escribir

$$\begin{aligned} (S_{N,t}f)(x) &= (S_{N,t'+1/2N+1/2N+\dots+1/2N}f)(x) \\ &= (S_{N,t'} \circ S_{N,1/2N}^{(k)})(f)(x) \\ &= (S_{N,1/2N}^{(k)})(f)(x) \end{aligned}$$

Sea $X = \{X_t\}_{0 \leq t \leq 1}$ el proceso de difusión cuyo semigrupo está dado por el generador infinitesimal

$$(\mathcal{G}f)(x) = \frac{1}{2}x(1-x)\frac{\partial^2}{\partial^2 x}f(x);$$

es decir, X será un proceso de Markov con trayectorias en $[0, 1]$ correspondientes al semigrupo generado por $\mathcal{G}$. Notaremos con $\mathcal{S} = \{S_t\}_{t \geq 0}$ a este semigrupo.

Proposición 4.3.4. Bajo las condiciones anteriores se tiene que

$$\lim_{N \rightarrow \infty} \sup_{0 \leq t \leq t_0} \sup_{x \in \tilde{E}_N} |(S_{N,t}f)(x) - (S_t f)(x)| = 0 \quad (4.5)$$

para toda $f \in C([0, 1])$ y todo $t_0 \geq 0$. Consecuentemente, si

$$\lim_{N \rightarrow \infty} X_0^{(N)} = X_0$$

en distribución, entonces $X^{(N)} \Rightarrow X$ débilmente en $D_E[0, +\infty)$.

Antes de realizar la prueba de esta proposición que será clave en el desarrollo de este trabajo, observemos qué es lo que esta prueba indica. Se tiene que con el reescalado descrito anteriormente los semigrupos asociados al proceso discreto y al proceso de difusión se hacen arbitrariamente chicos. Esto, sumado a la convergencia en distribución del proceso discreto al proceso de difusión en el valor inicial, nos dará la convergencia débil del proceso discreto al proceso de difusión. Esta proposición muestra de forma ejemplar el porqué de estudiar semigrupos de operadores y sus correspondientes generadores infinitesimales.

Demostración. La idea será utilizar la proposición (2.1.1) para probar (4.5). Para cada N consideramos $\mathcal{X}_N = D_{E_N}[0, +\infty)$ y consideraremos $\mathcal{X} = D_{[0,1]}[0, +\infty)$. El espacio de Skorokhod $D_I[0, +\infty)$ es un espacio de Banach equipado con la norma del supremo, $\|\cdot\|_\infty$. Utilizaremos que para cada $N \in \mathbb{N}$, las transformaciones lineales acotadas $\pi_N : \mathcal{X} \rightarrow \mathcal{X}_{t_N}$ que menciona la proposición (2.1.1) son todas iguales a la identidad. Consideraremos la sucesión $\{\varepsilon_N\}_{N \in \mathbb{N}}$ tal que $\varepsilon_N = \frac{1}{2N}$, para todo $N \in \mathbb{N}$. Notaremos con S_N a

$$(S_N f)(x) = \mathbb{E}_x(f(X_{1/2N})).$$

S_N es una contracción por la proposición (2.7). Por lo tanto, para probar (4.5), por la proposición (2.1.1) basta probar que

$$\lim_{N \rightarrow \infty} \sup_{x \in \tilde{E}_N} |2N(S_N - I)f(x) - (\mathcal{G}f)(x)| = 0$$

para toda $f \in C^2[0, 1]$. Asumiremos también, por simplicidad que f''' está acotada. Tenemos que

$$(S_N - I)f(x) = \mathbb{E}_x(f(X_{1/2N})) - f(x)$$

Realizando un desarrollo de Taylor de orden 2 de f alrededor de x tenemos que

$$\begin{aligned} (S_N - I)f(x) &= \mathbb{E}_x(f(X_{1/2N})) - f(x) \\ &= \mathbb{E}_x(f(x) + f'(x)(X_{1/2N} - x) + \frac{1}{2}f''(x)(X_{1/2N} - x)^2 + r(X_{1/2N})) \\ &\quad - f(x) \\ &= \mathbb{E}_x(f(x)) + f'(x)\mathbb{E}_x(X_{1/2N} - x) + \frac{1}{2}f''(x)\mathbb{E}_x((X_{1/2N} - x)^2) \\ &\quad + \mathbb{E}_x(r(X_{1/2N})) - f(x) \\ &= f'(x)\underbrace{\mathbb{E}_x(X_{1/2N} - x)}_{(1)} + \underbrace{\frac{1}{2}f''(x)\mathbb{E}_x((X_{1/2N} - x)^2) + \mathbb{E}_x(r(X_{1/2N}))}_{(2)}, \end{aligned}$$

donde $r(X_{1/2N})$ es el resto integral de Taylor, es decir:

$$r(X_{1/2N}) = \int_x^{X_{1/2N}} \frac{f^{(3)}(t)}{2}(X_{1/2N} - t)^2 dt$$

Analicemos el sumando (1):

$$\begin{aligned} \mathbb{E}_x(X_{1/2N} - x) &= \mathbb{E}(X_{1/2N}|X_0 = x) - x \\ &= \mathbb{E}\left(\frac{1}{2N}Z_1|Z_0 = 2Nx\right) - x \\ &= \frac{1}{2N}\mathbb{E}(Z_1|Z_0 = 2Nx) - x \\ &= \frac{1}{2N}2Nx - x \quad (Z_1|Z_0 \sim \text{Bin}(2N, Z_0/2N)) \\ &= 0 \end{aligned}$$

Para el sumando (2) se tiene

$$\begin{aligned} \mathbb{E}_x((X_{1/2N} - x)^2) &= \text{var}(X_{1/2N}|X_0 = x) \\ &= \frac{1}{(2N)^2}\text{var}(Z_1|Z_0 = 2Nx) \\ &= \frac{1}{(2N)^2}2Nx(1 - x) \\ &= \frac{x(1 - x)}{2N} \end{aligned}$$

Luego, por definición de $\mathcal{G}$

$$\begin{aligned}
|2N(S_N - I)f(x) - (\mathcal{G}f)(x)| &= \left| 2N \frac{1}{2} f''(x) \frac{x(1-x)}{2N} + 2N \mathbb{E}_x(r(X_{1/2N})) \right. \\
&\quad \left. - \frac{1}{2} x(1-x) f''(x) \right| \\
&= |2N \mathbb{E}_x(r(X_{1/2N}))| \xrightarrow{N \rightarrow \infty} 0
\end{aligned}$$

La prueba de que esto implica la convergencia débil en $D_{[0,1]}[0, +\infty)$ puede verse en [Ethier y Kurtz, 2005], p. 230. ■

A partir de la proposición anterior podemos obtener trayectorias de la aproximación por difusión sabiendo que el generador infinitesimal

$$(\mathcal{G}f)(x) = \frac{1}{2} x(1-x) \frac{\partial^2}{\partial^2 x} f(x)$$

corresponde a la solución de la ecuación diferencial estocástica $X = \{X_t\}$ tal que

$$\begin{cases} dX_t = \sqrt{X_t(1-X_t)} dW_t \\ X_0 = p_0 \in (0, 1) \end{cases}$$

Siguiendo la sección (3.3.1), utilizaremos el esquema de Milstein para simular soluciones de X . Dado un tiempo $T > 0$ y una precisión de n definimos $\Delta = T/n$ y obtenemos una grilla $\{\Delta i : i = 1, \dots, n\}$. El esquema de Milstein entonces será

$$\begin{cases} X_0 \leftarrow p_0 \\ X_{i+1} \leftarrow X_i + \sqrt{X_i(1-X_i)} (W_{\Delta(i+1)} - W_{\Delta i}) \\ \quad + \left(\frac{-2X_{i+1}}{4} \right) \left[(W_{\Delta(i+1)} - W_{\Delta i})^2 - \Delta \right] \end{cases}$$

con $W = \{W_t\}_{t \geq 0}$ un proceso de Wiener.

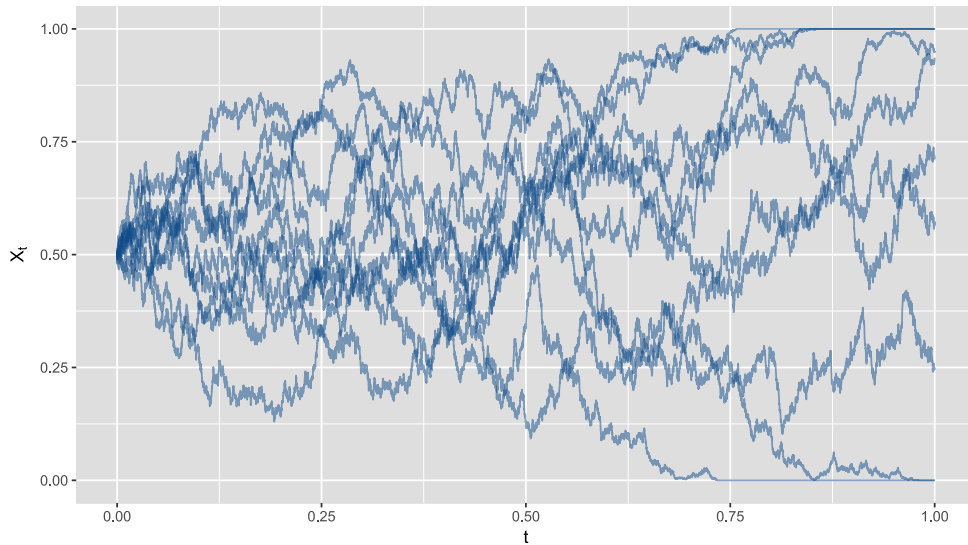


Figura 4.4: Trayectorias de la difusión límite de un proceso de Wright-Fisher en $[0, 1]$. Las trayectorias fueron obtenidas mediante el esquema de simulación de Milstein con una precisión para el proceso de Wiener de $n = 10^4$.

4.3.3. Tiempo de absorción

Sea $A = \{0, 2N\}$, el conjunto de los estados absorbentes en un proceso de Wright-Fisher neutral, $Z = \{Z_n\}$. Siguiendo el desarrollo presentado en la sección (4.3.2), para cada $N \in \mathbb{N}$ construimos el reescalado en tiempo y espacio $X^{(N)} = \{X_t^{(N)}\}_{t \geq 0}$ donde X

$$X_t^{(N)} = \frac{1}{2N} Z_{\lfloor 2Nt \rfloor}^{(N)}.$$

Para todo $N \in \mathbb{N}$ los estados 0 y 1 serán ahora absorbentes para el proceso $X^{(N)}$. Sea $X = \{X_t\}_{t \geq 0}$ la difusión límite hallada en la sección (4.3.2).

Se considera $I = [0, 1]$, definimos $\xi : D_I[0, +\infty) \rightarrow [0, +\infty)$ tal que

$$\xi(x) = \inf_{t \geq 0} \{x(t) \in \{0, 1\} \text{ o } x(t-) \in \{0, 1\}\}$$

con la convención de $\inf \emptyset = \infty$.

Definimos

$$\tau_N = \xi(X^{(N)}), \quad \tau = \xi(X)$$

Proposición 4.3.5. Si $X_0^{(N)} \xrightarrow{d} X_0$ entonces $(X^{(N)}, \tau_N) \Rightarrow (X, \tau)$ en $D_I[0, +\infty) \times [0, +\infty)$ y $\{\tau_N\}_{N \in \mathbb{N}}$ es uniformemente integrable.

Demostración. Ver [Ethier y Kurtz, 2005], página 419. ■

Corolario 4.3.2. Si $X_0^{(N)} \xrightarrow{d} X_0$ entonces

$$\lim_{N \rightarrow \infty} \mathbb{E}(\tau_N) = \mathbb{E}(\tau)$$

Demostración. Como se tiene que $X_0^{(N)} \xrightarrow{d} X_0$, por la proposición (4.3.5) se tiene que la sucesión $\{\tau_N\}_{N \in \mathbb{N}}$ es uniformemente integrable. Luego, por la proposición (5.3.1) se cumple lo pedido. ■

El corolario (4.3.2) implica que para estudiar el tiempo esperado para alcanzar un estado absorbente en el modelo de Wright-Fisher neutro, basta estudiar $\mathbb{E}(\tau)$.

Proposición 4.3.6. Sea $\mathcal{G}$ el operador diferencial definido en $C^2[0, 1]$ dado por

$$(\mathcal{G}f)(x) = \frac{1}{2}x(1-x)\frac{\partial^2}{\partial^2 x}f(x).$$

Sea g la única solución diferencial en $C[0, 1] \cap C^2([0, 1] \setminus \{0, 1\})$ que cumple

$$\begin{cases} \mathcal{G}(g) = -1 \\ g(1) = g(0) = 0 \end{cases}$$

Entonces, $\mathbb{E}_x(\tau) = g(x)$ para todo $x \in [0, 1]$. Consecuentemente,

$$\mathbb{E}_x(\tau) = -2[x \log(x) + (1-x) \log(1-x)]$$

Demostración. Por la observación (3.3.1) tenemos que $\mathbb{E}_x(\tau) = g(x)$ siendo g solución de

$$\begin{cases} \mathcal{G}(g) = -1 \\ g(1) = g(0) = 0 \end{cases}$$

Calculemos ahora la fórmula explícita para g . Sea g solución de la ecuación diferencial definida en (4.3.6). Para todo $x \in (0, 1)$, g cumple que

$$\frac{1}{2}x(1-x)g''(x) = -1$$

de donde se obtiene

$$g''(x) = -2 \left(\frac{1}{x(1-x)} \right)$$

El término de la derecha puede ser descompuesto en fracciones simples como

$$-2 \left(\frac{1}{x(1-x)} \right) = -2 \left(\frac{1}{x} + \frac{1}{1-x} \right)$$

Por lo tanto,

$$g''(x) = -2 \left(\frac{1}{x} + \frac{1}{1-x} \right)$$

Integrando a ambos lados se obtiene que

$$g'(x) = -2(\log(x) - \log(1-x)) + k_1$$

Volviendo a integrar a ambos lados se obtiene

$$g(x) = -2(x \log(x) - x + x + (1-x) \log(1-x)) + k_1x + k_2$$

Tomando $x \rightarrow 0$ y utilizando la condición de borde $g(0) = 0$ se obtiene que

$$\begin{aligned} \lim_{x \rightarrow 0} g(x) &= \lim_{x \rightarrow 0} -2(x \log(x) + (1-x) \log(1-x)) + k_1x + k_2 \\ &= k_2 = 0 \end{aligned}$$

Tomando ahora $x \rightarrow 1$ y utilizando la condición de borde $g(1) = 0$ se obtiene

$$\begin{aligned} \lim_{x \rightarrow 1} g(x) &= \lim_{x \rightarrow 1} -2(x \log(x) + (1-x) \log(1-x)) + k_1x \\ &= k_1 = 0 \end{aligned}$$

Por lo tanto:

$$g(x) = -2[x \log(x) + (1-x) \log(1-x)]$$

■

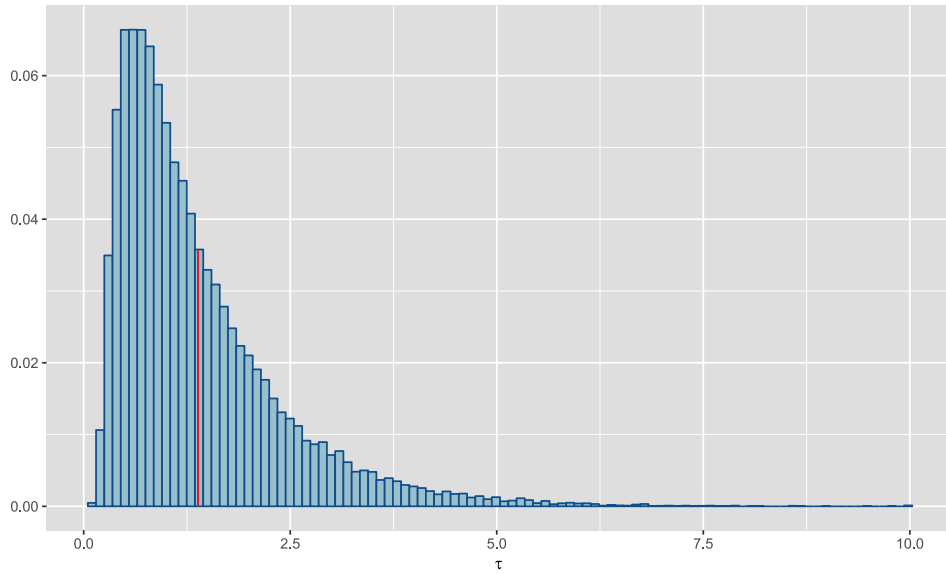


Figura 4.5: Densidad estimada $\tau = \inf\{t \geq 0 : X_t \in \{0, 1\}\}$ para el proceso de difusión $dX_t = \sqrt{X_t(1 - X_t)}dW_t$ y $X_0 = 0, 5$. En rojo, $\mathbb{E}(\tau) = -2\log(1/2)$. Para la estimación de la densidad se simuló 40,000 observaciones mediante el esquema de Milstein.

4.3.4. Distribución de X_t y distribución límite

Estamos interesados ahora en encontrar la distribución del proceso de difusión $X = \{X_t\}$ que aproxima un proceso de Wright-Fisher para cualquier $t \geq 0$ y en caso de que esta exista, encontrar la distribución límite del proceso.

En la figura (4.6) se observan trayectorias del proceso de Wright-Fisher para distintos $t \geq 0$ y partiendo del estado $X_0 = 0, 5$. Para $t = 0, 2$ la distribución de X_t es unimodal y simétrica y a medida que crece t y se acerca a 1 obtenemos dos modas en 0 y 1 correspondientes a los estados absorbentes del modelo. Para $t = 1$ se distingue que aproximadamente la mitad de las trayectorias han alcanzado uno de los dos estados absorbentes.

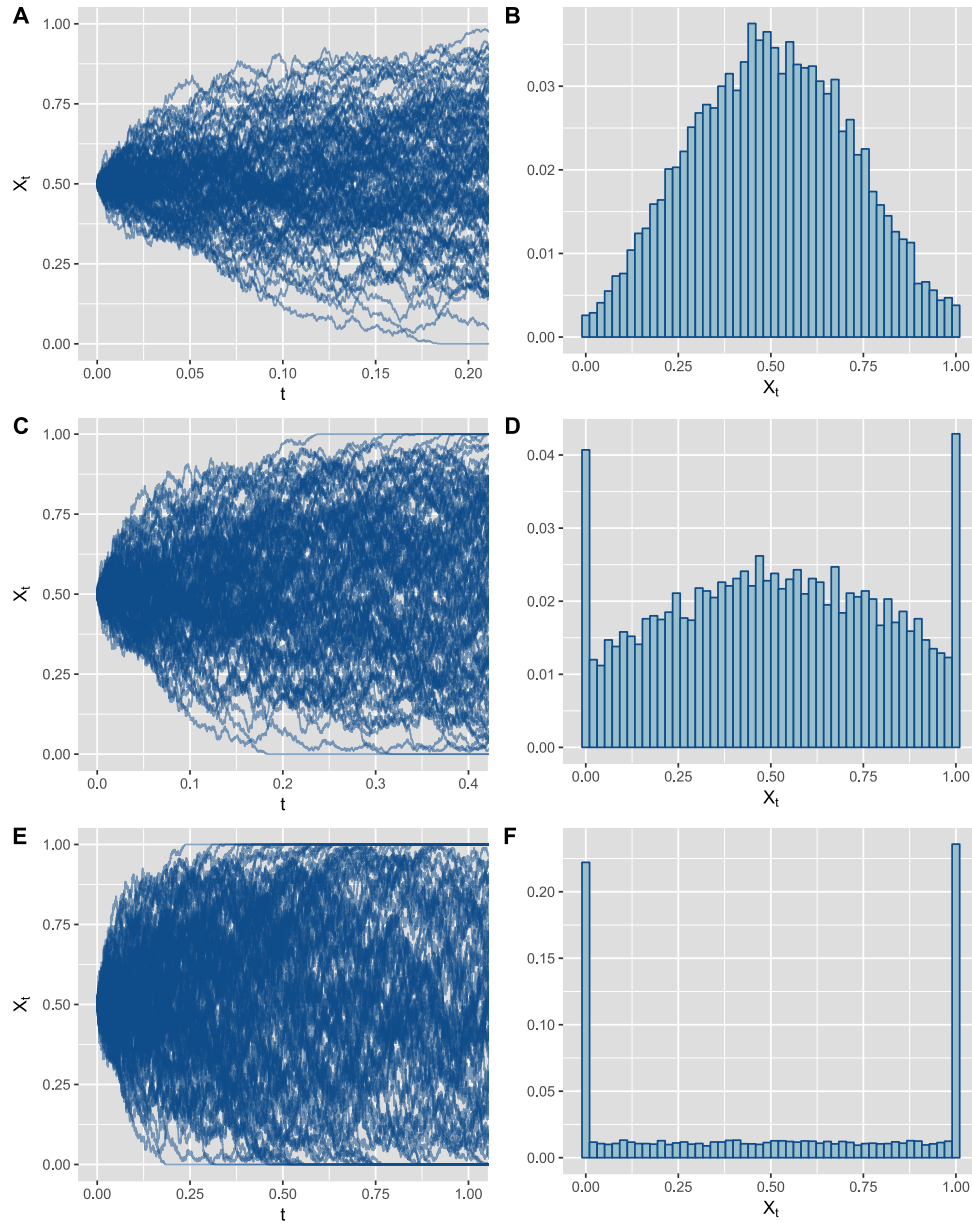


Figura 4.6: Densidad del proceso de difusión $X = \{X_t\}_{t \geq 0}$. (A), (C), (E): Trayectorias simuladas del proceso ($n = 100$) para $t \in [0; 0, 2]$, $t \in [0; 0, 4]$ y $t \in [0, 1]$, respectivamente. (B), (D), (F): Densidad estimada de X_t para $t = 0, 2, t = 0, 4$ y $t = 1$, respectivamente.

La distribución de X_t para valores de t cercanos a 0 es una distribución normal truncada (ver Apéndice (5.4)) como puede distinguirse en la figura (Apéndice 5.1). A medida que t crece y más trayectorias alcanzan los estados absorbentes, se pierde la distribución normal (test de Kolmogorov-Smirnov, $p < 0,01$).

De acuerdo a la figura (4.7) la distribución asintótica de X_t , dado que $X_0 = p \in [0, 1]$, es

$$\lim_{t \rightarrow \infty} X_t = (1 - p)\delta_0 + p\delta_1$$

donde δ_i es una delta de Dirac para $i = 0, 1$.

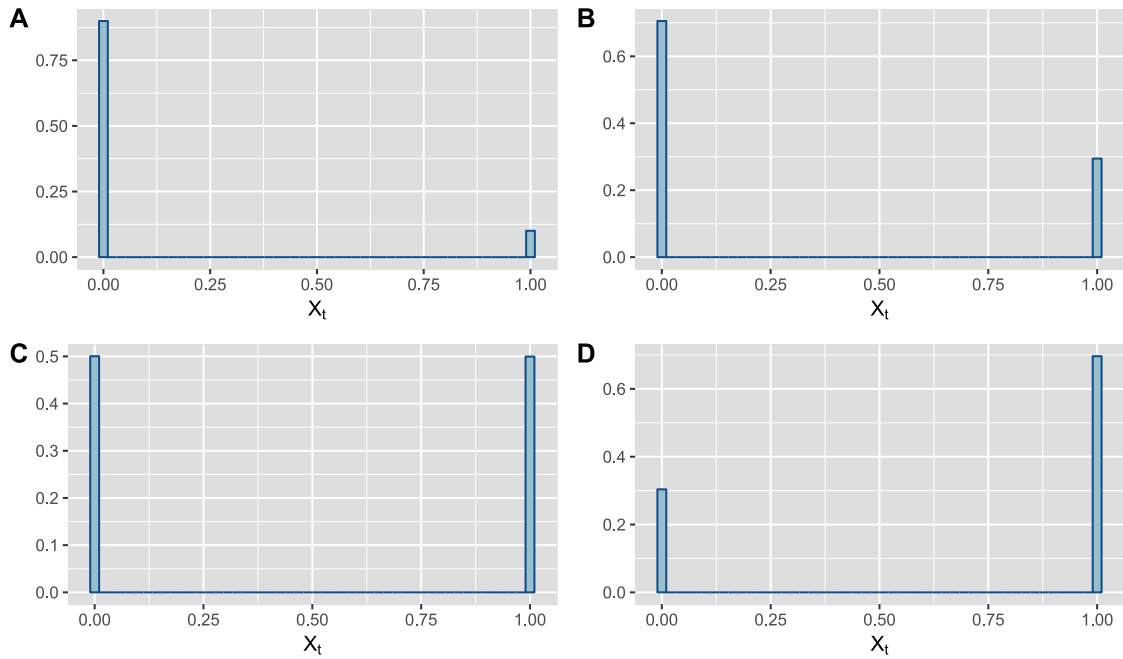


Figura 4.7: Distribución estimada de X_∞ . La distribución fue estimada en base a $k = 10^5$ simulaciones de X_t con $t = 100$. (A) $X_0 = 0,1$. (B) $X_0 = 0,3$. (C) $X_0 = 0,5$. (D) $X_0 = 0,7$.

4.4. Modelo de Wright-Fisher con selección natural

El modelo presentado a continuación es el resultado de relajar una más de las restricciones del modelo Hardy-Weinberg: al modelo de Wright-Fisher neutral se introducirá el concepto de eficacia relativa para estudiar la dinámica poblacional bajo un régimen de selección natural.

Al igual que en el modelo de Wright-Fisher neutral, estaremos interesados en estudiar un locus particular en una población diploide de tamaño N . En este locus existirán dos alelos en la población, A y a . De acuerdo al genotipo que posean los individuos sobrevivirán a la etapa adulta de forma diferencial; a la probabilidad de que un individuo sobreviva a la etapa adulta le llamaremos *viabilidad*. Pensaremos que los individuos AA tienen una viabilidad ω_{11} , los Aa y aA una viabilidad ω_{12} y los aa una viabilidad ω_{22} .

La manera en que modelaremos la evolución de una población de una generación a la siguiente será en dos etapas. En una primera etapa consideraremos a todo el *pool* génico y construiremos una población de adultos *infinita*. Supondremos que la probabilidad de obtener un alelo A será p y la probabilidad de obtener un alelo a será $q = 1 - p$. Al igual que en el modelo neutral supondremos que existe reproducción al azar. Por lo tanto obtendremos cigotos con las siguientes probabilidades:

$$P(AA) = p^2 \quad P(Aa) = 2pq \quad P(aA) = q^2$$

Sin embargo, no todos estos cigotos sobrevivirán a ser adultos y serán capaces de segregar gametos para construir la siguiente generación. Calculemos cuál será la probabilidad de observar un adulto del genotipo AA :

$$\begin{aligned}
P(AA|S) &= \frac{\overbrace{P(S|AA)}^{\omega_{11}} \overbrace{P(AA)}^{p^2}}{\underbrace{P(S|AA)}_{\omega_{11}} \underbrace{P(AA)}_{p^2} + \underbrace{P(S|AA)}_{\omega_{12}} \underbrace{P(AA)}_{2pq} + \underbrace{P(S|AA)}_{\omega_{22}} \underbrace{P(AA)}_{q^2}} \\
&= \frac{\omega_{11}p^2}{\omega_{11}p^2 + \omega_{12}2pq + \omega_{22}q^2}
\end{aligned}$$

Donde S es el evento "sobrevivir". A la cantidad $\omega_{11}p^2 + \omega_{12}2pq + \omega_{22}q^2$ suele denominarse *viabilidad promedio* o *fitness promedio* y se nota como $\bar{\omega}$. De la misma forma podemos calcular la probabilidad de observar un individuo adulto Aa y la de observar un individuo aa :

$$P(Aa|S) = \frac{\omega_{12}2pq}{\bar{\omega}}, \quad P(aa|S) = \frac{\omega_{22}q^2}{\bar{\omega}}$$

Calculemos las frecuencias del alelo A y a luego de que esta generación de adultos segreguen gametos. Para calcular la frecuencia A observamos que los individuos AA contribuyen con $p^2\omega_{11}/\bar{\omega}$ de los alelos, todos ellos de tipo A , y los individuos Aa contribuyen con $2pq\omega_{12}/\bar{\omega}$ alelos, la mitad de ellos de tipo A . Luego la frecuencia del alelo A , que notaremos como p_{sel} es

$$p_{sel} = \frac{p^2\omega_{11} + pq\omega_{12}}{\bar{\omega}}$$

De manera análoga se obtiene que la frecuencia del alelo a , que notaremos como q_{sel} , es

$$q_{sel} = \frac{pq\omega_{12} + q^2\omega_{22}}{\bar{\omega}}$$

La siguiente etapa del modelo es obtener una población de tamaño $2N$. Para esto obtendremos $2N$ realizaciones independientes de una distribución Bernoulli de parámetro p_{sel} . Podemos construir entonces el siguiente proceso que describe cómo es la dinámica poblacional:

Sea $Z_n : \Omega \rightarrow \{0, 1, \dots, 2N\}$ el número de alelos A en la generación n . Como el número de alelos A de la generación $n+1$ sólo depende del número de alelos de la generación n , $\{Z_n\}_{n \in \mathbb{N}}$ es una cadena de Márkov con espacio de estados $E = \{0, 1, \dots, 2N\}$. Las probabilidades de transición son

$$p_{ij} = P(Z_{n+1} = j | Z_n = i) = \binom{2N}{j} (p_{sel})^j (1 - p_{sel})^{2N-j}$$

con

$$p_{sel} = \frac{\left(\frac{i}{2N}\right)^2 \omega_{11} + \left(\frac{i}{2N}\right) \left(1 - \frac{i}{2N}\right) \omega_{12}}{\frac{i}{2N}^2 \omega_{11} + 2 \left(\frac{i}{2N}\right) \left(1 - \frac{i}{2N}\right) \omega_{12} + \left(1 - \frac{i}{2N}\right)^2 \omega_{22}}$$

Definición 4.4.1 (Proceso de Wright-Fisher con selección). *Diremos que una cadena de Markov en tiempo discreto $Z = \{Z_n\}_{n \in \mathbb{N}}$ es un **proceso de Wright-Fisher con selección** de parámetros N , z_0 y $\omega = (\omega_{11}, \omega_{12}, \omega_{22})$ sobre $E = \{0, 1, 2, \dots, 2N\}$ Si*

- i) la distribución inicial es $\mathbf{P}(Z_0 = z_0) = 1$ y,
 ii) la matriz de transición $\mathbb{P} = ((p_{ij}))_{i,j=0,1,\dots,2N}$ es tal que

$$p_{ij} = \mathbf{P}(Z_{n+1} = j | Z_n = i) = \binom{2N}{j} (p_{sel})^j (1 - p_{sel})^{2N-j}$$

con

$$p_{sel} = \frac{\left(\frac{i}{2N}\right)^2 \omega_{11} + \left(\frac{i}{2N}\right) \left(1 - \frac{i}{2N}\right) \omega_{12}}{\frac{i}{2N}^2 \omega_{11} + 2 \left(\frac{i}{2N}\right) \left(1 - \frac{i}{2N}\right) \omega_{12} + \left(1 - \frac{i}{2N}\right)^2 \omega_{22}} \quad (4.6)$$

Observación 4.4.1. Es usual que las viabilidades $\omega_{11}, \omega_{12}, \omega_{22}$ sean consideradas como viabilidades *relativas* en vez de *absolutas*. Esto puede ser útil para determinar cuán viable es un genotipo con respecto a otro.

Supongamos por ejemplo que $\omega_{11} = 0,9$, $\omega_{12} = 1$ y $\omega_{22} = 0,7$. Por ejemplo, tomando el alelo AA como el alelo de referencia, podemos calcular la viabilidad del alelo Aa relativa al alelo AA como

$$\frac{\omega_{12}}{\omega_{11}} = \frac{1}{0,9} \approx 1,1$$

y la viabilidad del alelo aa relativa al alelo AA como

$$\frac{\omega_{22}}{\omega_{11}} = \frac{0,7}{0,9} \approx 0,78.$$

Nótese que la ecuación (4.6) no se modifica si en lugar de las viabilidades absolutas se las toma como relativas.

Observación 4.4.2. El supuesto de que las viabilidades de un genotipo son constantes en el tiempo puede ser un supuesto fuerte para modelar la evolución de algunas poblaciones. Una exposición clara y rigurosa de cómo incluir selección fluctuante en el tiempo al modelo de Wright-Fisher puede encontrarse en [Weber, 1986].

4.4.1. Aproximación por difusión

Aproximaremos el proceso de Wright-Fisher con selección por un proceso de difusión límite de forma análoga a lo obtenido en la sección (4.3.2):

Se considera un vector de viabilidades (absolutas o relativas) $\omega = (\omega_{11}, \omega_{12}, \omega_{22})$. Para cada $N \in \mathbb{N}$ se construye el proceso de Wright-Fisher con selección, $Z^{(N)} = \{Z_n^{(N)}\}_{n \in \mathbb{N}}$, de parámetros N , ω y p_0 . Se realizará la siguiente reparametrización de las viabilidades,

$$\omega_{ij} = 1 + \frac{\sigma_{ij}}{2N},$$

con $i, j = 1, 2$ y $\sigma_{ij} \in \mathbb{R}$. Para cada $N \in \mathbb{N}$ encajaremos el proceso $Z^{(N)}$ en un proceso continuo realizando el reescalado $X^{(N)} = \{X_t^{(N)}\}_{t \geq 0}$ tal que

$$X_t^{(N)} = \frac{1}{2N} Z_{\lfloor 2Nt \rfloor}^{(N)}$$

un reescalado de tiempo y espacio del proceso original. Para este proceso el operador de transición $\mathcal{S}_N = \{S_{N,t}\}_{t \geq 0}$ como

$$(S_{N,t}f)(x) = \mathbb{E}(f(X_t^{(N)}) | X_0 = x),$$

para toda $f : \{0, \dots, 2N\} \rightarrow \mathbb{R}$ continua.

Consideramos el operador diferencial $\mathcal{G} : C^2[0, +\infty) \rightarrow C^2[0, +\infty)$ tal que

$$(\mathcal{G}f)(x) = a(x) \frac{\partial}{\partial x} f'(x) + \frac{1}{2} x(1-x) \frac{\partial^2}{\partial^2 x} f''(x)$$

con

$$a(x) = x^2(1-x)\sigma_{11} + x(1-x)(1-2x)\sigma_{12} - x(1-x)^2\sigma_{22}$$

Proposición 4.4.1. Bajo las condiciones anteriores se tiene que

$$\lim_{N \rightarrow \infty} \sup_{0 \leq t \leq t_0} \sup_{x \in \tilde{E}_N} |(S_{N,t}f)(x) - (S_t f)(x)| = 0 \quad (4.7)$$

para toda $f \in C([0, 1])$ y todo $t_0 \geq 0$. Consecuentemente, si

$$\lim_{N \rightarrow \infty} X_0^{(N)} = X_0$$

en distribución, entonces $X^{(N)} \Rightarrow X$ débilmente en $D_E[0, +\infty)$.

Demostración. La prueba será análoga a la realizada para el caso neutral dado por la proposición (4.3.4). Para cada N definimos S_N como

$$(S_N f)(x) = \mathbb{E}_x(f(X_{1/2N})).$$

Luego, por la proposición (2.1.1), para probar (4.7) basta probar que

$$\lim_{N \rightarrow \infty} \sup_{x \in \tilde{E}_N} |2N(S_N - I)f(x) - (\mathcal{G}f)(x)| = 0$$

para toda $f \in C^2([0, 1])$. Se tiene que

$$(S_N - I)f(x) = \mathbb{E}(f(X_{1/2N}) | X_0 = x) - f(x)$$

Realizando un desarrollo de Taylor de orden 2 de f centrado en x se tiene que

$$\begin{aligned} (S_N - I)f(x) &= \mathbb{E}_x(f(X_{1/2N})) - f(x) \\ &= \mathbb{E}_x \left(f(x) + f'(x)(X_{1/2N} - x) + f''(x)(X_{1/2N} - x)^2 + r(X_{1/2N}) \right) \\ &\quad - f(x) \\ &= \mathbb{E}_x(f(x)) + f'(x)\mathbb{E}_x(X_{1/2N} - x) + f''(x)\mathbb{E}_x((X_{1/2N} - x)^2) \\ &\quad + \mathbb{E}_x(r(X_{1/2N})) - f(x) \end{aligned}$$

$$\begin{aligned}
&= f'(x) \underbrace{\mathbb{E}_x(X_{1/2N} - x)}_{(1)} + \frac{f''(x)}{2} \underbrace{\mathbb{E}_x((X_{1/2N} - x)^2)}_{(2)} \\
&\quad + \underbrace{\mathbb{E}_x(r(X_{1/2N}))}_{(3)}
\end{aligned}$$

Analizaremos la expresión en (1) para entender cómo afecta la reparametrización de ω_{ij} . Las expresiones en (2) y (3) serán análogas a la (1) y a lo expuesto en la proposición (4.3.4).

$$\begin{aligned}
\mathbb{E}_x(X_{1/2N} - x) &= \mathbb{E}(X_{1/2N} | X_0 = x) - x \\
&= \frac{1}{2N} \mathbb{E}(Z_1 | Z_0 = 2Nx) - x \\
&= \frac{1}{2N} \left(2N \frac{x^2\omega_{11} + x(1-x)\omega_{12}}{x^2\omega_{11} + 2x(1-x)\omega_{12} + (1-x)^2\omega_{22}} \right) - x \\
&= \frac{x^2\omega_{11} + x(1-x)\omega_{12} - x^3\omega_{11} - 2x^2(1-x)\omega_{12} - x(1-x)^2\omega_{22}}{x^2\omega_{11} + 2x(1-x)\omega_{12} + (1-x)^2\omega_{22}} \\
&= \frac{x^2(1-x)\omega_{11} + x(1-x)(1-2x)\omega_{12} - x(1-x)^2\omega_{22}}{x^2\omega_{11} + 2x(1-x)\omega_{12} + (1-x)^2\omega_{22}}
\end{aligned}$$

Escribiendo la reparametrización ω_{ij} :

$$\omega_{ij} = 1 + \frac{\sigma_{ij}}{2N}$$

para todo $i, j = 1, 2$, tenemos que

$$\begin{aligned}
x^2(1-x)\omega_{11} + x(1-x)(1-2x)\omega_{12} - x(1-x)^2\omega_{22} &= \\
&\quad \overbrace{x^2(1-x) + x(1-x)(1-2x) - x(1-x)^2}^{=0} \\
&\quad + \frac{x^2(1-x)\sigma_{11} + x(1-x)(1-2x)\sigma_{12} - x(1-x)^2\sigma_{22}}{2N}
\end{aligned}$$

y

$$\begin{aligned}
x^2\omega_{11} + 2x(1-x)\omega_{12} + (1-x)^2\omega_{22} &= \overbrace{x^2 + 2x(1-x) + (1-x)^2}^{=1} \\
&\quad + \frac{x^2 + 2x(1-x) + (1-x)^2}{2N} \\
&= \frac{2N + x^2 + 2x(1-x) + (1-x)^2}{2N}
\end{aligned}$$

Por lo tanto

$$2N\mathbb{E}(X_{1/2N} - x | X_0 = x) = 2N \frac{x^2(1-x)\sigma_{11} + x(1-x)(1-2x)\sigma_{12} + x(1-x)^2\sigma_{22}}{2N + x^2 + 2x(1-x) + (1-x)^2}$$

que tiende con $N \rightarrow \infty$ a

$$a(x) = x^2(1-x)\sigma_{11} + x(1-x)(1-2x)\sigma_{12} + x(1-x)^2\sigma_{22},$$

La proposición se concluye de forma análoga a (4.3.4). ■

A partir de la proposición anterior podemos obtener trayectorias de la aproximación por difusión sabiendo que el generador infinitesimal

$$(\mathcal{G}f)(x) = a(x)\frac{\partial}{\partial x}f'(x) + \frac{1}{2}x(1-x)\frac{\partial^2}{\partial^2 x}f''(x)$$

con

$$a(x) = x^2(1-x)\sigma_{11} + x(1-x)(1-2x)\sigma_{12} - x(1-x)^2\sigma_{22}$$

corresponde a la solución de la ecuación diferencial estocástica $X = \{X_t\}$ tal que

$$\begin{cases} dX_t = (X_t^2(1-X_t)\sigma_{11} + X_t(1-X_t)(1-2X_t)\sigma_{12} - X_t(1-X_t)^2\sigma_{22})dt \\ \quad + \sqrt{X_t(1-X_t)} dW_t \\ X_0 = p_0 \in (0, 1) \end{cases} \quad (4.8)$$

Siguiendo la sección (3.3.1), utilizaremos el esquema de Milstein para simular soluciones de X . Dado un tiempo $T > 0$ y una precisión de n definimos $\Delta = T/n$ y obtenemos una grilla $\{\Delta i : i = 1, \dots, n\}$. El esquema de Milstein entonces será

$$\begin{cases} X_0 \leftarrow p_0 \\ X_{i+1} \leftarrow (X_i^2(1-X_i)\sigma_{11} + X_i(1-X_i)(1-2X_i)\sigma_{12} - X_i(1-X_i)^2\sigma_{22})\Delta \\ \quad + \sqrt{X_i(1-X_i)} (W_{\Delta(i+1)} - W_{\Delta i}) \\ \quad + \left(\frac{-2X_i+1}{4}\right) \left[(W_{\Delta(i+1)} - W_{\Delta i})^2 - \Delta\right] \end{cases}$$

con $W = \{W_t\}_{t \geq 0}$ un proceso de Wiener.

Dado el fitness de un alelo ω_{ij} y un tamaño poblacional podemos obtener mediante la reparametrización

$$\omega_{ij} = 1 + \frac{\sigma_{ij}}{2N}$$

el coeficiente σ_{ij} de la difusión límite como

$$\sigma_{ij} = 2N(\omega_{ij} - 1).$$

Así, por ejemplo, si la viabilidad del alelo AA es de 0,8 podemos obtener que

$$\sigma_{11} = -2N0,2$$

Tipos de selección

Analizaremos casos especiales de selección modificando los coeficientes de selección σ_{ij} y veremos cuáles son las implicancias de estas modificaciones en la dinámica de las trayectorias de la difusión límite $X = \{X_t\}_{t \geq 0}$.

- **Selección balanceadora.** En este caso, el fitness del genotipo Aa es mayor que el de los genotipos homocigotas, AA y aa . Expresando las eficacias de los genotipos AA y aa relativas a las del genotipo Aa se obtiene que las eficacias son

$$\begin{array}{ccc} \omega_{11} & \omega_{12} & \omega_{22} \\ 1 - s_1 & 1 & 1 - s_2 \end{array}$$

con $s_1, s_2 \in [0, 1]$. Por lo tanto, la difusión límite en (4.8) es, en este caso,

$$\begin{cases} dX_t = [X_t(1 - X_t)(-X_t 2N s_1 + (1 - X_t) 2N s_2)]dt \\ \quad + \sqrt{X_t(1 - X_t)} dW_t \\ X_0 = p_0 \in (0, 1) \end{cases} \quad (4.9)$$

- **Selección disruptiva.** En este caso el fitness del genotipo heterocigoto Aa es menor al de los genotipos homocigotos AA y aa . Expresando las eficacias de los genotipos AA y aa relativas a la de Aa se obtiene que las eficacias son

$$\begin{array}{ccc} \omega_{11} & \omega_{12} & \omega_{22} \\ 1 + s_1 & 1 & 1 + s_2 \end{array}$$

con $s_1, s_2 \in [0, 1]$. Por lo tanto, la difusión límite en (4.8) es, en este caso,

$$\begin{cases} dX_t = [X_t(1 - X_t)(X_t 2N s_1 - (1 - X_t) 2N s_2)]dt \\ \quad + \sqrt{X_t(1 - X_t)} dW_t \\ X_0 = p_0 \in (0, 1) \end{cases} \quad (4.10)$$

- **El alelo A es dominante.** Como consecuencia de la dominancia del alelo A la eficacia de los genotipos AA y Aa es igual. Expresando las eficacias relativas al genotipo AA se obtienen las eficacias

$$\begin{array}{ccc} \omega_{11} & \omega_{12} & \omega_{22} \\ 1 & 1 & 1 - s \end{array}$$

con $s \in [0, 1]$. Por lo tanto, la difusión límite en (4.8) es, en este caso,

$$\begin{cases} dX_t = X_t(1 - X_t)^2 2N s dt + \sqrt{X_t(1 - X_t)} dW_t \\ X_0 = p_0 \in (0, 1) \end{cases} \quad (4.11)$$

$$\begin{array}{ccc} \omega_{11} & \omega_{12} & \omega_{22} \\ 1-s & 1 & 1 \end{array}$$

- **El alelo a es dominante.** Como consecuencia de la dominancia del alelo a la eficacia de los genotipos Aa y aa es igual. Expresando las eficacias relativas al genotipo aa se obtienen las eficacias

con $s \in [0, 1]$. Por lo tanto, la difusión límite en (4.8) es, en este caso,

$$\begin{cases} dX_t = X_t^2(1 - X_t)2Ns dt + \sqrt{X_t(1 - X_t)} dW_t \\ X_0 = p_0 \in (0, 1) \end{cases} \quad (4.12)$$

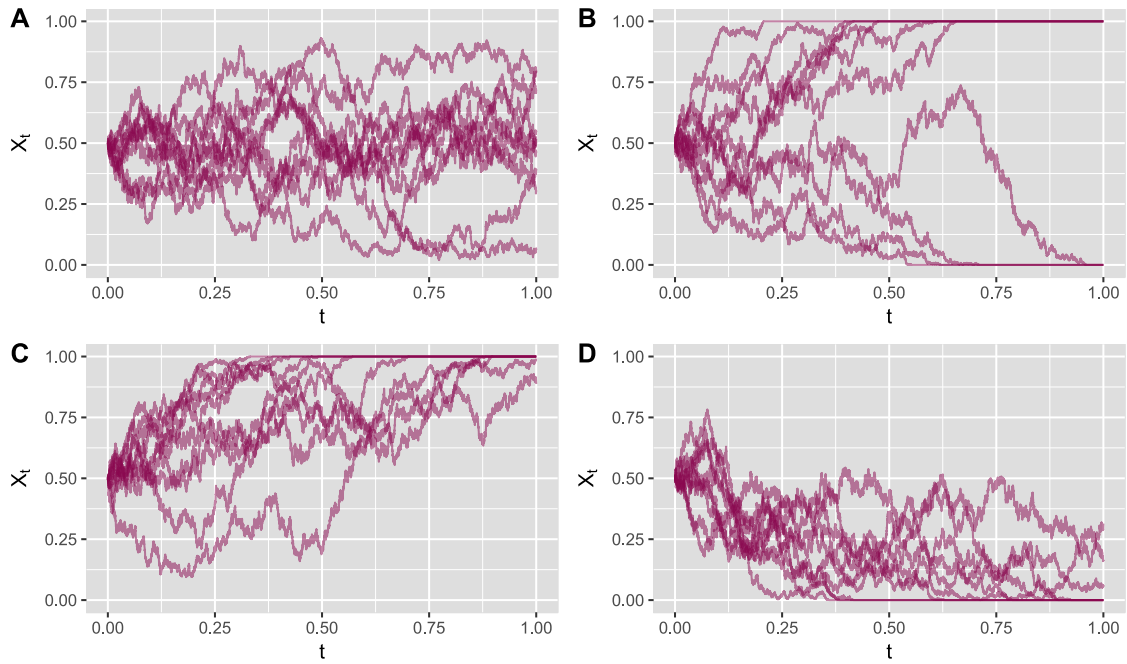


Figura 4.8: Trayectorias de la difusión $X = \{X_t\}_{t \geq 0}$ correspondiente al proceso de Wright-Fisher con selección. (A) Selección balanceadora. Las trayectorias corresponden a la ecuación (4.9) con $s_1 = s_2 = 0,1$ y $N = 100$. (B) Selección disruptiva. Las trayectorias corresponden a la ecuación (4.10) con $s_1 = s_2 = 0,1$ y $N = 100$. (C) A dominante. Las trayectorias corresponden a la ecuación (4.11) con $s = 0,1$ y $N = 100$. (D) a dominante. Las trayectorias corresponden a la ecuación (4.12) con $s = 0,1$ y $N = 100$.

4.4.2. Tiempo de absorción

Sea $X = \{X_t\}_{t \geq 0}$ la difusión límite definida por la ecuación (4.8). Definimos el tiempo de parada

$$\tau = \inf_{t \geq 0} \{X_t \in \{0, 1\}\}.$$

Estamos interesados en calcular $\mathbb{E}_x(\tau)$. Esto lo dará la siguiente proposición.

Proposición 4.4.2. Sea $\mathcal{G} : C^2[0, 1] \rightarrow C^2[0, 1]$ dado por

$$(\mathcal{G}f)(x) = a(x)\frac{\partial}{\partial x}f'(x) + \frac{1}{2}x(1-x)\frac{\partial^2}{\partial^2 x}f''(x)$$

con

$$a(x) = x^2(1-x)\sigma_{11} + x(1-x)(1-2x)\sigma_{12} - x(1-x)^2\sigma_{22}.$$

Sea g la única solución diferencial en $C[0, 1] \cap C^2([0, 1] \setminus \{0, 1\})$ que cumple

$$\begin{cases} \mathcal{G}(g) = -1 \\ g(1) = g(0) = 0 \end{cases}$$

Entonces, $\mathbb{E}_x(\tau) = g(x)$ para todo $x \in [0, 1]$.

Demostración. La prueba es consecuencia inmediata de la observación (4.3.6).

Se consideran $a, b : [0, 1] \rightarrow \mathbb{R}$ tal que

$$a(x) = x^2(1-x)\sigma_{11} + x(1-x)(1-2x)\sigma_{12} - x(1-x)^2\sigma_{22}$$

y

$$b(x) = \frac{1}{2}x(1-x)$$

Sea g solución de la ecuación diferencial (4.4.2). g es, por lo tanto, solución de la ecuación equivalente

$$\begin{cases} g''(x) + \frac{a(x)}{b(x)}g'(x) = \frac{-1}{b(x)} \\ g(1) = g(0) = 0 \end{cases} \quad (4.13)$$

Definimos $u(x) = g'(x)$. La ecuación (4.13) puede ser escrita como

$$\begin{cases} u'(x) + \frac{a(x)}{b(x)}u(x) = \frac{-1}{b(x)} \\ g(1) = g(0) = 0 \end{cases} \quad (4.14)$$

Para resolver esta ecuación definiremos el factor integrando $I(x)$ tal que

$$I(x) = e^{\int \frac{a(x)}{b(x)} dx}$$

donde con $\int \frac{a(x)}{b(x)} dx$ nos referimos a una primitiva de $\frac{a(x)}{b(x)}$. Nótese que $I(x)$ cumple que

$$(I(x)u(x))' = I(x)\frac{a(x)}{b(x)} + I(x)u'(x)$$

Multiplicando por $I(x)$ a la ecuación (4.14) se obtiene

$$I(x) \left(u'(x) + \frac{a(x)}{b(x)}u(x) \right) = I(x)\frac{-1}{b(x)}$$

Integrando a ambos lados de la ecuación observamos que

$$\int I(x) \left(u'(x) + \frac{a(x)}{b(x)}u(x) \right) dx = I(x)u(x) = \int I(x)\frac{-1}{b(x)} dx.$$

Se concluye que

$$u(x) = \frac{1}{I(x)} \int I(x) \frac{-1}{b(x)} dx$$

Para hallar la fórmula explícita de g basta tomar una primitiva de u y evaluarla en las condiciones de borde. ■

El resultado obtenido en la proposición (4.4.2) es, en buena medida, un resultado decepcionante, comparado con el resultado análogo obtenido para el proceso neutral en la proposición (4.3.6). Esto se debe a la estructura más compleja del generador infinitesimal de la difusión límite del proceso de Wright-Fisher con selección. Para obtener una idea de la esperanza de τ utilizaremos simulaciones de la difusión para varios esquemas de selección.

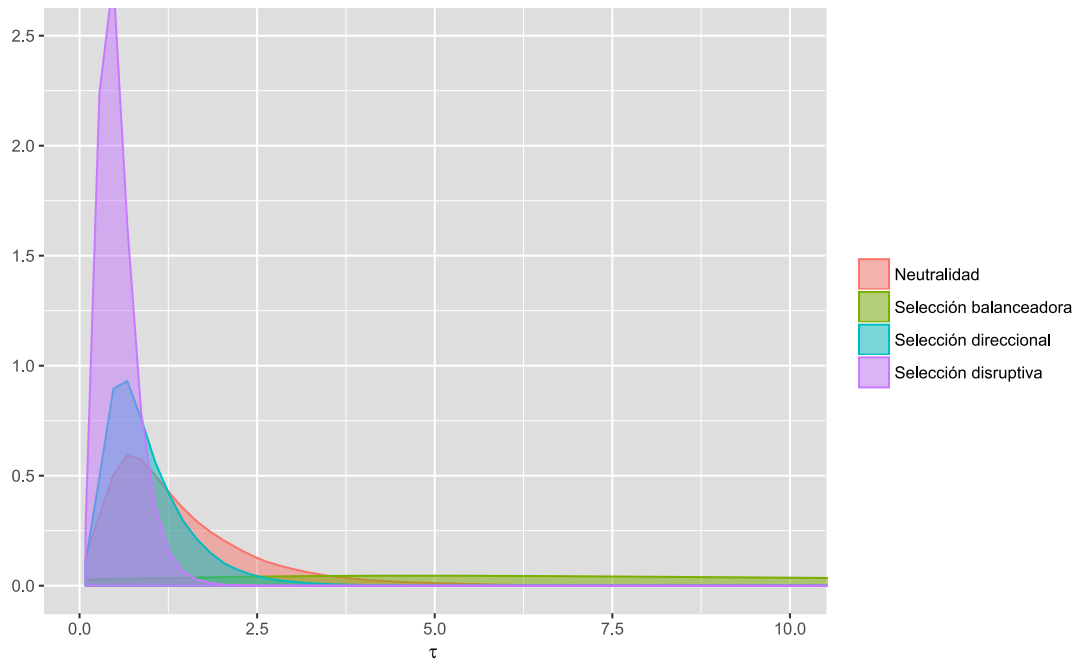


Figura 4.9: Densidad estimada de τ para diversos esquemas de selección. Los valores de τ fueron obtenidos simulando trayectorias de la difusión límite del proceso de Wright-Fisher con selección, $X = \{X_t\}_{t \geq 0}$, mediante el esquema de Milstein con una precisión de $n = 10^6$. Neutralidad: corresponde a la difusión límite del proceso de Wright-Fisher neutro. Selección balanceadora: corresponde a la difusión que cumple la ecuación (4.9) con $s_1 = s_2 = -0,1$ y $N = 100$. Selección disruptiva: corresponde a la difusión que cumple la ecuación (4.10) con $s_1 = s_2 = 0,1$ y $N = 100$. Selección direccional: corresponde a la difusión que cumple la ecuación (4.11) con $s = 0,1$ y $N = 100$.

En la figura (4.9) se muestra la densidad de τ para distintos la difusión límite del proceso de Wright-Fisher siguiendo distintos esquemas de selección (ver nota a pie de figura (4.9) para más información sobre los parámetros de selección). Las densidades fueron calculadas simulando trayectorias hasta alcanzar cualquiera de los dos estados absorbentes y partiendo del estado $X_0 = 0,5$. Los esquemas de selección disruptiva y direccional arrojaron valores menores de $\mathbb{E}(\tau)$ que el obtenido bajo neutralidad ($\mathbb{E}(\tau) = 0,54$ para selección disruptiva y $\mathbb{E}(\tau) = 0,94$ para selección direccional). La fuerza que repele a las

trayectorias de los estados absorbentes en el esquema de selección balanceadora dio como resultado que $\widehat{\mathbb{E}(\tau)}$ fuera considerablemente mayor comparado con el resultado obtenido bajo neutralidad ($\widehat{\mathbb{E}(\tau)} = 15,23$).

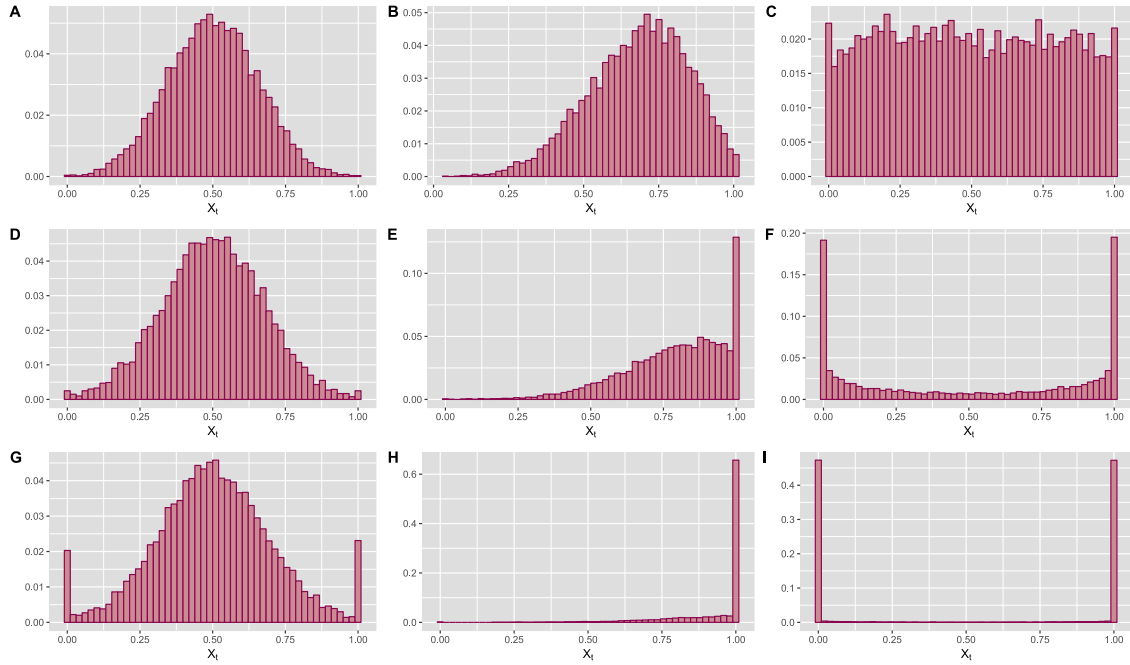


Figura 4.10: Densidad estimada de X_t para diversos esquemas de selección. Los valores de X_t fueron obtenidos simulando trayectorias de la difusión límite del proceso de Wright-Fisher con selección, $X = \{X_t\}_{t \geq 0}$, mediante el esquema de Milstein con una precisión de $n = 1 \times 10^6$. La fila superior corresponde a la densidad de X_t para $t = 0, 2$; la fila central corresponde a la densidad de X_t para $t = 0, 4$; la fila inferior corresponde a la densidad de X_t para $t = 1$. Figuras A, D y G: corresponden a la difusión que cumple la ecuación (4.9) con $s_1 = s_2 = -0,1$ y $N = 100$. Figuras B, E y H: corresponden a la difusión que cumple la ecuación (4.11) con $s = 0,1$ y $N = 100$. Figuras C, F y I: corresponden a la difusión que cumple la ecuación (4.10) con $s_1 = s_2 = 0,1$ y $N = 100$.

4.4.3. Distribución de X_t y distribución límite

Podemos hallar, mediante simulación, la distribución de X_t para cada $t \geq 0$. En la figura (4.10) se observan estimaciones de la densidad de X_t para $t = 0, 2$, $t = 0, 4$, $t = 1$ y para varios esquemas de selección partiendo del estado inicial $X_0 = 0,5$ y parámetros específicos (ver pie de figura para más detalles). La probabilidad estimada de alcanzar el estado 0 bajo selección direccional en contra del genotipo aa (figuras C, F, I) es de aproximadamente 0,001 en tiempo $t = 1$. Por el contrario, en base a $n = 10000$ trayectorias se estima que $\mathbf{P}(X_t = 1) \approx 0,65$. Resultados análogos se alcanzan si se utiliza un esquema de selección direccional en contra del AA . Las selecciones balanceadora (figuras A, D, G) y disruptiva (figuras B, E, H) llevan a la a las trayectorias de X a ambos estados absorbentes con probabilidades equivalentes. La diferencia entre la selección disruptiva y la selección balanceadora se encuentra en la velocidad con la que se alcanzan los estados absorbentes: mientras que la probabilidad de alcanzar un estado absorbente es aproximadamente 0,94 para un regimen de selección disruptiva, la probabilidad es aproximadamente 0,04 para

selección balanceadora.

4.5. Modelo de Wright-Fisher con selección natural y mutación

El modelo presentado a continuación surge quitando uno más de los supuestos del modelo de Hardy-Weinberg: estudiaremos cuál es el efecto de agregar una probabilidad no nula de que en cada generación exista mutación.

Se considera una población diploide de tamaño N . Analizaremos la dinámica de un locus autosómico que presenta dos alelos en la población, A y a . Consideraremos que existe un vector $\omega = (\omega_{11}, \omega_{12}, \omega_{22})$ correspondientes a las viabilidades de los genotipos AA , Aa y aa , respectivamente. A diferencia de lo expuesto en la sección (4.4), supondremos que existe una probabilidad $\mu_1 \geq 0$ de que un alelo A mute a un alelo a y una probabilidad $\mu_2 \geq 0$ de que un alelo a mute a un alelo A . La manera de modelar la dinámica temporal de la frecuencia de los alelos A y a se realizará en tres pasos:

En un primer paso consideraremos todo el pool génico y supondremos que la frecuencia del alelo A es p y la del alelo a es $q = 1 - p$. A continuación, formaremos cigotos. Para esto, calcularemos la probabilidad de obtener un alelo A y un alelo a teniendo en cuenta las probabilidades de mutación:

$$\begin{aligned} P(A) &= P(\text{extraer } A \text{ y que éste no mute a } a) + P(\text{extraer } a \text{ y que éste mute a } A) \\ &= p(1 - \mu_1) + q\mu_2 \end{aligned}$$

y

$$\begin{aligned} P(a) &= P(\text{extraer } a \text{ y que éste no mute a } A) + P(\text{extraer } A \text{ y que éste mute a } a) \\ &= q(1 - \mu_2) + p\mu_1. \end{aligned}$$

Definiremos las frecuencias de A y a , tras mutación, $p^{(mut)}$ y $q^{(mut)}$, como

$$p_{mut} = p(1 - \mu_1) + q\mu_2, \quad q_{mut} = q(1 - \mu_2) + p\mu_1.$$

En la siguiente etapa formaremos cigotos y dejaremos que estos se desarrollen en adultos. Siguiendo un desarrollo análogo al visto en la sección (4.4) tenemos

$$P(AA|S) = \frac{\omega_{11}p_{mut}^2}{\bar{\omega}}, \quad P(Aa|S) = \frac{\omega_{12}2p_{mut}q_{mut}}{\bar{\omega}}, \quad P(aa|S) = \frac{\omega_{22}q_{mut}^2}{\bar{\omega}}$$

con

$$\bar{\omega} = \omega_{11}p_{mut}^2 + \omega_{12}2p_{mut}q_{mut} + \omega_{22}q_{mut}^2.$$

La última etapa del modelo es obtener una población de tamaño $2N$. Para esto obtendremos $2N$ realizaciones independientes de una distribución Bernoulli de parámetro p_{sel} . Podemos construir entonces el siguiente proceso que describe cómo es la dinámica poblacional:

Sea $Z_n : \Omega \rightarrow \{0, 1, \dots, 2N\}$ el número de alelos A en la generación n . Como el número de alelos A de la generación $n+1$ sólo depende del número de alelos de la generación n , $\{Z_n\}_{n \in \mathbb{N}}$ es una cadena de Márkov discreta con espacio de estados $E = \{0, 1, \dots, 2N\}$. Las probabilidades de transición son

$$p_{ij} = \mathbf{P}(Z_{n+1} = j | Z_n = i) = \binom{2N}{j} (p_{\text{sel}})^j (1 - p_{\text{sel}})^{2N-j}$$

con

$$p_{\text{sel}} = \frac{p_{\text{mut}}^2 w_{11} + p_{\text{mut}}(1 - p_{\text{mut}})w_{12}}{p_{\text{mut}}^2 w_{11} + 2p_{\text{mut}}(1 - p_{\text{mut}})w_{12} + (1 - p_{\text{mut}})^2 w_{22}}$$

y

$$p_{\text{mut}} = \frac{i}{2N} (1 - \mu_1) + (1 - \frac{i}{2N})\mu_2.$$

Definición 4.5.1 (Proceso de Wright-Fisher con selección y mutación). *Diremos que una cadena de Markov en tiempo discreto $Z = \{Z_n\}_{n \in \mathbb{N}}$ es un **proceso de Wright-Fisher con selección y mutación** de parámetros N , z_0 , $\boldsymbol{\omega} = (\omega_{11}, \omega_{12}, \omega_{22})$ y $\boldsymbol{\mu} = (\mu_1, \mu_2)$ sobre $E = \{0, 1, 2, \dots, 2N\}$ Si*

- i) la distribución inicial es $\mathbf{P}(Z_0 = z_0) = 1$ y,
- ii) la matriz de transición $\mathbb{P} = ((p_{ij}))_{i,j=0,1,\dots,2N}$ es tal que

$$p_{ij} = \mathbf{P}(Z_{n+1} = j | Z_n = i) = \binom{2N}{j} (p_{\text{sel}})^j (1 - p_{\text{sel}})^{2N-j}$$

con

$$p_{\text{sel}} = \frac{p_{\text{mut}}^2 w_{11} + p_{\text{mut}}(1 - p_{\text{mut}})w_{12}}{p_{\text{mut}}^2 w_{11} + 2p_{\text{mut}}(1 - p_{\text{mut}})w_{12} + (1 - p_{\text{mut}})^2 w_{22}}$$

y

$$p_{\text{mut}} = \frac{i}{2N} (1 - \mu_1) + (1 - \frac{i}{2N})\mu_2.$$

La principal diferencia entre el modelo de Wright-Fisher con mutación y los modelos sin mutación radica en la estructura de los estados de la cadena de Markov subyacente: en un proceso de Wright-Fisher con mutación ya no existirán estados absorbentes y la cadena resultará irreducible. Esto implicará que ya no será de interés el tiempo esperado de alcanzar los estados absorbentes y, por el contrario, el interés estará en estudiar la distribución estacionaria de Z . Al igual que en la secciones anteriores, la herramienta para estudiar el proceso de Wright-Fisher con mutación será la aproximación por difusión.

4.5.1. Aproximación por difusión

Aproximaremos al proceso de Wright-Fisher con selección y mutación por un proceso de difusión límite de forma análoga a lo obtenido en las secciones (4.3.2) y (4.4.1).

Sean $\omega = (\omega_{11}, \omega_{12}, \omega_{22})$ un vector de viabilidades y $\mu = (\mu_1, \mu_2)$ un vector de probabilidades de mutación. Para cada $N \in \mathbb{N}$ consideraremos el proceso de Wright-Fisher con selección de parámetros N, ω, μ y p_0 , $Z_N = \{Z_n^{(N)}\}_{n \in \mathbb{N}}$. Al igual que en la sección (4.4.1) asumiremos que las viabilidades ω_{ij} son $O(2N)$ realizando la reparametrización

$$\omega_{ij} = 1 + \frac{\sigma_{ij}}{2N}. \quad (4.15)$$

Consideraremos también que las mutaciones μ_i son también $O(2N)$ realizando la reparametrización

$$\mu_i = \frac{u_i}{2N} \quad (4.16)$$

Para cada $N \in \mathbb{N}$ encajaremos al proceso $Z^{(N)}$ en un proceso de tiempo continuo $X^{(N)} = \{X_t^{(N)}\}_{t \geq 0}$ tal que

$$X_t^{(N)} = \frac{1}{2N} Z_{\lfloor 2Nt \rfloor}^{(N)}.$$

Para cada $N \in \mathbb{N}$ notaremos con $\mathcal{S}_N = \{S_{N,t}\}_{t \geq 0}$ al semigrupo de operadores de transición

$$(S_{N,t}f)(x) = \mathbb{E}(f(X_t^{(N)}) | X_0 = x)$$

para toda $f : \{0, \dots, 2N\} \rightarrow \mathbb{R}$ continua.

Consideramos el operador diferencial $\mathcal{G} : C^2[0, +\infty) \rightarrow C^2[0, +\infty)$ tal que

$$(\mathcal{G}f)(x) = a(x) \frac{\partial}{\partial x} f'(x) + \frac{1}{2} x(1-x) \frac{\partial^2}{\partial^2 x} f''(x)$$

con

$$a(x) = -u_1x + u_2(1-x) + x^2(1-x)\sigma_{11} + x(1-x)(1-2x)\sigma_{12} - x(1-x)^2\sigma_{22}$$

Proposición 4.5.1. Bajo las condiciones anteriores se tiene que

$$\lim_{N \rightarrow \infty} \sup_{0 \leq t \leq t_0} \sup_{x \in \tilde{E}_N} |(S_{N,t}f)(x) - (S_t f)(x)| = 0 \quad (4.17)$$

para toda $f \in C([0, 1])$ y todo $t_0 \geq 0$. Consecuentemente, si

$$\lim_{N \rightarrow \infty} X_0^{(N)} = X_0$$

en distribución, entonces $X^{(N)} \Rightarrow X$ débilmente en $D_E[0, +\infty)$.

Demostración. Análoga a la proposición (4.4.1). ■

A partir de la proposición anterior podemos obtener trayectorias de la aproximación por difusión sabiendo que el generador infinitesimal

$$(\mathcal{G}f)(x) = a(x)\frac{\partial}{\partial x}f'(x) + \frac{1}{2}x(1-x)\frac{\partial^2}{\partial^2 x}f''(x)$$

con

$$a(x) = -u_1x + u_2(1-x) + x^2(1-x)\sigma_{11} + x(1-x)(1-2x)\sigma_{12} - x(1-x)^2\sigma_{22}$$

corresponde a la solución de la ecuación diferencial estocástica $X = \{X_t\}_{t \geq 0}$ tal que

$$\begin{cases} dX_t = (-u_1X_t + u_2(1-X_t) + \\ \quad X_t^2(1-X_t)\sigma_{11} + X_t(1-X_t)(1-2X_t)\sigma_{12} - X_t(1-X_t)^2\sigma_{22})dt \\ \quad + \sqrt{X_t(1-X_t)} dW_t \\ X_0 = p_0 \in (0, 1) \end{cases} \quad (4.18)$$

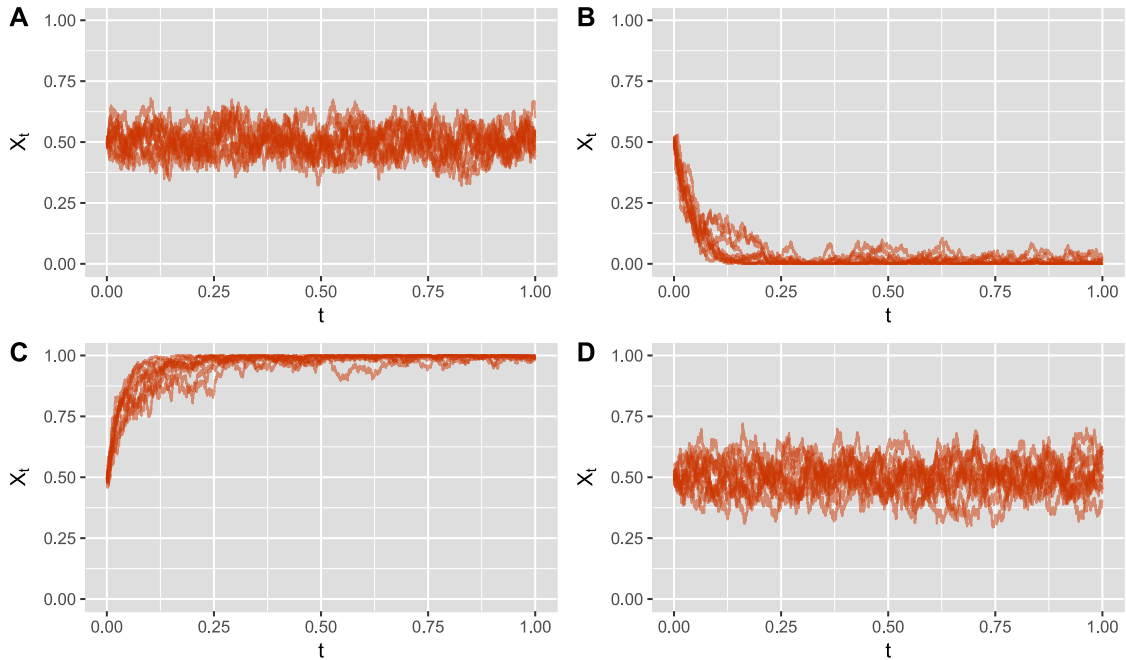


Figura 4.11: Trayectorias de la difusión $X = \{X_t\}_{t \geq 0}$ correspondiente al proceso de Wright-Fisher con selección y mutación. Las trayectorias fueron simuladas mediante el esquema de Milstein con una precisión de $n = 10^5$. (A) Soluciones de la ecuación (4.18) con parámetros $N = 100$, $u_1 = u_2 = 2N \times 0,1$ y $\sigma_{11} = \sigma_{12} = \sigma_{22} = 0$. (B) Soluciones de la ecuación (4.18) con parámetros $N = 100$, $u_1 = 2N \times 0,1$, $u_2 = 2N \times 0,001$ y $\sigma_{11} = \sigma_{12} = \sigma_{22} = 0$. (C) Soluciones de la ecuación (4.18) con parámetros $N = 100$, $u_1 = 2N \times 0,001$, $u_2 = 2N \times 0,1$ y $\sigma_{11} = \sigma_{12} = \sigma_{22} = 0$. (D) Soluciones de la ecuación (4.18) con parámetros $N = 100$, $u_1 = u_2 = \sigma_{11} = \sigma_{12} = \sigma_{22} = 2N \times 0,1$.

4.5.2. Distribución estacionaria

Proposición 4.5.2. Se considera $X = \{X_t\}_{t \geq 0}$ la difusión límite de un proceso de Wright-Fisher con selección y mutación tal que $\mu_1, \mu_2 > 0$. Entonces X tiene una única distribución estacionaria $\nu \in \mathcal{P}([0, 1])$. Más aun, ν es absolutamente continua con respecto a la medida de Lesbegue en $[0, 1]$ y su densidad f es la única solución de la ecuación

$$-(af)' + \frac{1}{2}(bf)'' = 0$$

tal que

$$\int_0^1 f(x) dx = 1$$

con

$$a(x) = -u_1x + u_2(1-x) + x^2(1-x)\sigma_{11} + x(1-x)(1-2x)\sigma_{12} - x(1-x)^2\sigma_{22}$$

y

$$b(x) = x(1-x).$$

Demostración. Ver [Ethier y Kurtz, 2005]. ■

En la figura (4.12) se observa la densidad estimada de la distribución estacionaria para varias tasas de mutación.

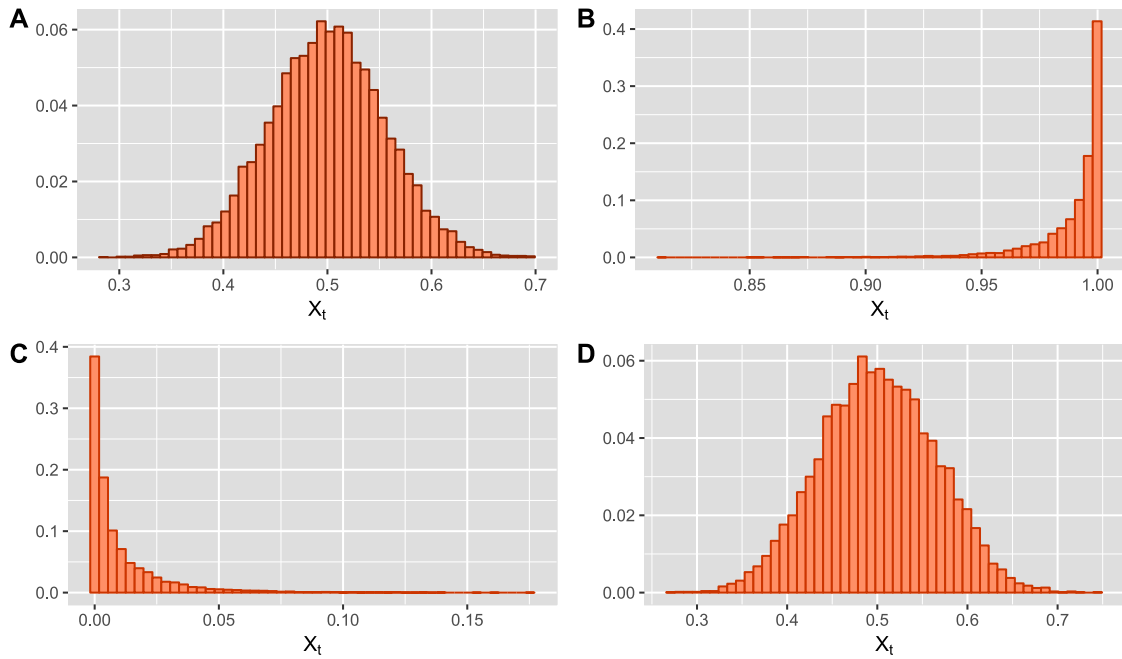


Figura 4.12: Densidad estimada de la distribución estacionaria de la difusión $X = \{X_t\}_{t \geq 0}$. Fueron simuladas $k = 10^5$ trayectorias para estimar la densidad de $t = 100$. (A) Corresponde a la difusión (4.18) con parámetros $N = 100$, $u_1 = u_2 = 2N0,1$ y $\sigma_{11} = \sigma_{12} = \sigma_{22} = 0$. (B) Corresponde a la difusión (4.18) con parámetros $N = 100$, $u_1 = 2N0,001$, $u_2 = 2N0,1$ y $\sigma_{11} = \sigma_{12} = \sigma_{22} = 0$. (C) Corresponde a la difusión (4.18) con parámetros $N = 100$, $u_1 = 2N0,1$, $u_2 = 2N0,001$ y $\sigma_{11} = \sigma_{12} = \sigma_{22} = 0$. (D) Corresponde a la difusión (4.18) con parámetros $N = 100$, $u_1 = u_2 = \sigma_{11} = \sigma_{22} = 2N0,001$ y $\sigma_{12} = 0$.

Las densidades para los casos donde $\sigma_{11} = \sigma_{12} = \sigma_{22}$ es teóricamente una distribución Beta (ver [Ethier y Kurtz, 2005]). Las densidades expuestas en la figura (4.12) se ajustan una distribución Beta (test de Kolmogorov-Smirnov, $p > 0,01$).

Comentarios finales

Las herramientas brindadas por los procesos de difusión resultan útiles para estudiar el proceso de Wright-Fisher discreto. El estudio de la difusión que aproxima al proceso discreto permite estudiar algunos aspectos de este último que las herramientas propias de las cadenas de Markov no permiten. Por otra parte, resultados como el de la proposición (4.3.4) dan la pauta de que se pueden estudiar otros resultados sobre la difusión límite resolviendo ecuaciones diferenciales determinísticas apropiadas. En particular, cabe la pregunta de si se puede estudiar la distribución de X_t para un t fijo resolviendo alguna ecuación diferencial determinística; resolver esto puede ser útil para dar un sustento teórico a algunas aseveraciones realizadas en algunos artículos de genética de poblaciones (por ejemplo, en [Nicholson et al., 2002]). Da esta monografía, por lo tanto, un puntapié inicial para un estudio más exhaustivo de este proceso fundamental en genética de poblaciones.

Capítulo 5

Apéndice

5.1. Cadenas de Markov discretas

Definición 5.1.1 (Cadena de Markov discreta). Sea $(\Omega, \mathcal{A}, \mathbf{P})$ un espacio de probabilidad y $X_n : \Omega \rightarrow E$ una variable aleatoria para cada $n \in \mathbb{N}$ que toma valores en un espacio de estados E numerable. Se dice que $\{X_n\}_{n \in \mathbb{N}}$ es una **cadena de Markov**, si

$$\mathbf{P}(X_{n+1} = i_{n+1} | X_n = i_n, X_{n-1} = i_{n-1}, \dots, X_1 = i_1, X_0 = i_0) = \mathbf{P}(X_{n+1} = i_{n+1} | X_n = i_n),$$

para todo $n \in \mathbb{N}$ e $i_0, i_1, \dots, i_{n-1}, i_n, i_{n+1} \in E$.

Definición 5.1.2 (cadena de Markov homogénea). Se dice que una cadena de Markov $\{X_n\}_{n \in \mathbb{N}}$ es **homogénea** si

$$\mathbf{P}(X_{n+1} = j | X_n = i) = \mathbf{P}(X_1 = j | X_0 = i),$$

para todo $n \in \mathbb{N}$.

Si $\{X_n\}_{n \in \mathbb{N}}$ es una cadena de Márkov homogénea notaremos

$$p_{ij} = \mathbf{P}(X_1 = j | X_0 = i),$$

y llamaremos a p_{ij} la *probabilidad de transición* del estado i al estado j .

Definición 5.1.3 (Distribución inicial). Dada una cadena de Markov $\{X_n\}_{n \in \mathbb{N}}$ definida en un espacio de estados $E = (0, 1, 2, \dots)$, se define su **distribución inicial** al vector $\lambda = (\lambda_0, \lambda_1, \lambda_2, \dots) \in \mathbb{R}^{|E|}$ tal que

$$\lambda_j = \mathbf{P}(X_0 = j)$$

Definición 5.1.4 (matriz de transición). Dada una cadena de Márkov $X = \{X_n\}_{n \in \mathbb{N}}$ homogénea definida en un espacio de estados $E = (0, 1, 2, \dots)$, se define su **matriz de transición** $\mathbb{P} = ((p_{ij})) \in \mathbb{R}^{|E| \times |E|}$ donde p_{ij} es la probabilidad de transición del estado i al estado j . Esto es, $\mathbb{P}$ es la matriz (eventualmente infinita):

$$\mathbf{P} = \begin{pmatrix} p_{00} & p_{01} & p_{02} & \dots \\ p_{10} & p_{11} & p_{12} & \dots \\ p_{20} & p_{21} & p_{22} & \dots \\ \vdots & \vdots & \vdots & \\ p_{i0} & p_{i1} & p_{i2} & \dots \\ \vdots & \vdots & \vdots & \end{pmatrix}$$

Proposición 5.1.1. Sea $X = \{X_n\}_{n \in \mathbb{N}}$ una cadena de Márkov homogénea en un espacio de estados E con matriz de transición $\mathbf{P}$ y distribución inicial λ . Entonces se cumple que

$$\mathbf{P}(X_0 = i_0, X_1 = i_1, \dots, X_N = i_N) = \lambda_{i_0} p_{i_0 i_1} p_{i_1 i_2} \dots p_{i_{N-1} i_N}$$

Demostración. Supongamos que $\{X_n\}_{n \in \mathbb{N}}$ es una cadena de Márkov homogénea. Entonces, aplicando la definición de probabilidad condicionada se tiene que

$$\begin{aligned} \mathbf{P}(X_n = i_n, X_{n-1} = i_{n-1}, \dots, X_1 = i_1, X_0 = i_0) = \\ \mathbf{P}(X_n = i_n | X_{n-1} = i_{n-1}, \dots, X_1 = i_1, X_0 = i_0) \mathbf{P}(X_{n-1} = i_{n-1}, \dots, X_1 = i_1, X_0 = i_0) \end{aligned}$$

Como X es una cadena de Markov

$$\begin{aligned} \mathbf{P}(X_n = i_n | X_{n-1} = i_{n-1}, \dots, X_1 = i_1, X_0 = i_0) &= \mathbf{P}(X_n = i_n | X_{n-1} = i_{n-1}) \\ &= p_{i_{n-1} i_n} \end{aligned}$$

Luego

$$\begin{aligned} \mathbf{P}(X_n = i_n, X_{n-1} = i_{n-1}, \dots, X_1 = i_1, X_0 = i_0) = \\ p_{i_{n-1} i_n} \mathbf{P}(X_{n-1} = i_{n-1}, \dots, X_1 = i_1, X_0 = i_0) \end{aligned}$$

Aplicando $n - 1$ veces la definición de probabilidad condicionada y la hipótesis de que $\{X_n\}$ es una cadena de Márkov para el término $\mathbf{P}(X_{n-1} = i_{n-1}, \dots, X_1 = i_1, X_0 = i_0)$ se concluye que

$$\mathbf{P}(X_0 = i_0, X_1 = i_1, \dots, X_n = i_n) = \lambda_{i_0} p_{i_0 i_1} p_{i_1 i_2} \dots p_{i_{n-1} i_n}$$

■

Observación 5.1.1. Este último resultado implica que toda cadena de Markov $\{X_n\}_{n \in \mathbb{N}}$ homogénea queda totalmente determinada por su distribución inicial λ y su matriz de transición $\mathbf{P}$. Esto se debe a que dados índices $j_1, \dots, j_k \in \mathbb{N}$ tales que $j_1 < j_2 < \dots < j_k$ podemos calcular la distribución de $(X_{j_1}, \dots, X_{j_k})$ utilizando la ley de probabilidad total para obtener la distribución de $(X_0, X_1, \dots, X_{j_k})$ y luego aplicando el resultado anterior.

Denotaremos $p_{ij}^{(n)}$ a la probabilidad de transición del estado i al estado j en n pasos:

$$p_{ij}^{(n)} = \mathbf{P}(X_{m+n} = j | X_m = i) = \mathbf{P}(X_n = j | X_0 = i)$$

Proposición 5.1.2 (Ecuación de Chapman-Kolmogorov). Dada una cadena de Márkov $X = \{X_n\}_{n \in \mathbb{N}}$ homogénea definida en un espacio de estados numerable E se cumple que

$$p_{ij}^{(n+m)} = \sum_{k \in E} p_{ik}^{(n)} p_{kj}^{(m)}$$

Demostración.

$$\begin{aligned} p_{ij}^{(n+m)} &= \mathbf{P}(X_{n+m} = j | X_0 = i) = \sum_{k \in E} \mathbf{P}(X_{n+m} = j, X_n = k | X_0 = i) \\ &= \sum_{k \in E} \frac{\mathbf{P}(X_{n+m} = j, X_n = k, X_0 = i)}{\mathbf{P}(X_0 = i)} \\ &= \sum_{k \in E} \underbrace{\mathbf{P}(X_{n+m} = j | X_n = k, X_0 = i)}_{\mathbf{P}(X_{n+m} = j | X_n = k)} \frac{\mathbf{P}(X_n = k, X_0 = i)}{\mathbf{P}(X_0 = i)} \\ &= \sum_{k \in E} \mathbf{P}(X_{n+m} = j | X_n = k) \mathbf{P}(X_n = k | X_0 = i) \\ &= \sum_{k \in E} p_{ik}^{(n)} p_{kj}^{(m)}, \end{aligned}$$

donde en la última igualdad se utilizó que

$$\mathbf{P}(X_{n+m} = j | X_n = k) = \mathbf{P}(X_m = j | X_0 = k)$$

ya que la cadena es homogénea. ■

Definición 5.1.5. Dada una cadena de Márkov $X = \{X_n\}_{n \in \mathbb{N}}$ homogénea definida en un espacio de estados numerable E , se dice que de un estado i se va a un estado j , y se nota como $i \rightarrow j$ si

$$\mathbf{P}(\exists n \in \mathbb{N} : X_n = j | X_0 = i) > 0$$

Se dice que i se comunica con j y se nota como $i \leftrightarrow j$ si $i \rightarrow j$ y $j \rightarrow i$.

Teorema 5.1.1. Sea una cadena de Márkov $X = \{X_n\}_{n \in \mathbb{N}}$ homogénea definida en un espacio de estados numerable E . Son equivalentes:

1. $i \rightarrow j$
2. Existe un conjunto de estados $i = i_0, i_1, \dots, i_n = j$ tales que

$$p_{i_0 i_1} p_{i_1 i_2} \cdots p_{i_{n-1} i_n} > 0$$

3. Existe $n_0 \in \mathbb{N}$ tal que $p_{ij}^{(n)} > 0$.

Demostración. (1) $\Rightarrow$ (3) Se tiene que

$$p_{ij}^{(n)} = \mathbf{P}(X_n = j | X_0 = i)$$

$$\begin{aligned}
&\leq \mathbf{P} \left(\bigcup_{n=0}^{\infty} \{X_n = j\} \middle| X_0 = i \right) \\
&\leq \sum_{n=0}^{\infty} \mathbf{P}(X_n = j | X_0 = i) \\
&= \sum_{n=0}^{\infty} p_{ij}^{(n)}
\end{aligned}$$

Por hipótesis tenemos que

$$\mathbf{P}(\exists n \in \mathbb{N} : X_n = j | X_0 = i) > 0.$$

Esto es equivalente a

$$\mathbf{P} \left(\bigcup_{n=0}^{\infty} \{X_n = j\} \middle| X_0 = i \right) > 0$$

Luego por la inecuación en (5.1) se tiene que $\sum_{n=0}^{\infty} p_{ij}^{(n)} > 0$ y por lo tanto debe existir $n_0 \in \mathbb{N}$ tal que $p_{ij}^{n_0} > 0$.

(3) $\Rightarrow$ (1) Por hipótesis se tiene que existe $n_0 \in \mathbb{N}$ tal que $p_{ij}^{n_0} > 0$. Luego, por la inecuación vista en (5.1) se cumple que

$$0 < p_{ij}^{n_0} \leq \mathbf{P} \left(\bigcup_{n=0}^{\infty} \{X_n = j\} \middle| X_0 = i \right),$$

esto es, $i \rightarrow j$.

(1) $\Leftrightarrow$ (2) Se tiene que

$$p_{ij}^{(n)} = \sum_{i_1, \dots, i_{n-1}} p_{ii_1} p_{i_1 i_2} \cdots p_{i_{n-1} j}$$

Luego, por la inecuación (5.1) se cumple lo pedido. ■

Proposición 5.1.3. Dada una cadena de Márkov $X = \{X_n\}_{n \in \mathbb{N}}$ la relación $i \leftrightarrow j$ define una relación de equivalencia en $E \times E$.

Demostración. • **Reflexividad.** Dado $i \in E$ se tiene que

$$\mathbf{P}(X_0 = i | X_0 = i) = 1,$$

luego existe $n \in \mathbb{N}$ tal que $i \leftrightarrow i$.

• **Simetría.** Se sigue trivialmente de la definición de $i \leftrightarrow j$.

• **Transitividad.** Sean $i, j, k \in E$ tales que $i \rightarrow j$ y $j \rightarrow k$. Por teorema 5.1.1 tenemos que existen $n_0, n_1 \in \mathbb{N}$ tales que

$$p_{ij}^{n_0} > 0, \quad p_{jk}^{n_1} > 0.$$

Luego, por teorema (5.1.2)

$$0 < p_{ij}^{n_0} p_{jk}^{n_1} \leq \sum_{l \in E} p_{il}^{n_0} p_{lk}^{n_1} = p_{ij}^{(n_0+n_1)}$$

Lo que implica que $i \rightarrow k$. La prueba de que $k \rightarrow i$ es análoga. ■

Definición 5.1.6 (Clase cerrada, estado absorbente y cadena irreducible). *Sea una cadena de Márkov $X = \{X_n\}_{n \in \mathbb{N}}$ homogénea definida en un espacio de estados numerable $(E, \leftrightarrow)$.*

- *Diremos que una clase de equivalencia C es cerrada si dado $i \in C$ y $j \in E$ tal que $i \rightarrow j$ esto implica que $j \in C$.*
- *Un estado $i \in E$ se dice absorbente si $\{i\}$ es una clase cerrada.*
- *Diremos que X es una cadena irreducible si E es la única clase de equivalencia.*

Supongamos que X_n es una cadena de Markov en tiempo discreto definida en un espacio de estados E finito. Sea $A \subset E$, $A \neq \emptyset$. Definimos un *tiempo de llegada*, τ_A a la variable aleatoria

$$\tau_A = \inf_{n \geq 0} \{X_n \in A\}, \quad \text{o,} \quad \tau_A = \infty \quad \text{si} \quad \{X_n \in A\} = \emptyset$$

Estamos interesados en la esperanza de la variable τ_A , condicionado a que se partió de un estado i :

$$k_i^A = \mathbb{E}(\tau_A | X_0 = i) = \sum_{n=0}^{\infty} n \mathbf{P}(\tau_A = n | X_0 = i)$$

Proposición 5.1.4. El vector de esperanzas de tiempos de llegada $k^A \in \mathbb{R}^{|E|}$ definido como

$$k^A = \{k_i^A : i \in E\}$$

es solución la mínima solución no negativa de

$$\begin{cases} k_i^A = 0 & \text{si } i \in A \\ k_i^A = 1 + \sum_{j \notin A} p_{ij} k_j^A & \text{si } i \notin A \end{cases}$$

Demostración. Ver [Norris, 1997], p. 17. ■

5.2. Espacio de Skorokhod

Las trayectorias del proceso estocástico definido en la sección (4.3.2) no son continuas: presentan saltos en los puntos $t = \lfloor i/2N \rfloor$, con $i \in \mathbb{N}$. Sin embargo, las trayectorias de este proceso tienen la propiedad de tener límite por izquierda y ser continuas por derecha. Las funciones que cumplan tengan esta propiedad son las llamadas funciones *càdlàg*:

Definición 5.2.1 (Función *càdlàg*). Sea (X, d) un espacio métrico. Una función $f : [0, \infty) \rightarrow X$ se dice que es *càdlàg* (del francés ‘continue à droite, limite à gauche’) si cumple que

a) f es continua por derecha. Esto es, para todo $t \in [0, \infty)$ se cumple que

$$f(t) = \lim_{h \rightarrow 0} f(t+h)$$

b) f tiene límite por izquierda. Esto es, para todo $t \in [0, \infty)$ se cumple que

$$\exists \lim_{h \rightarrow 0} f(t-h)$$

Definición 5.2.2 (Espacio de Skorokhod). Dado (X, d) un espacio métrico, definimos como espacio de Skorokhod, y lo notamos como $D_X[0, +\infty)$, al conjunto

$$D_X[0, +\infty) = \{f : [0, +\infty) \rightarrow X : f \text{ es càdlàg}\}$$

5.3. Integrabilidad uniforme

Definición 5.3.1 (Integrabilidad uniforme). Una sucesión de variables aleatorias $\{X_n\}_{n \in \mathbb{N}}$ se dice que es *uniformemente integrable* si

$$\lim_{\alpha \rightarrow \infty} \sup_{n \in \mathbb{N}} \int_{\{|X_n| \geq \alpha\}} |X_n| d\mathbf{P} = 0$$

Proposición 5.3.1. Sea $\{X_n\}$ una sucesión de variables aleatorias uniformemente integrables y sea X tal que $X_n \xrightarrow{d} X$. Entonces

$$\lim_{n \rightarrow \infty} \mathbb{E}(X_n) = \mathbb{E}(X)$$

Demostración. Ver [Billingsley, 1968], p. 32. ■

5.4. Distribución normal truncada

Sea Z una variable aleatoria normal de parámetros μ y σ^2 . Notamos con ϕ a la densidad de Z , esto es

$$\phi(x; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

y, con Φ a la función de distribución de Z , esto es

$$\Phi(x; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^x e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2} dt$$

Dada Z y $a, b \in \mathbb{R}$ tales que $a < b$ diremos que una variable aleatoria X tiene distribución normal truncada entre a y b si su función de densidad $f : \mathbb{R} \rightarrow \mathbb{R}$ es tal que

$$f(x; \mu, \sigma^2, a, b) = \begin{cases} 0 & \text{si } x < a \\ \frac{\phi(x; \mu, \sigma^2)}{\Phi(b; \mu, \sigma^2) - \Phi(a; \mu, \sigma^2)} & \text{si } a \leq x < b \\ 0 & \text{si } x > b \end{cases}$$

También es posible definir una distribución truncada unilateralmente tomando $a = -\infty$ o $b = +\infty$ y definiendo $\Phi(a; \mu, \sigma^2) = 0$ y $\Phi(b; \mu, \sigma^2) = 1$.

5.5. Bondad de ajuste de la difusión límite del proceso de Wright-Fisher a una normal truncada

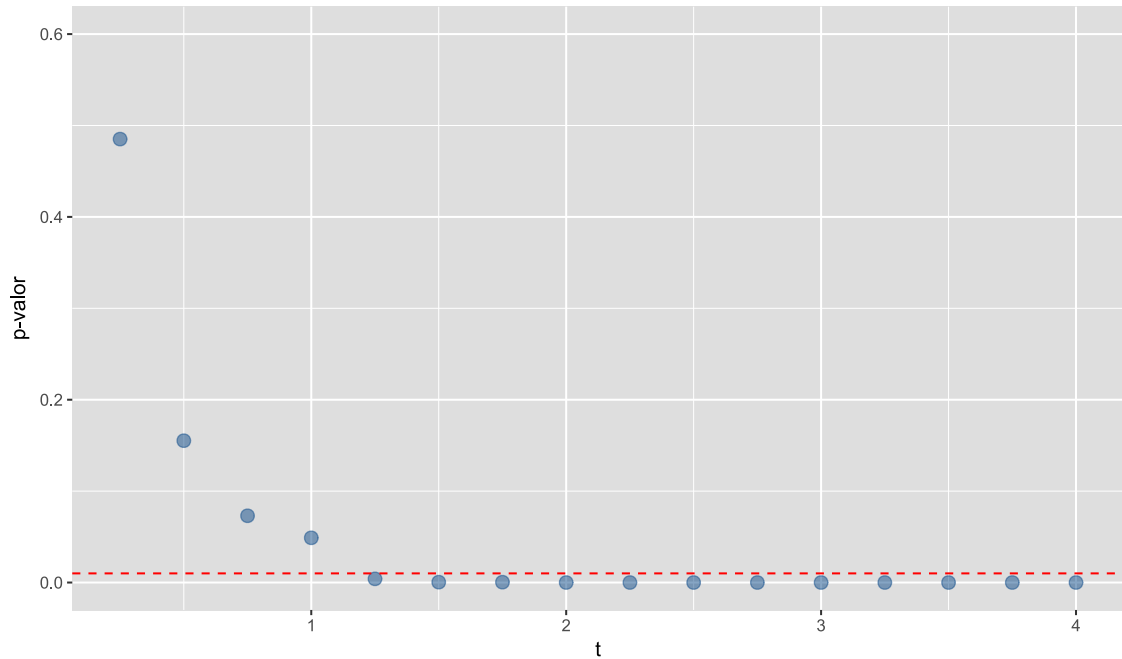


Figura 5.1: p -valores de bondad de ajuste normal de X_t para distintos valores de t . Para cada tiempo t se simulaban 10,000 trayectorias y se realizó un test de Kolmogorov-Smirnov tomando como hipótesis nula que la distribución de X_t es una normal truncada de parámetros $a = 0$, $b = 1$, $\mu = \bar{X}_n$ y $\sigma = S_n$ (ver (5.4)). La línea roja punteada corresponde a un p -valor de 0,01.

Bibliografía

- [Billingsley, 1968] Billingsley, P. (1968). *Convergence of Probability Measures*. John Wiley & Sons.
- [Borodin, 2017] Borodin, A. N. (2017). *Stochastic Processes*. Birkhäuser.
- [Brzezniak y Zastawniak, 1999] Brzezniak, Z. y Zastawniak, T. (1999). *Basic Stochastic Processes*. Springer.
- [Carsouw, 2012] Carsouw, M. F. J. (2012). Wright-fisher Evolution. Bachelor thesis.
- [Durrett, 2008] Durrett, R. (2008). *Probability Models for DNA Sequence Evolution*. Springer.
- [Engel y Nagel, 2006] Engel, K.-J. y Nagel, R. (2006). *A Short Course on Operator Semigroups*. Springer-Verlag New York.
- [Ethier y Kurtz, 2005] Ethier, S. N. y Kurtz, T. G. (2005). *Markov Processes: Characterization and Convergence*. John Wiley & Sons.
- [Ethier y Nagylaki, 1988] Ethier, S. N. y Nagylaki, T. (1988). Diffusion Approximations of Markov Chains with Two Time Scales and Applications to Population Genetics, ii. *Advances in Applied Probability*, 20.
- [Ewens, 2004] Ewens, W. J. (2004). *Mathematical Population Genetics*. Springer.
- [Feller, 1951] Feller, W. (1951). Diffusion processes in genetics. In *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, pages 227–246. University of California Press.
- [Fisher, 1922] Fisher, R. (1922). On the Dominance Ratio. *Proceedings of the Royal Society of Edinburgh*, 42:321–341.
- [Futuyma, 2005] Futuyma, D. J. (2005). *Evolution*. Sinauer Associates, Inc.
- [Griffiths et al., 2004] Griffiths, A. J. F., Wessler, S. R., Carroll, S. B., y Doebley, J. (2004). *An Introduction to Genetic Analysis*. W. H. Freeman.
- [Haldane, 1924] Haldane, J. B. S. (1924). A Mathematical Theory of Natural and Artificial Selection - Part 1. *Transactions of the Cambridge Philosophical Society*, 23:19–41.

- [Hartl y Clark, 1997] Hartl, D. L. y Clark, A. G. (1997). *Principles of Population Genetics*. Sinauer Associates, Inc.
- [Herron y Freeman, 2014] Herron, J. C. y Freeman, S. (2014). *Evolutionary Analysis*. Pearson.
- [Hille y Philips, 1957] Hille, E. y Philips, R. S. (1957). *Functional Analysis and Semi-Groups*. American Mathematical Society.
- [Howie, 1976] Howie, J. M. (1976). *An Introduction to Semigroup Theory*. New York: Academic Press.
- [Kimura, 1964] Kimura, M. (1964). Diffusion Models in Population Genetics. *Journal of Applied Probability*, 1.
- [Kloeden y Platen, 1995] Kloeden, P. E. y Platen, E. (1995). *Numerical Solutions of Stochastic Differential Equations*. Springer.
- [Kreyszig, 1978] Kreyszig, E. (1978). *Introductory Functional Analysis with Applications*. John Wiley & Sons.
- [Moore, 2001] Moore, R. (2001). The “Rediscover” of Mendel’s Work. *Bioscene*, 27(2):13–26.
- [Nicholson et al., 2002] Nicholson, G., Smith, A. V., Jónsson, F., Gústafsson, O., Stefánsson, K., y Donnelly, P. (2002). Assessing population differentiation and isolation from single-nucleotide polymorphism data. *Journal of the Royal Statistical Society*.
- [Norris, 1997] Norris, J. R. (1997). *Markov Chains*. Cambridge University Press.
- [Petrov y Mordecki, 2008] Petrov, V. V. y Mordecki, E. (2008). *Teoría de la probabilidad*. Dirac.
- [Portin y Wilkins, 2017] Portin, P. y Wilkins, A. (2017). The Evolving Definition of the Term “gene”. *Genetics*, 205(4):1353–1364.
- [Shiryaev, 1996] Shiryaev, A. N. (1996). *Probability*. Springer.
- [Weber, 1986] Weber, N. C. (1986). The Wright-Fisher Model with Varying Selection. *Journal of Applied Probability*, 23.
- [Wright, 1931] Wright, S. (1931). Evolution in Mendelian Populations. *Genetics*, 16:97–159.
- [Çınlar, 2011] Çınlar, E. (2011). *Probability and Stochastics*. Springer.