

UNIVERSIDAD NACIONAL AUTÓNOMA DE  
MÉXICO

---

POSGRADO EN CIENCIAS MATEMÁTICAS  
FACULTAD DE CIENCIAS

CORRELACIÓN TEMPORAL EN LA VALUACIÓN  
DE DERIVADOS

T E S I S  
QUE PARA OBTENER EL TÍTULO ACADÉMICO DE:  
M A E S T R O E N C I E N C I A S  
P R E S E N T A:  
ACT. ENRIQUE MÉNDEZ RÍOS

DIRECTOR DE TESIS: PABLO PADILLA LONGORIA

MÉXICO, D.F.

Noviembre de 2005



Universidad Nacional  
Autónoma de México



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

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

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

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

cualquier función. En particular estudiaremos una correlación del tipo exponencial (ecuación de *Ornstein-Uhlenbeck*); pues, en el caso límite, cuando  $\Phi(\tau)$  es la delta de Dirac, tendremos el movimiento browniano recuperando de esta forma la fórmula de Black-Scholes. Los ambientes cambiantes se refieren a la propiedad de estar en uno u otro estado regida por un proceso de nacimiento y muerte. Primero se logra la valuación del derivado en dos ambientes (solución analítica), para después agregar la correlación temporal al modelo. Pensamos que la valuación de la opción en varios ambientes cambiantes podría ser de gran utilidad; imaginemos por ejemplo la situación de distinguir entre los ambientes de poco riesgo, estable y muy riesgoso; la utilidad del modelo se hace presente al incorporar o aproximar tales eventos que en otras circunstancias aparecerían indiferentes en el modelo simple. La correlación temporal,  $\tau$ , tiene efectos significativos en el precio del derivado. ¿Debería ser mayor o menor el precio del derivado al incrementar  $\tau$ ? El modelo modulado de Markov proporciona la respuesta. Finalmente, la manipulación de los resultados se facilita con los códigos en MATLAB que se utilizaron para realizar los cálculos numéricos. La intención no es que los códigos sean eficientes, más bien que sean prácticos.

## Agradecimientos

Mi más grande respeto y agradecimiento para P. Padilla por sus innumerables comentarios y su permanente apoyo durante el desarrollo de este trabajo. También agradezco a M. Bladt por sus productivas observaciones que me ayudaron a la mejor comprensión de sus artículos. De igual forma a Eduardo, Ramsés y Luis Antonio. La recopilación de la información no hubiera sido posible sin la amabilidad del personal de la biblioteca. A mis amigos y compañeros de estudio.

# Índice general

<b>Introducción</b>	<b>III</b>
<b>1. Fundamentos Matemáticos</b>	<b>1</b>
1.1. Cadenas de Markov Discretas . . . . .	1
1.2. Cadenas de Markov Continuas . . . . .	5
1.2.1. Proceso Puntual . . . . .	5
1.2.2. Proceso Poisson . . . . .	6
1.2.3. Distribución de una cadena de Markov Homogénea . . . . .	7
1.2.4. Matriz de Intensidad . . . . .	8
1.2.5. Sistema Diferencial de Kolmogorov . . . . .	10
1.2.6. Cadena Incluida y Tiempos de transición . . . . .	14
<b>2. Ecuaciones Diferenciales Estocásticas</b>	<b>21</b>
2.1. El Movimiento Browniano . . . . .	21
2.2. La Integral de Itô . . . . .	22
2.3. El Generador Infinitesimal . . . . .	28
<b>3. Aplicación</b>	<b>35</b>
3.1. Opciones . . . . .	35
3.2. Futuros . . . . .	37
3.3. Contratos Adelantados . . . . .	38
3.4. La Fórmula de Samuelson . . . . .	39
3.5. El Modelo Modulado de Markov . . . . .	40
3.6. Valuación del derivado para 2 estados . . . . .	42
3.7. Aproximación del derivado para $n$ estados . . . . .	44
3.7.1. La desigualdad de Jensen . . . . .	46
3.7.2. La griega $\Theta$ . . . . .	47
3.8. Correlación temporal en el sistema . . . . .	58
3.8.1. Aproximación unificada de ruido coloreado . . . . .	61
3.9. Escape de un pozo biestable . . . . .	64
3.9.1. El tiempo de escape es exponencial . . . . .	65
<b>Conclusión</b>	<b>74</b>

<b>A.</b>	<b>75</b>
A.1. Espacios de Probabilidad . . . . .	75
<b>B.</b>	<b>77</b>
B.1. Espacios de Banach y de Hilbert . . . . .	77
<b>C.</b>	<b>81</b>
C.1. Fórmulas de Green . . . . .	81
<b>Bibliografía</b>	<b>84</b>

# Introduccion

La llegada de las matematicas al campo de las finanzas ha sido a un ritmo bastante rapido y a veces hasta explosivo en los ultimos 20 años. Cuando se vio la utilidad de la medida equivalente de las martingalas en el precio y en la optimizacion de los portafolios, tecnicas poderosas como el Analisis Estocastico y el Control Estocastico han tenido que ver con casi todos los aspectos de las Finanzas: el arbitraje, la cobertura de portafolios, la optimizacion de las inversiones, los mercados incompletos, la tasa de equilibrio, los costos de transaccion y las tasas de interes, entre otros. Al mismo tiempo, el desarrollo de metodos numericos y analiticos sofisticados, basados primordialmente en ecuaciones diferenciales parciales y en sus soluciones, ha ayudado a incrementar la importancia de estos desarrollos en la practica diaria de las finanzas.

La valuacion de derivados financieros es uno de los grandes acontecimientos de la economia financiera moderna. Basada en la reconocida condicion de no arbitraje, el modelo de valuacion de opciones de Black-Scholes y Merton gano una inmediata aceptacion entre academicos e investigadores que no tiene parangon en la historia de la ciencia economica.

La correcta valuacion de las opciones ha desconcertado a los economistas por varios años. A principios de 1900 con su revolucionaria tesis “La Teoria de la Especulacion,” Louis Bachelier [3] (alumno de Poincare) describio la forma de valuar opciones. Sorprendentemente, un componente de la formula que el concibio para este proposito anticipo el modelo que Albert Einstein usara en su descripcion del movimiento Browniano, la dinamica aleatoria de particulas en un fluido. Sin embargo, la formula de Bachelier contenia supuestos financieros irreales tales como la existencia de valores negativos<sup>1</sup> para el precio de las acciones.

Otros academicos, incluyendo al premio Nobel Paul Samuelson, trataron de atacar el problema. Fracasaron en la tediosa labor de calcular la prima de riesgo: un descuento en el precio de la opcion que compensa la aversion al riesgo por parte del inversionista y los movimientos inciertos de las acciones en el mercado.

Black y Scholes por una parte, y Merton por la otra, dilucidaron la formula de valuacion de las opciones construyendo un portafolio hipotetico en el cual un cambio en el precio de la accion se compensaba con un cambio en el valor de la opcion—una estrategia llamada cobertura (*hedging*).

---

<sup>1</sup>Válido para algunos tipos de opciones *spread*.

En realidad, la formula de Black-Scholes se extrae de una ecuacion diferencial parcial demostrando que el precio justo para una opcion es la que traer a una tasa de rendimiento libre de riesgo en el portafolios mencionado. Esta estrategia propuesta por Black, Scholes y Merton ha sido invaluable para las instituciones bancarias y financieras que pueden usarla para proteger sus portafolios de las variaciones del mercado—un seguro contra una pronunciada ca da en los precios de las acciones, por ejemplo.

En los 70's, el art culo escrito por Black-Scholes tuvo di cultades para ser publicado. Cuando aparecio en *The Journal of Political Economy* en 1973, su impacto en los mercados financieros fue inmediato. En meses, su formula ya estaba programada en las calculadoras. *Wall Street* la amaba, ya que un corredor pod a resolver la ecuacion simplemente proporcionando algunas variables: el precio de la accion, la tasa de interes y la fecha de expiracion de la opcion. La unica variable que no pod a obtenerse facilmente era la volatilidad de la accion, la desviacion estandar del precio de la accion. Este valor sin embargo pod a estimarse. De igual manera, si se conoc a el valor de la opcion se pod a estimar un valor para la volatilidad, la cual se pod a usar para determinar si una opcion estaba sub o sobrevaluada con respecto al valor de la accion en el mercado.

Los inversionistas que compran opciones estan basicamente comprando la volatilidad—ya sea para especular o para protegerse de las turbulencias del mercado. Entre mas altibajos haya en el mercado mas valor tiene la accion. Un inversionista que especule con una opcion *call* (una opcion para comprar una accion) puede perder solamente el precio de compra, llamada prima, si el precio de la accion no alcanza al precio en el cual el comprador puede ejercer su derecho de compra de la accion. En contraste, si el precio de la accion sobrepasa al precio de ejercicio, el potencial de ganancias es ilimitado.

Aunque esto pueda reducirse a operaciones en una calculadora, las matematicas detras de la ecuacion de Black-Scholes es el calculo estocastico, un descendiente del trabajo de Bachelier y Einstein. Estas ecuaciones no se estudiaban en los programas de administracion de empresas; es aqu donde entran los cient cos de Wall Street: los fsicos, los matematicos, los programadores y los econometristas que ahora juegan un papel muy importante en la industria financiera.

Este trabajo se organizo de la manera siguiente. El cap tulo primero proporciona las nociones basicas de los procesos de Markov discretos y continuos, con especial enfasis en la matriz de intensidad, cuya importancia se reflejara en el ultimo cap tulo. La mayor parte de los conceptos se tomaron del libro de Bremaud [6], en donde se encontrara la demostracion de los teoremas presentados. El segundo cap tulo presenta el movimiento Browniano y el lema de Itô, un tema que no debe de faltar en el ambito financiero. El libro *Stochastic Differential Equations* de Øksendal [19] es un texto invaluable. El cap tulo tercero, la aplicacion, se basa en los art culos [13] y [12] escritos por M. Bladt y P. Padilla; en donde se valuan opciones en ambientes cambiantes. Nosotros trataremos de extender el numero de ambientes y agregar correlacion temporal. Por correlacion temporal nos referimos a una variable estocastica  $\eta(t)$  (ruido coloreado) con media cero y funcion de correlacion  $E[\eta(t)\eta(t + \tau)] = \tau^{-\alpha}$ , siendo  $\alpha > 0$

cualquier funcion. En particular estudiaremos una correlacion del tipo exponencial (ecuacion de *Ornstein-Uhlenbeck*); pues, en el caso limite, cuando  $\delta(t)$  es la delta de Dirac, tendremos el movimiento browniano recuperando de esta forma la formula de Black-Scholes. Los ambientes cambiantes se refieren a la propiedad de estar en uno u otro estado regida por un proceso de nacimiento y muerte. Primero se logra la valuacion del derivado en dos ambientes (solucion analitica), para despues agregar la correlacion temporal al modelo. Pensamos que la valuacion de la opcion en varios ambientes cambiantes podria ser de gran utilidad; imaginemos por ejemplo la situacion de distinguir entre los ambientes de poco riesgo, estable y muy riesgoso; la utilidad del modelo se hace presente al incorporar o aproximar tales eventos que en otras circunstancias aparecerian indiferentes en el modelo simple. La correlacion temporal,  $\rho$ , tiene efectos significativos en el precio del derivado. ¿Deberia ser mayor o menor el precio del derivado al incrementar  $\rho$ ? El modelo modulado de Markov proporciona la respuesta. Finalmente, la manipulacion de los resultados se facilita con los codigos en MATLAB que se utilizaron para realizar los calculos numericos. La intencion no es que los codigos sean eficientes, mas bien que sean practicos.

## Agradecimientos

Mi mas grande respeto y agradecimiento para P. Padilla por sus innumerables comentarios y su permanente apoyo durante el desarrollo de este trabajo. Tambien agradezco a M. Bladt por sus productivas observaciones que me ayudaron a la mejor comprension de sus articulos. De igual forma a Eduardo, Ramses y Luis Antonio. La recopilacion de la informacion no hubiera sido posible sin la amabilidad del personal de la biblioteca. A mis amigos y compañeros de estudio.



# Capítulo 1

## Fundamentos Matemáticos

*Los procesos estocásticos son un concepto fundamental en la teoría financiera. Describen fenómenos aleatorios que evolucionan en el tiempo, tales como el precio de las acciones, las tasas de rendimientos, el valor de un portafolio, etc. Presentaremos las ecuaciones diferenciales de Kolmogorov, la matriz de intensidad y el proceso de nacimiento y muerte como conceptos importantes en el desarrollo del trabajo.*

### 1.1. Cadenas de Markov Discretas

Los procesos de Markov constituyen la clase más importante de procesos estocásticos que son útiles para modelar fenómenos con cierto tipo de dependencia que aparece con frecuencia en diversas aplicaciones. Esta dependencia expresa que la evolución probabilista en el futuro, conociendo el pasado y el presente, depende solamente del estado presente; pero esta limitada cantidad de memoria es suficiente para producir una gran diversidad de comportamientos. La observación de los precios de las acciones o la posición de una partícula de difusión y muchos otros procesos observados en el tiempo son a menudo modelados por un proceso estocástico. Un proceso estocástico es cualquier colección de variables aleatorias  $\{X(t)\}$  dependientes del tiempo  $t$ . El tiempo puede ser discreto, por ejemplo,  $t = 0, 1, 2, \dots$ , o continuo,  $t \geq 0$ . En cualquier momento  $t$  describe la observación de una variable aleatoria que denotaremos por  $X_t$  o  $X(t)$ .

En la práctica, observamos solamente un resultado del proceso, una trayectoria, de una gran cantidad de posibles trayectorias. Una trayectoria es una función del tiempo  $t$ ,  $x_t = x(t)$ ,  $0 \leq t \leq T$ . Para hacer cálculos de un futuro incierto, se necesita conocer las propiedades colectivas de todas las trayectorias lo que se logra con las distribuciones de probabilidad de la variable aleatoria  $X(t)$ .

Una sucesión  $\{X_n\}_{n \geq 0}$  de variables aleatorias con valores en un conjunto  $E$  se llama un *proceso estocástico discreto* con *espacio de estados*  $E$ , contable, y sus elementos se denotan por  $i, j, k, \dots$ . Si  $X_n = i$ , se dice que el proceso está en

el estado  $i$  al tiempo  $n$ .

**Definición 1.1.1 (Cadena de Markov Homogenea).** Sea  $\{X_n\}_{n \geq 0}$  un proceso estocastico discreto con espacio de estados contable  $E = \{i, j, k, \dots\}$ . Si para todos los enteros  $n \geq 0$  y todos los estados  $i_0, i_1, \dots, i_{n-1}, i, j$ ,

$$\mathbb{P}(X_{n+1} = j | X_n = i, X_{n-1} = i_{n-1}, \dots, X_0 = i_0) = \mathbb{P}(X_{n+1} = j | X_n = i) \quad (1.1)$$

donde ambos lados estan bien de nidos, entonces este proceso estocastico es llamado una *cadena de Markov* o un *proceso de Markov*, y se le llama una *cadena homogenea* si el lado derecho de (1.1) es independiente de  $n$ .

La propiedad (1.1) es la *propiedad de Markov*. La matriz  $\mathbf{P} = (p_{ij})_{i,j \in E}$ , donde

$$p_{ij} = \mathbb{P}(X_{n+1} = j | X_n = i), \quad (1.2)$$

es la *matriz de transición* de la cadena homogenea de Markov. Como sus entradas son probabilidades, y como la transicion de cualquier estado  $i$  debe estar en *algún* estado, se sigue que

$$p_{ij} \geq 0, \quad \sum_{j \in E} p_{ij} = 1$$

para todos los estados  $i, j$ . El espacio de estados puede ser infinito, y por lo tanto tal matriz no es en general de las estudiadas en algebra lineal. Sin embargo, satisfacen las operaciones basicas de suma y multiplicacion. Así,  $p_i = (p_{ij})_{j \in E}$  es un vector de probabilidades para cada  $i$

**Definición 1.1.2 (Distribucion de una cadena de Markov homogenea).** A la variable aleatoria  $X_0$  se le conoce como *estado inicial*, y a su distribucion de probabilidad  $\nu$ ,

$$\nu(i) = \mathbb{P}(X_0 = i) \quad (1.3)$$

como su *distribución inicial*. De la definición de probabilidad condicional, de la propiedad de Markov, de la propiedad de homogeneidad de la cadena y de la definición de matriz de transicion:

$$\mathbb{P}(X_0 = i_0, X_1 = i_1, \dots, X_k = i_k) = \nu(i_0)p_{i_0 i_1} \dots p_{i_{k-1} i_k}. \quad (1.4)$$

La ecuacion (1.4) es la *ley de probabilidad* o la distribucion finito dimensional de la cadena homogenea de Markov.

La distribucion al tiempo  $n$  de la cadena es el vector  $\nu_n$ , donde

$$\nu_n(i) = \mathbb{P}(X_n = i).$$

Del teorema de probabilidades totales,  $\nu_{n+1}(j) = \sum_{i \in E} \nu_n(i)p_{ij}$ , que en forma matricial es  $\nu_{n+1}^\top = \nu_n^\top \mathbf{P}$ . Iterando esta ecuacion se tiene

$$\nu_n^\top = \nu_0^\top \mathbf{P}^n. \quad (1.5)$$

La matriz  $\mathbf{P}^n$  se llama *matriz de transición de  $n$  pasos* con termino general

$$p_{ij}(m) = \mathbb{P}(X_{m+n} = j | X_n = i).$$

Ahora, observando que  $\mathbb{P}(X_{m+n} = j | X_n = i) = \sum_{k \in E} \mathbb{P}(X_{m+n} = j, X_{n+1} = k | X_n = i)$  y de la propiedad de Markov, se encuentra que el lado derecho de esta ecuacion es

$$\sum_{i_1 \dots i_{n-1} \in E} p_{ii_1} p_{i_1 i_2} \dots p_{i_{n-1} j},$$

que es el termino general de la  $n$ -esima potencia de  $\mathbf{P}$ .

**Definicion 1.1.3 (Comunicacion de estados).** Se dice que el estado  $j$  es *accesible* desde  $i$  si existe  $M \geq 0$  tal que  $p_{ij}(M) > 0$ . En particular, un estado  $i$  es siempre accesible desde s mismo, ya que  $p_{ii}(0) = 1$ . Se dice que los estados  $i$  y  $j$  *se comunican* si  $i$  es accesible desde  $j$  y  $j$  es accesible desde  $i$ , y lo denotamos por  $i \leftrightarrow j$ .

Claramente, la comunicacion entre estados es *reflexiva, simétrica y transitiva*; por tanto la relacion de comunicacion ( $\leftrightarrow$ ) es una relacion de equivalencia, y genera una particion del espacio de estados  $E$  en clases de equivalencia disjuntas llamadas *clases de comunicaci3n*.

**Definicion 1.1.4 (Clase irreducible).** Si existe solo una clase de comunicacion, entonces la cadena y su matriz de trasi3n se diran *irreducibles*.

**Definicion 1.1.5 (Per odo).** El *período* del estado  $i$  se define como el maximo comun divisor de todos los enteros  $n \geq 1$ , para los cuales  $p_{ii}(n) > 0$ . Cuando el per odo es 1, decimos que el estado es aperiodico.

**Teorema 1.1.1 (Propiedad fuerte de Markov).** Sea  $\{X_n\}_{n \geq 0}$  una cadena de Markov homogénea con espacio de estados contable  $E$  y matriz de transición  $\mathbf{P}$ . Sea  $\tau$  un tiempo de paro con respecto a esta cadena. Entonces para cualquier estado  $i \in E$ , si  $X_\tau = i$ , se cumple lo siguiente:

(a) El proceso después de  $\tau$  y el proceso antes de  $\tau$  son independientes:

$$\mathbb{P}(X_{\tau+k} = j | X_\tau = i, X_{\tau \wedge n} = i_n) = \mathbb{P}(X_{\tau+k} = j | X_\tau = i).$$

(b) El proceso después de  $\tau$  es una cadena de Markov homogénea:

$$\begin{aligned} \mathbb{P}(X_{\tau+n+1} = k | X_{\tau+n} = j, X_{\tau+n-1} = i_{n-1}, \dots, X_\tau = i) \\ = \mathbb{P}(X_{\tau+n+1} = k | X_{\tau+n} = j) = p_{jk} \end{aligned}$$

con matriz de transición  $\mathbf{P}$ .

**Definicion 1.1.6 (Distribucion Estacionaria).** Una distribucion de probabilidad  $\boldsymbol{\pi}$  que satisface

$$\boldsymbol{\pi}^\top = \boldsymbol{\pi}^\top \mathbf{P} \tag{1.6}$$

se llama *distribuci3n estacionaria* de la matriz de transicion  $\mathbf{P}$ .

Iterando (1.6) se obtiene  $\boldsymbol{\pi}^\top = \boldsymbol{\pi}^\top \mathbf{P}^n$  para todo  $n \geq 0$ . Si la distribución inicial es  $\boldsymbol{\nu} = \boldsymbol{\pi}$ , entonces  $\boldsymbol{\nu}_n = \boldsymbol{\pi}$  para toda  $n \geq 0$ . Así, si la cadena comienza con una distribución estacionaria, mantiene la misma distribución todo el tiempo. Aun más,

$$\mathbb{P}(X_n = i_0, X_{n+1} = i_1, \dots, X_{n+k} = i_k) = \pi(i_0)p_{i_0 i_1} \dots p_{i_{k-1} i_k}$$

no depende de  $n$ . En este sentido la cadena es *estacionaria*.

En la teoría de las cadenas de Markov, un tiempo de paro (ver apéndice A.1.4) importante es el *tiempo de regreso* al estado  $i$ ,

$$T_i = \inf\{n \geq 1 : X_n = i\},$$

donde  $T_i = \infty$  si  $X_n \neq i$  para todo  $n \geq 1$ .

Observe que  $T_i \geq 1$ , y en particular,  $X_0 = i$  no implica que  $T_i = 0$ . Esta es la razón por la cual a  $T_i$  se le conoce como el tiempo de *regreso* a  $i$ , y no como el tiempo de llegada (*hitting time*) a  $i$ . Este último es  $S_i = T_i$  si  $X_0 \neq i$ , y  $S_i = 0$  si  $X_0 = i$ .

Asociado a cada estado está el número de visitas al estado  $i$  estrictamente después del tiempo cero

$$N_i = \sum_{n \geq 1} I_{\{X_n = i\}}.$$

La siguiente definición identifica a los estados con su naturaleza interna. Sea

$$f_{ii}^{(n)} = \mathbb{P}(X_n = i, X_r \neq i \ (r = 1, 2, \dots, n-1) | X_0 = i)$$

y

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

Es decir,  $f_{ii}^{(n)}$  es la probabilidad que, comenzando en el estado  $i$ , el proceso regrese a  $i$  por primera vez en  $n$  pasos, y  $f_{ii}^*$  es la probabilidad que, comenzando en  $i$ , el regreso al estado inicial ocurra en un tiempo finito.

**Definición 1.1.7 (Estado recurrente).** Un estado  $i \in E$  es *recurrente* si  $f_{ii}^* = 1$ , y es *transitorio* si  $f_{ii}^* < 1$ .

Cuando el estado  $i$  es recurrente, definimos

$$\mu_i = \mu_{ii} = \sum_{n=1}^{\infty} n f_{ii}^{(n)}$$

que es la esperanza matemática del número de pasos que se requieren para el primer regreso al estado  $i$ . El número de pasos que se requieren para el primer retorno al mismo estado se llama el *tiempo de recurrencia*;  $\mu_i$ , entonces, puede ser llamado *el tiempo de recurrencia promedio* del estado  $i$ .

**Teorema 1.1.2.** *Si la cadena de Markov es irreducible y aperiódica, entonces para cualesquiera par de estados  $i$  y  $j$*

$$\lim_{n \rightarrow \infty} p_{ij}(n) = \pi(j)$$

*existe y es independiente de  $i$ , además  $\pi(j) = \mu_j^{-1}$ .*

El siguiente resultado que consideraremos concierne al comportamiento límite de las sumas normalizadas de las probabilidades de transición de  $n$  pasos. El teorema se debe a Kolmogorov.

**Teorema 1.1.3 (Teorema Ergodico).** *Considere una cadena de Markov arbitraria con estados numerables; entonces el límite*

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n p_{ij}(k) = \frac{1}{\mu_j} = \pi(j)$$

*existe para cualquier par de estados  $i$  y  $j$ .*

Podemos interpretar a  $\pi(j)$  como la fracción de tiempo que el proceso se mantiene en el estado  $j$  a lo largo del tiempo.

## 1.2. Cadenas de Markov Continuas

### 1.2.1. Proceso Puntual

Un *proceso puntual aleatorio* es un conjunto aleatorio contable de puntos en la recta real. En la mayoría de las aplicaciones un *punto* de un proceso puntual es el tiempo en que ocurre algún evento, y este es el motivo por el cual los puntos son también llamados *eventos*. Por ejemplo, los tiempos de llegada de los clientes a un mostrador, las tareas que llegan a la unidad de procesamiento central en una computadora son eventos de un proceso puntual. En biología, un evento puede ser el tiempo de nacimiento de un organismo. En general, los procesos puntuales en la línea real aparecen en los modelos estocásticos donde el estado del sistema cambia con la llegada de un evento. El ejemplo más común es el Proceso Poisson y la cadena de Markov continua.

**Definición 1.2.1 (Proceso Puntual Aleatorio).** Un *proceso puntual aleatorio* en la recta real positiva es una sucesión  $\{T_n\}_{n \geq 0}$  de variables aleatorias no negativas tales que, casi seguramente,

- a)  $T_0 \doteq 0$
- b)  $0 < T_1 < T_2 < \dots$ ,
- c)  $\lim_{n \rightarrow \infty} T_n = +\infty$

La sucesion  $\{S_n\}_{n \geq 0}$  de nida por

$$S_n = T_{n+1} - T_n$$

es llamada la sucesion de interarribos. Para cualquier intervalo  $(a, b]$  en  $\mathbb{R}_+$ ,

$$N((a, b]) \doteq \sum_{n \geq 1} I_{(a, b]}(T_n)$$

es una variable aleatoria en los enteros que cuenta los eventos ocurridos en el intervalo  $(a, b]$ .

### 1.2.2. Proceso Poisson

**Definicion 1.2.2 (Proceso Poisson Homogeneo).** Un proceso puntual  $N$  en la recta real positiva es llamado *un proceso Poisson Homogéneo* con *intensidad*  $\lambda > 0$  si

- Para todos los tiempos  $t_i, i \in [1, k]$ , tal que  $0 \leq t_1 \leq t_2 \leq \dots \leq t_k$ , las variables aleatorias  $N(t_i, t_{i+1}), i \in [1, k-1]$ , son independientes.
- Para cualquier intervalo  $(a, b] \subset \mathbb{R}_+$ ,  $N(a, b]$  es una variable aleatoria Poisson con media  $\lambda(b-a)$ .

As , para todo  $k \geq 0$ ,

$$\mathbb{P}(N(a, b] = k) = e^{-\lambda(b-a)} \frac{[\lambda(b-a)]^k}{k!}$$

y, en particular,

$$\mathbb{E}N(a, b] = \lambda(b-a).$$

En este sentido,  $\lambda$  es la densidad promedio de puntos.

La condicion a) es la propiedad de *independencia de incrementos* del proceso Poisson, en particular, implica que para cualquier intervalo  $(a, b]$ , la variable aleatoria  $N(a, b]$  es independiente de  $N(s)$ ,  $s \in (0, a]$ . Por esta razon, se dice que los procesos Poisson no tienen memoria: los incrementos de un proceso Poisson homogeneo no tienen memoria del pasado. Tambien se observa que el proceso Poisson no permite la acumulacion de puntos o explosiones; a) y b) de la de nicion 1.2.2 impiden que haya multiples puntos o explosiones.

**Teorema 1.2.1.** *La sucesión de interarribos  $\{S_n\}_{n \geq 0}$  de un proceso Poisson homogéneo con intensidad  $\lambda > 0$  es independiente e idénticamente distribuida, con distribución exponencial de parámetro  $\lambda$ .*

Otra cantidad de interes es  $W_n$ , el tiempo de arribo del  $n$ -esimo evento, tambien llamado *tiempo de espera* del  $n$ -esimo evento. Entonces,

$$W_n = \sum_{i=1}^n S_i$$

y por el teorema 1.2.1  $W_n$  tiene un distribucion gamma con parametros  $n$  y  $\lambda$ .

**Teorema 1.2.2.** Sea  $\{N_i\}_{i \geq 1}$  una familia de Procesos Poisson independientes con intensidades positivas  $\{\lambda_i\}_{i \geq 1}$ . Entonces

- a) dos procesos distintos de esta familia no tienen puntos en común, y  
 b) si

$$\sum_{i=1}^{\infty} \lambda_i = \lambda < \infty,$$

entonces

$$N(t) \doteq \sum_{i=1}^{\infty} N_i(t)$$

define un proceso de conteo del proceso Poisson con intensidad  $\lambda$ .

El siguiente resultado muestra a los procesos Poisson compitiendo por la producción del primer evento.

**Teorema 1.2.3 (Teorema de Competencia).** Dadas las condiciones del teorema 1.2.2, sea  $Z$  el primer evento de  $N(t) \doteq \sum_{i=1}^{\infty} N_i(t)$  y  $J$  el índice del proceso Poisson que lo produjo; es decir,  $Z$  es el primer evento de  $N_J$ . Entonces

$$\mathbb{P}(J = i, Z \geq a) = \mathbb{P}(J = i) \mathbb{P}(Z \geq a) = \frac{\lambda_i}{\lambda} e^{-\lambda a}.$$

En particular,  $J$  y  $Z$  son independientes,  $\mathbb{P}(J = i) = \frac{\lambda_i}{\lambda}$ , y  $Z$  es exponencial con media  $\lambda^{-1}$ .

### 1.2.3. Distribución de una cadena de Markov Homogénea

El enfoque tradicional de las cadenas de Markov continuas se basa en el *semigrupo de transición*, y entonces el objeto matemático principal es el *generador infinitesimal*. El semigrupo de transición es el análogo continuo a la matriz de transición para el tiempo discreto.

**Definición 1.2.3 (Cadenas de Markov en tiempo Continuo).** El proceso estocástico  $\{X_t\}_{t \geq 0}$ , indexado por  $\mathbb{R}_+$ , se llama *cadena de Markov en tiempo continuo* si para todo  $i, j, i_1, \dots, i_k \in E$ ; todo  $t, s \geq 0$ ; y todo  $s_1, \dots, s_k \geq 0$  con  $s_l \leq s$  para  $l \in [1, k]$ ,

$$\begin{aligned} \mathbb{P}(X(t+s) = j | X(s) = i, X(s_1) = i_1, \dots, X(s_k) = i_k) \\ = \mathbb{P}(X(t+s) = j | X(s) = i), \end{aligned} \quad (1.7)$$

donde ambos lados están bien definidos, y es homogénea si el lado derecho de (1.7) es independiente de  $s$ .

Sea pues

a)  $\mathbf{P}(t) \doteq \{p_{ij}(t)\}_{i,j \in E}$ ,

donde

$$p_{ij}(t) \doteq \mathbb{P}(X(t+s) = j | X(s) = i).$$

b) De forma similar a las cadenas discretas de Markov, obtenemos las ecuaciones de Chapman-Kolmogorov

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

que en forma compacta se lee

$$\mathbf{P}(t+s) = \mathbf{P}(t)\mathbf{P}(s).$$

c) Claramente  $\mathbf{P}(0) = \mathbf{I}$ , siendo  $\mathbf{I}$  la matriz identidad.

La familia  $\{\mathbf{P}(t)\}_{t \geq 0}$  es el *semigrupo de transición* de la cadena continua de Markov.

La distribución al tiempo  $t$  de  $X(t)$  es el vector  $\boldsymbol{\mu}(t) = \{\mu_i(t)\}_{i \in E}$  donde  $\mu_i(t) = \mathbb{P}(X(t) = i)$ . Esta se obtiene de la distribución inicial por la fórmula

$$\boldsymbol{\mu}(t)^\top = \boldsymbol{\mu}(0)^\top \mathbf{P}(t).$$

#### 1.2.4. Matriz de Intensidad

**Teorema 1.2.4.** *Sea  $\{\mathbf{P}(t)\}_{t \geq 0}$  un semigrupo de transición continuo en el espacio de estados contable  $E$ . Para cualquier estado  $i$ , existe*

$$q_i \doteq \lim_{h \rightarrow 0} \frac{1 - p_{ii}(h)}{h} \in [0, \infty],$$

y para cualquier pareja  $i, j$  de estados diferentes, existe

$$q_{ij} \doteq \lim_{h \rightarrow 0} \frac{p_{ij}(h)}{h} \in [0, \infty).$$

Además se tiene para cada estado  $i$

$$q_{ii} \doteq -q_i.$$

**Definición 1.2.4 (Matriz de Intensidad o Generador Infinitesimal).** La matriz

$$\mathbf{\Lambda} = \{q_{ij}\}_{i,j \in E}$$

es llamada *matriz de intensidad* o *generador infinitesimal* del semigrupo.

En notación compacta

$$\mathbf{\Lambda} = \lim_{h \rightarrow 0} \frac{\mathbf{P}(h) - \mathbf{P}(0)}{h}.$$

Así, en este sentido, la matriz de intensidad  $\mathbf{\Lambda}$  es la derivada en 0 de la función matricial  $t \rightarrow \mathbf{P}(t)$ . Por esta razón  $\mathbf{\Lambda}$  recibe el nombre de generador infinitesimal o matriz de intensidad.



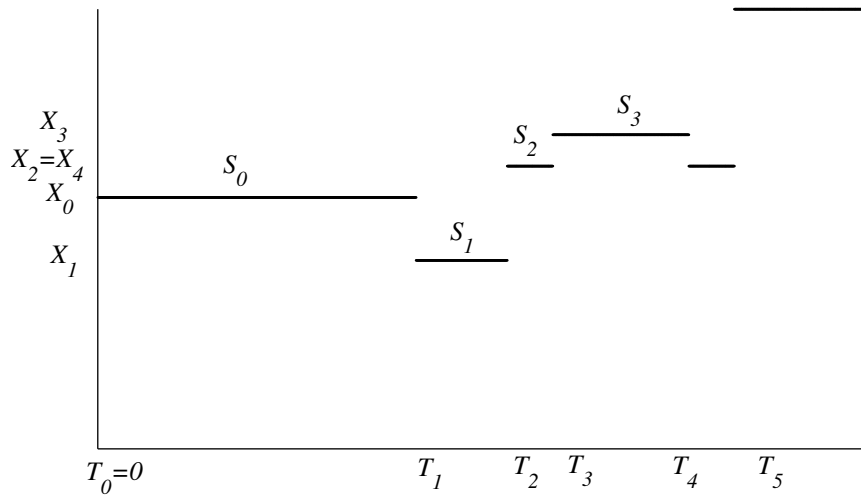


Figura 1.1: Trayectorias de un proceso de brincos Markoviano

**Definición 1.2.5 (Conservatividad y Estabilidad).** Si para cualquier estado  $i$

$$q_i < \infty,$$

entonces se llama al semigrupo  $\{\mathbf{P}(t)\}$  *estable*, y si para para cualquier estado  $i$ ,

$$q_i = \sum_{\substack{j \in E \\ j \neq i}} q_{ij} \tag{1.8}$$

se le llama *conservador*.

La razon de este nombre viene de la igualdad

$$\sum_{j \in E} p_{ij}(h) = 1,$$

o equivalentemente,

$$\frac{1 - p_{ii}(h)}{h} = \sum_{\substack{j \in E \\ j \neq i}} \frac{p_{ij}(h)}{h},$$

lo que conduce a

$$q_i = \lim_{h \rightarrow 0} \sum_{\substack{j \in E \\ j \neq i}} \frac{p_{ij}(h)}{h}.$$

y si se permite el intercambio del límite y suma se obtiene (1.8).

Como caso interesante, los procesos de brinco son estables y conservativos. Nos enfocaremos ahora en la estructura del *brinco* Markoviano.

**Definición 1.2.6 (Brinco Markoviano).** Un proceso estocástico  $\{X_t\}_{t \geq 0}$  que toma valores en un espacio de estados  $E$  se llama *brinco Markoviano* si para casi todo  $\omega \in \Omega$  y todo  $t \geq 0$  (ver de nición A.1.1), existe  $\epsilon(t, \omega) > 0$  tal que

$$X(t+s, \omega) = X(t, \omega) \quad \forall s \in [t, t + \epsilon(t, \omega)).$$

Un proceso de brinco Markoviano es por de nición una cadena de Markov continua, tambien observamos que para un brinco Markoviano, existe una sucesion tiempos  $\{T_n\}_{n \geq 0}$  donde

$$T_0 = 0 < T_1 < T_2 < \dots$$

y una sucesion  $\{X_n\}_{n \geq 0}$  tal que

$$X(t) = X_n \quad \text{si} \quad T_n \leq t < T_{n+1}.$$

As queda descrito el proceso  $\{X_t\}_{t \geq 0}$  en el intervalo  $[0, T_\infty)$ , donde

$$T_\infty = \lim_{n \rightarrow \infty} T_n$$

es el tiempo de *explosión*.

El tiempo de espera en cada estado es positivo, razon por la que las trayectorias son constantes por pedazos, ver Figura 1.1. Para un proceso de brinco Markoviano, denotamos los tiempos de brinco por  $T_0 = 0 < T_1 < T_2 < \dots$ , los tiempos de espera (*sojourn times*) por  $S_n = T_{n+1} - T_n$  y la sucesion de estados visitados por  $X_0, X_1, X_2, \dots$ . As, las trayectorias son constantes entre  $T_n$  consecutivos y se de ne el valor al tiempo  $T_n$  por continuidad a la derecha, es decir,  $X(T_n) = X_n$ .

**Teorema 1.2.5.** *Un proceso de brinco Markoviano es estable y conservador.*

### 1.2.5. Sistema Diferencial de Kolmogorov

Dadas las propiedades de semigrupo, para todo  $t \geq 0$  y todo  $h \geq 0$

$$\frac{\mathbf{P}(t+h) - \mathbf{P}(t)}{h} = \mathbf{P}(t) \frac{\mathbf{P}(h) - \mathbf{I}}{h} = \frac{\mathbf{P}(h) - \mathbf{I}}{h} \mathbf{P}(t). \quad (1.9)$$

Por tanto, si se puede tomar el l mite de (1.9), que es el caso cuando  $E$  es finito, se obtiene el sistema diferencial

$$\frac{d}{dt} \mathbf{P}(t) = \mathbf{P}(t) \mathbf{\Lambda} = \mathbf{\Lambda} \mathbf{P}(t),$$

donde  $\mathbf{\Lambda}$  es el generador infinitesimal. La ecuacion

$$\frac{d}{dt} \mathbf{P}(t) = \mathbf{\Lambda} \mathbf{P}(t) \quad (1.10)$$

se puede escribir expl citamente. Para todo  $i, j \in E$ ,

$$\frac{d}{dt}p_{ij}(t) = -q_i p_{ij}(t) + \sum_{\substack{k \in E \\ k \neq i}} q_{ik} p_{kj}(t).$$

El sistema (1.10) se llama la *ecuación diferencial regresiva de Kolmogorov*. La *ecuación diferencial progresiva* es

$$\frac{d}{dt}\mathbf{P}(t) = \mathbf{P}(t)\mathbf{\Lambda}, \quad (1.11)$$

es decir para todo  $i, j \in E$ ,

$$\frac{d}{dt}p_{ij}(t) = -p_{ij}(t)q_j + \sum_{\substack{k \in E \\ k \neq j}} p_{ik}(t)q_{kj}.$$

**Teorema 1.2.6 (Ecuacion Regresiva).** *Si el semigrupo continuo  $\{\mathbf{P}(t)\}_{t \geq 0}$  es estable y conservador, entonces se satisface la ecuación diferencial regresiva de Kolmogorov (1.10).*

**Teorema 1.2.7 (Ecuacion Progresiva).** *Bajo las condiciones del teorema 1.2.6 y si, además, para todos los estados  $i$  y todo  $t \geq 0$ ,*

$$\sum_{k \in E} p_{ik}(t)q_k < \infty, \quad (1.12)$$

*entonces se satisface la ecuación diferencial progresiva de Kolmogorov (1.11).*

La condicion (1.12) se satisface trivialmente cuando el espacio de estados es nito, o cuando

$$\sup_{i \in E} q_i < \infty.$$

La distribucion al tiempo  $t$  de la cadena es el vector columna

$$\boldsymbol{\mu}(t) = \{\mu_i(t)\}_{i \in E} = \{\mathbb{P}(X(t) = i)\}_{i \in E}.$$

Y satisface para todo  $t, s \geq 0$ ,

$$\boldsymbol{\mu}^\top(t+s) = \boldsymbol{\mu}^\top(t)\mathbf{P}(s),$$

es decir, para todo  $i, j \in E$ ,

$$\mu_i(t+s) = \sum_{j \in E} \mu_j(t)p_{ji}(s).$$

**Teorema 1.2.8.** *Bajo las condiciones del teorema 1.2.6, y si además, para todo  $t \geq 0$ ,*

$$\sum_{k \in E} q_k \mu_k(t) < \infty,$$

*entonces el sistema diferencial global de Kolmogorov*

$$\frac{d}{dt} \boldsymbol{\mu}^\top(t) = \boldsymbol{\mu}^\top(t) \mathbf{\Lambda}$$

*se satisface, es decir, para todo  $i, j \in E$ ,*

$$\frac{d}{dt} \mu_i(t) = -\mu_i(t) q_i + \sum_{\substack{j \in E \\ j \neq i}} \mu_j(t) q_{ji}.$$

Si  $\boldsymbol{\pi}$  es una distribución estacionaria de una cadena continua, estable y conservadora, y si  $\sum_{i \in E} \pi(i) q_i < \infty$ , entonces, según el teorema 1.2.8,  $\boldsymbol{\pi}$  satisface el sistema diferencial de Kolmogorov

$$\boldsymbol{\pi}^\top \mathbf{\Lambda} = 0,$$

es decir, en forma expandida,

$$\pi(i) q_i = \sum_{\substack{j \in E \\ j \neq i}} \pi(j) q_{ji}.$$

*Resumiendo:* Para una cadena de Markov continua con espacio de estados finito  $E$  y matriz de intensidad  $\mathbf{\Lambda}$ , la condición  $\boldsymbol{\pi}^\top \mathbf{\Lambda} = 0$  es necesaria y suficiente para que la distribución de probabilidad  $\boldsymbol{\pi}$  sea una distribución estacionaria.

Si el espacio de estados es finito, una solución de (1.10) o (1.11) con condición inicial  $\mathbf{P}(0) = \mathbf{I}$  es

$$\mathbf{P}(t) = e^{t\mathbf{\Lambda}},$$

donde la exponencial de la matriz  $t\mathbf{\Lambda}$  (pensada como un número) está definida por

$$e^{t\mathbf{\Lambda}} \doteq \sum_{n=0}^{\infty} \frac{t^n}{n!} \mathbf{\Lambda}^n = \mathbf{I} + \sum_{n=1}^{\infty} \frac{t^n}{n!} \mathbf{\Lambda}^n. \quad (1.13)$$

Cuando  $\mathbf{\Lambda}$  es una matriz finita, la serie en (1.13) es convergente, y es solución única a las ecuaciones regresiva y progresiva. Sin embargo, cuando  $\mathbf{\Lambda}$  es de dimensión infinita, no se puede decir mucho, pues la serie puede no converger del todo. Cuando  $\mathbf{\Lambda}$  es una matriz finita y es diagonalizable, podemos escribir a  $\mathbf{\Lambda}$  en la forma

$$\mathbf{\Lambda} = \mathbf{VDV}^{-1}$$

donde  $\mathbf{D}$  es una matriz diagonal con diferentes eigenvalores  $\{\lambda_1, \lambda_2, \dots, \lambda_n\}$  de  $\mathbf{\Lambda}$  como elementos. As, tenemos

$$\begin{aligned} \mathbf{P}(t) = e^{t\mathbf{V}\mathbf{D}\mathbf{V}^{-1}} &= \mathbf{I} + \sum_{n=1}^{\infty} \frac{\mathbf{V}(\mathbf{D}t)^n \mathbf{V}^{-1}}{n!} \\ &= \mathbf{V} \left[ \mathbf{I} + \sum_{n=1}^{\infty} \frac{(\mathbf{D}t)^n}{n!} \right] \mathbf{V}^{-1} \\ &= \mathbf{V} e^{\mathbf{D}t} \mathbf{V}^{-1}, \end{aligned} \quad (1.14)$$

donde

$$e^{\mathbf{D}t} = \begin{pmatrix} e^{\lambda_1 t} & & 0 \\ & e^{\lambda_2 t} & \\ & & \ddots \\ 0 & & & e^{\lambda_n t} \end{pmatrix}.$$

**Teorema 1.2.9.** Sea  $\{X(t)\}_{t \geq 0}$  un brinco Markoviano con espacio de estados contable  $E$  y semigrupo de transición  $\{\mathbf{P}(t)\}_{t \geq 0}$ . Entonces

- El semigrupo  $\{\mathbf{P}(t)\}_{t \geq 0}$  es continuo, estable y conservador.
- Se satisfacen las dos ecuaciones diferenciales de Kolmogorov.
- Una condición necesaria y suficiente para que una distribución de probabilidad  $\pi$  en  $E$  sea distribución estacionaria es que  $\pi^t \mathbf{\Lambda} = 0$ .

**Ejemplo 1.2.1.** Suponga que  $E$  tiene solamente dos estados 1, 2 y que tanto  $\lambda(1)$  como  $\lambda(2)$  no son cero. La igualdad  $0 = |\mathbf{\Lambda} - \lambda I|$ , implica que  $\lambda = 0$  y  $\lambda = -\lambda(1) - \lambda(2) \neq 0$ , con correspondientes eigenvectores  $(1, 1)^T$ ,  $(\lambda(1), \lambda(2))^T$ . Entonces podemos escribir a  $\mathbf{\Lambda}$  como

$$\mathbf{\Lambda} = \begin{pmatrix} -\lambda(1) & \lambda(1) \\ \lambda(2) & -\lambda(2) \end{pmatrix} = \mathbf{V}\mathbf{D}\mathbf{V}^{-1},$$

donde

$$\mathbf{V} = \begin{pmatrix} 1 & \lambda(1) \\ 1 & -\lambda(2) \end{pmatrix}, \quad \mathbf{V}^{-1} = \frac{1}{-\lambda} \begin{pmatrix} \lambda(2) & \lambda(1) \\ 1 & -1 \end{pmatrix}, \quad \mathbf{D} = \begin{pmatrix} 0 & 0 \\ 0 & \lambda \end{pmatrix} = \text{diag}\{0, \lambda\}.$$

De esta forma  $\mathbf{\Lambda}^n = \mathbf{V}\mathbf{D}^n \mathbf{V}^{-1}$ . Así,

$$\mathbf{P}(t) = e^{t\mathbf{\Lambda}} = \mathbf{V} \text{diag}\{e^{t0}, e^{t\lambda}\} \mathbf{V}^{-1},$$

es decir

$$\mathbf{P}(t) = \frac{1}{\lambda(1) + \lambda(2)} \begin{pmatrix} \lambda(2) + \lambda(1)e^{\lambda t} & \lambda(1) - \lambda(1)e^{\lambda t} \\ \lambda(2) - \lambda(2)e^{\lambda t} & \lambda(1) + \lambda(2)e^{\lambda t} \end{pmatrix}.$$

### 1.2.6. Cadena Incluida y Tiempos de transición

Sea  $\{\tau_n\}_{n \geq 1}$  una sucesión de tiempos de transición de un proceso de brincos Markoviano  $\{X(t)\}_{t \geq 0}$  donde  $\tau_0 = 0$ ; y  $\tau_n = \infty$  si hay estrictamente menos de  $n$  transiciones en  $(0, \infty)$ . Ahora, para cada  $n \geq 0$   $\tau_n$  es un tiempo de paro con respecto a  $\{X(t)\}_{t \geq 0}$ .

El proceso  $\{X_n\}_{n \geq 0}$  con valores en  $E_\Delta = E \cup \{\Delta\}$ , donde  $\Delta$  es un elemento arbitrario que no está en  $E$ , se define por

$$X_n = X(\tau_n),$$

con la convención de que  $X(\infty) = \Delta$ , y es llamado la *cadena incluida* del proceso de brincos.

La importancia del siguiente teorema radica en la facilidad de encontrar la matriz de transición  $\mathbf{P}$  y mostrar que los tiempos de permanencia en los estados tiene una distribución exponencial.

**Teorema 1.2.10 (Estructura Regenerativa).** *Sea  $\{X(t)\}_{t \geq 0}$  un proceso de brincos Markoviano con generador infinitesimal  $\mathbf{A}$ , sucesión de tiempos de transición  $\{\tau_n\}_{n \geq 1}$  y cadena incluida  $\{X_n\}_{n \geq 0}$ . Entonces*

- a)  $\{X_n\}_{n \geq 0}$  es una cadena de Markov homogénea con espacio de estados  $E_\Delta$  con matriz de transición dada por  $p_{\Delta\Delta} = 1$ ,  $p_{i\Delta} = 1$  si  $i \in E$  y  $q_i = 0$ ; por  $p_{i\Delta} = 0$  si  $i \in E$  y  $q_i > 0$ ; y si  $q_i > 0$  y  $j \neq i$ , por

$$p_{ij} = \frac{q_{ij}}{q_i}.$$

- b) Dado  $\{X_n\}_{n \geq 0}$ , la sucesión  $\{\tau_{n+1} - \tau_n\}_{n \geq 0}$  es independiente, y para todo  $n \geq 0$  y  $a \in \mathbb{R}_+$ ,

$$\mathbb{P}(\tau_{n+1} - \tau_n \leq a | \{X_k\}_{k \geq 0}) = 1 - e^{-q_{X_n} a}. \quad (1.15)$$

Identificamos al parámetro  $q_i$  como el parámetro de intensidad de la distribución exponencial cuando el proceso estando en el estado  $i$  al tiempo  $t$ , sale de  $i$  antes de  $t + dt$  con probabilidad  $q_i dt$ . El estado siguiente  $j$  es seleccionado independientemente del tiempo de salida del estado  $i$  de acuerdo a  $p_{ij}$ . Así, si el estado  $i$  es recurrente, entonces será visitado un número infinito de veces, y cada visita durará un tiempo exponencial con parámetro  $q_i$ .

**Definición 1.2.7 (Esencial, Permanente).** Un estado  $i \in E$  tal que  $q_i = 0$  se llama *permanente* o absorbente; *esencial* o estable si  $0 < q_i < \infty$ , e instantáneo si  $q_i = \infty$ .

En vista de (1.15), si  $X(\tau_n) = i$ , un estado permanente, entonces  $\tau_{n+1} - \tau_n = \infty$ ; es decir, ya no hay más transiciones en una distancia finita y el estado permanece ahí para siempre, de ahí la terminología. Si  $i$  es esencial o estable permanece ahí un tiempo positivo pero finito. Finalmente, si  $i$  es instantáneo el proceso brincará tan rápido como haya entrado.

**Teorema 1.2.11 (El Generador Infinitesimal Caracteriza al Semigrupo).** *Dos brincos Markovianos con el mismo generador infinitesimal y la misma distribución inicial son probabilísticamente equivalentes; es decir, tienen el mismo semigrupo de transición.*

**Definición 1.2.8 (Generatrices).** Sea  $\mathbf{\Lambda} = \{q_{ij}\}_{i,j \in E}$  una matriz con entradas que satisfacen que para todo  $i, j \in E$ ,

$$q_i \in [0, \infty), \quad q_{ij} \in [0, \infty), \quad \sum_{\substack{k \in E \\ k \neq i}} q_{ik} = q_i,$$

donde  $q_i = -q_{ii}$ . Esta matriz es llamada *generatriz estable y conservativa en E*; y se le llama *generatriz esencial* si, además,  $q_i > 0$  para todo  $i \in E$ .

**Ejemplo 1.2.2 (Cadenas Uniformes de Markov).** *Sea  $\{\hat{X}_n\}_{n \geq 0}$  una cadena de Markov discreta con espacio de estados contable  $E$  y matriz de transición  $\mathbf{K} = \{k_{ij}\}_{i,j \in E}$ ; sea  $\{T_n\}_{n \geq 1}$  un proceso Poisson en  $\mathbb{R}_+$  con intensidad  $\lambda > 0$  y proceso de conteo asociado  $N$ . Suponga que  $\{\hat{X}_n\}_{n \geq 0}$  y  $N$  son independientes. Al proceso  $\{X_t\}_{t \geq 0}$  con valores en  $E$ , y definido por*

$$X(t) = \hat{X}_{N(t)}$$

*se le llama una cadena uniforme de Markov. Al proceso Poisson  $N$  se le conoce como reloj, y a la cadena  $\{\hat{X}_n\}_{n \geq 0}$  se le llama cadena subordinada.*

*Observamos que  $X(T_n) = \hat{X}_n$  para todo  $n \geq 0$ , y que el proceso  $\{X_t\}_{t \geq 0}$  es una cadena continua de Markov. Su semigrupo de transición es*

$$\begin{aligned} p_{ij}(t) &= \mathbb{P}_i(X(t) = j) \\ &= \mathbb{P}_i(\hat{X}_{N(t)} = j) \\ &= \sum_{n=0}^{\infty} \mathbb{P}_i(\hat{X}_n = j, N(t) = n) \\ &= \sum_{n=0}^{\infty} \mathbb{P}_i(N(t) = n) \mathbb{P}_i(\hat{X}_n = j) \\ &= \sum_{n=0}^{\infty} e^{-\lambda t} \frac{(\lambda t)^n}{n!} k_{ij}(n), \end{aligned}$$

*es decir,*

$$\mathbf{P}(t) = \sum_{n=0}^{\infty} e^{-\lambda t} \frac{(\lambda t)^n}{n!} \mathbf{K}^n.$$

*Su matriz de intensidad  $\mathbf{\Lambda}$  se puede calcular observando que*

$$\mathbf{P}(t) = e^{-\lambda t} e^{\lambda t \mathbf{K}},$$

*y como la matriz de intensidad es la derivado en 0*

$$\mathbf{\Lambda} = \lambda e^{-\lambda t} e^{\lambda t \mathbf{K}} (\mathbf{K} - I) |_{t=0}$$

$$\mathbf{\Lambda} = \lambda (\mathbf{K} - I)$$

es decir,

$$q_i = \lambda(1 - k_{ii}),$$

y para  $i \neq j$

$$q_{ij} = \lambda k_{ij};$$

$\lambda$  es la intensidad del reloj.

Su estructura regenerativa se puede expresar para un estado esencial  $i \in E$  como

$$p_{ij} = \frac{q_{ij}}{q_i} = \frac{\lambda k_{ij}}{\lambda(1 - k_{ii})} = \frac{k_{ij}}{1 - k_{ii}} \quad \text{para } j \neq i \text{ con } k_{ii} < 1.$$

**Ejemplo 1.2.3 (Los Procesos de Nacimiento y Muerte).** Un proceso de nacimiento y muerte es un brinco Markoviano que toma valores en  $\mathbb{N} \cup \{0\}$ , y con generador infinitesimal de la forma

$$q_{i,i+1} = \beta_i \quad q_{i,i-1} = \delta_i I_{\{i \geq 1\}},$$

y  $q_{ij} = 0$  si  $j \notin \{i-1, i, i+1\}$ .

$$\mathbf{A} = \begin{pmatrix} -\beta_0 & \beta_0 & 0 & 0 & \dots \\ \delta_1 & -(\beta_1 + \delta_1) & \beta_1 & 0 & 0 & \dots \\ 0 & \delta_2 & -(\beta_2 + \delta_2) & \beta_2 & 0 & \dots \\ 0 & 0 & \delta_3 & -(\beta_3 + \delta_3) & \beta_3 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

Los parámetros  $\beta_i$  y  $\delta_i$  son las intensidades de nacimiento y muerte en el instante en que el tamaño de la población es  $i$ . Un nacimiento incrementa el tamaño de la población en una unidad, una muerte lo decrementa en uno. Como  $p_n(t) = \mathbb{P}(X(t) = n)$  es un vector columna, entonces por el teorema 1.2.8 el sistema diferencial que satisface es,

$$\begin{aligned} \dot{p}_0(t) &= -\beta_0 p_0(t) + \delta_1 p_1(t), \\ \dot{p}_n(t) &= \beta_{n-1} p_{n-1}(t) - (\beta_n + \delta_n) p_n(t) + \delta_{n+1} p_{n+1}(t), \quad \text{para } n \geq 1. \end{aligned}$$

Una condición necesaria y suficiente para que la probabilidad  $\pi$  sea distribución estacionaria del proceso de nacimiento y muerte es que

$$\begin{aligned} 0 &= -\beta_0 \pi(0) - \delta_1 \pi(1), \\ 0 &= \beta_{n-1} \pi(n-1) - (\beta_n + \delta_n) \pi(n) + \delta_{n+1} \pi(n+1), \quad \text{para } n \geq 1. \end{aligned}$$

Para un valor fijo  $\pi(0)$  se encuentra de manera recursiva una solución única del sistema:

$$\pi(n) = \pi(0) \frac{\beta_0 \beta_1 \dots \beta_{n-1}}{\delta_1 \delta_2 \dots \delta_n} \quad \text{para } n \geq 1.$$

Y para que  $\pi$  sea una distribución de probabilidad, se tiene que cumplir

$$\pi(0) \left( 1 + \sum_{n=1}^{\infty} \frac{\beta_0 \beta_1 \dots \beta_{n-1}}{\delta_1 \delta_2 \dots \delta_n} \right) = 1,$$

y esto es posible si y sólo si

$$1 + \sum_{n=1}^{\infty} \frac{\beta_0 \beta_1 \dots \beta_{n-1}}{\delta_1 \delta_2 \dots \delta_n} < \infty.$$



La matriz de transición  $\mathbf{P}$  viene dada por la estructura regenerativa  $p_{i,i+1} = q_{i,i+1}/q_i = \beta_i/(\beta_i + \delta_i)$ ,  $p_{i,i-1} = q_{i,i-1}/q_i = \delta_i/(\beta_i + \delta_i)$  y  $p_{0,1} = 1$ , es decir

$$\mathbf{P} = \begin{pmatrix} 0 & 1 & 0 & 0 & \dots & \dots \\ \frac{\delta_1}{\beta_1 + \delta_1} & 0 & \frac{\beta_1}{\beta_1 + \delta_1} & 0 & 0 & \dots \\ 0 & \frac{\delta_2}{\beta_2 + \delta_2} & 0 & \frac{\beta_2}{\beta_2 + \delta_2} & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

Los procesos de nacimiento y muerte son modelos importantes en biología (de donde obviamente viene la terminología), demografía, pero también en teoría de colas, donde aparecen como colas  $M/M/1/\infty$ ,  $M/M/K/0$ , entre otros modelos de espera.

**Ejemplo 1.2.4.** Considere un proceso de brincos Markoviano con espacio de estados  $E = \{1, 2\}$ , matriz de intensidad

$$\mathbf{\Lambda} = \begin{pmatrix} -\lambda_1 & \lambda_1 \\ \lambda_2 & -\lambda_2 \end{pmatrix}$$

y que inicia en el estado uno con probabilidad uno. Nos interesa calcular la densidad  $h(x)$  del tiempo de permanencia en el estado uno hasta el tiempo  $x$ , en un intervalo de tiempo  $[0, t]$ . La densidad  $h(x)$  se puede descomponer en una densidad  $f(x)$  y la probabilidad de que el proceso permanezca en el mismo estado todo el tiempo hasta  $t$  (no sale del estado uno);  $e^{-\lambda_1 t}$ .

Definimos

$$p_n(x; \lambda) = \frac{\lambda^n x^n}{n!} e^{-\lambda x}$$

$$g_n(x; \lambda) = \frac{\lambda^n x^{n-1}}{(n-1)!} e^{-\lambda x}$$

La probabilidad Poisson con intensidad  $\lambda$  de los primeros  $n$  arribos antes del tiempo  $x$  está representada por  $p_n(x; \lambda)$ , mientras que  $g_n(x; \lambda)$  es la densidad gamma de la suma de  $n$  variables aleatorias exponenciales independientes e idénticamente distribuidas con intensidad  $\lambda$ .

Calcularemos  $f(x)$  mediante el siguiente argumento. Se ha observado un tiempo de ocupación  $x$  en el estado uno y  $n$  transiciones del estado 1 al estado 2. Por tanto ha habido  $n$  ó  $n - 1$  transiciones del estado 2 al estado 1. En el primer caso el proceso estará en el estado 1 y estará en 2 en el otro.

Si hubo  $n$  transiciones de 1 a 2 y el proceso terminó en el estado 1, podemos interpretar un cambio de estado como un arribo, y por tanto se observaron  $n$  arribos al tiempo  $x$ , entonces la densidad es la dada por una Poisson. El tiempo restante  $t - x$  que el proceso pasa en el estado 2 es la suma de  $n$  variables aleatorias exponenciales independientes e idénticamente distribuidas con intensidad  $\lambda_2$ , por tanto con densidad gamma.

Si hubo  $n$  transiciones de 1 a 2 y el proceso terminó en el estado 2, el tiempo  $x$  que el proceso pasa en el estado 1 es la suma de  $n$  variables aleatorias exponenciales independientes e idénticamente distribuidas con intensidad  $\lambda_1$ , y el tiempo de permanencia en el estado 2 es la densidad de una Poisson por argumentos similares.

Así,

$$\begin{aligned}
f(x) &= \sum_{n=1}^{\infty} p_n(x; \lambda_1) g_n(t-x; \lambda_2) + \sum_{n=1}^{\infty} g_n(x; \lambda_1) p_{n-1}(t-x; \lambda_2) \\
&= \sum_{n=1}^{\infty} \frac{\lambda_1^n x^n}{n!} e^{-\lambda_1 x} \frac{\lambda_2^n (t-x)^{n-1}}{(n-1)!} e^{-\lambda_2(t-x)} \\
&\quad + \sum_{n=1}^{\infty} \frac{\lambda_1^n x^{n-1}}{(n-1)!} e^{-\lambda_1 x} \frac{\lambda_2^{n-1} (t-x)^{n-1}}{(n-1)!} e^{-\lambda_2(t-x)} \\
&= e^{-\lambda_1 x} e^{-\lambda_2(t-x)} \left[ \frac{\lambda_1^1 x^1}{1!} \frac{\lambda_2^1 (t-x)^0}{0!} \right. \\
&\quad + \frac{\lambda_1^2 x^2}{2!} \frac{\lambda_2^2 (t-x)^1}{1!} + \frac{\lambda_1^3 x^3}{3!} \frac{\lambda_2^3 (t-x)^2}{2!} + \dots \\
&\quad + \frac{\lambda_1^1 x^0}{0!} \frac{\lambda_2^0 (t-x)^0}{0!} \\
&\quad \left. + \frac{\lambda_1^2 x^1}{1!} \frac{\lambda_2^1 (t-x)^1}{1!} + \frac{\lambda_1^3 x^2}{2!} \frac{\lambda_2^2 (t-x)^2}{2!} + \dots \right] \\
&= e^{-\lambda_1 x} e^{-\lambda_2(t-x)} \left[ \sqrt{\frac{\lambda_1 \lambda_2 x}{t-x}} I_1 \left( 2\sqrt{\lambda_1 \lambda_2 x(t-x)} \right) \right. \\
&\quad \left. + \lambda_1 I_0 \left( 2\sqrt{\lambda_1 \lambda_2 x(t-x)} \right) \right] \tag{1.16}
\end{aligned}$$

donde

$$I_n(x) = \sum_{k=1}^{\infty} \frac{(x/2)^{2k+n}}{k!(n+k)!}$$

es la función de Bessel modificada.

La figura 1.2 muestra a la función  $f$  para distintos valores de  $(\lambda_1, \lambda_2, t)$ .

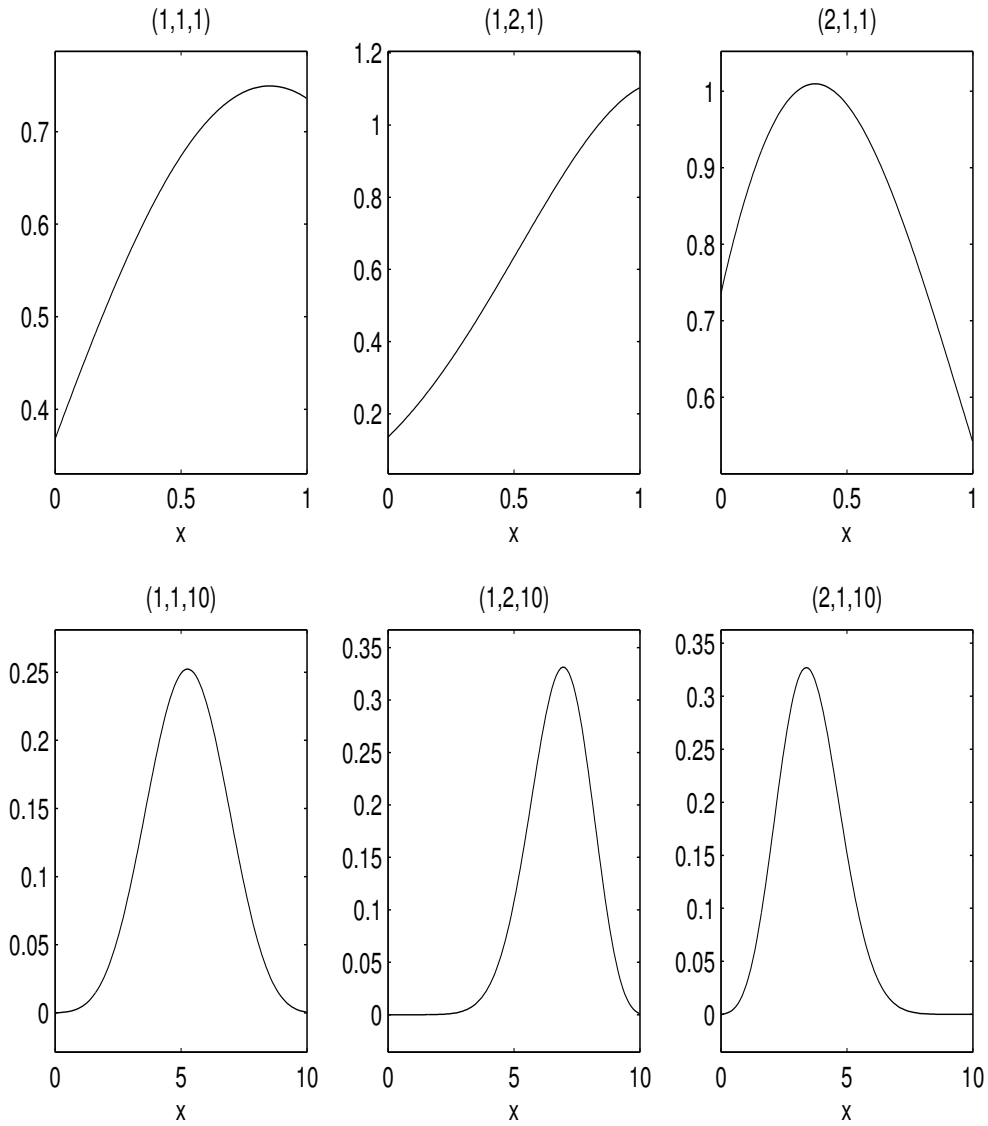


Figura 1.2: Las funciones de densidad  $f(x)$  para distintos valores de  $(\lambda_1, \lambda_2, t)$



## Capítulo 2

# Ecuaciones Diferenciales Estocásticas

*Indudablemente, una de las herramientas más importantes en el ambiente financiero y en el cálculo estocástico es la integral de Itô. Igualmente importantes son las ecuaciones de Kolmogorov que se deducen a partir de la fórmula de Dynkin. La solución de la ecuación progresiva proporciona la evolución de la densidad en el tiempo, mientras que la ecuación regresiva encuentra aplicaciones en problemas de tiempos de escape, en donde se encuentra la probabilidad de que una partícula deje una región en un tiempo dado.*

### 2.1. El Movimiento Browniano

Como es bien sabido, Robert Brown describió el movimiento de partículas de polen suspendidas en un líquido en 1828. Observó que una partícula se movía de una manera irregular y aleatoria. Albert Einstein en 1905 argumentó que el movimiento de la partícula se debía a un bombardeo de las moléculas del líquido, y obtuvo la ecuación para el movimiento browniano. Pocos años después Langevin describió el movimiento de las partículas mediante una ecuación diferencial estocástica, y es por eso que en física se les conoce como ecuaciones de Langevin. La base matemática del movimiento browniano como proceso estocástico se debió a Norbert Wiener y a Paul Lévy, por lo que a este proceso se le conoce también como proceso de Wiener.

**Definición 2.1.1 (Movimiento Browniano).** El movimiento browniano en  $\mathbb{R}$  es un proceso estocástico  $\{B_t : t \geq 0\}$  con las siguientes propiedades:

- a)  $B_0 = 0$  para casi seguramente.
- b) (Incrementos Normales)  $B_t - B_s$  tiene una distribución normal con media 0 y varianza  $(t - s)$ . Esto implica que con  $s = 0$ ,  $B_t - B_0$  tiene una distribución  $N(0, t)$ .

- c) (Independencia de los incrementos)  $B_{t_1}, B_{t_2} - B_{t_1}, \dots, B_{t_k} - B_{t_{k-1}}$  son independientes para todo  $0 \leq t_1 < t_2 < \dots < t_k$ .
- d) (Trayectorias continuas)  $t \rightarrow B_t(\omega)$ ,  $t \geq 0$  es una trayectoria continua para toda  $\omega \in \Omega$ .

Las propiedades b) y c) determinan todas las distribuciones finito dimensionales, y son gaussianas. La propiedad d) requiere de una construccion que muestre que una version del proceso con trayectorias continuas existe.

El movimiento browniano de dimension  $m$  es el vector  $B(t, \omega) = (B_1(t, \omega), \dots, B_m(t, \omega))^T$  en donde cada una de las coordenadas  $B_i(t, \omega)$  es un movimiento browniano unidimensional mutuamente independiente. Algunas de sus propiedades se obtienen del movimiento browniano unidimensional:  $\mathbb{E}[B_t^T B_t] = mt$ ,  $\mathbb{E}[B_t^T B_s] = m \min(t, s)$ ,  $\mathbb{E}[(B_t - B_s)^T (B_t - B_s)] = m(t - s)$ ,  $\mathbb{E}[(B_t - B_s)(B_t - B_s)^T] = (t - s)I_{m \times m}$ , de donde se concluye que  $(B_t - B_s)$  tiene una distribucion normal  $N(0, (t - s)I_{m \times m})$ . Su densidad de transicion es,

$$p_t(x, y) = \frac{1}{\sqrt{2\pi t}} e^{-|x-y|^2/2t},$$

donde  $x, y$  son vectores  $m$ -dimensionales y  $|x|^2$  es la longitud de  $x$ .

## 2.2. La Integral de Itô

Ahora, para describir el movimiento observado por Robert Brown, si  $b(t, x) \in \mathbb{R}^3$  es la velocidad del fluido en el punto  $x$  al tiempo  $t$ , entonces un modelo matematico razonable para la posicion  $X_t$  de la partucula al tiempo  $t$  ser a una ecuacion diferencial de la forma

$$\frac{dX_t}{dt} = b(t, X_t) + \sigma(t, X_t)\xi_t,$$

donde  $\xi_t \in \mathbb{R}^3$  denota el ruido blanco y  $\sigma(t, X_t) \in \mathbb{R}^{3 \times 3}$ . La interpretacion de Itô<sup>1</sup> de esta ecuacion es

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dB_t,$$

donde  $B_t$  es un movimiento browniano en  $\mathbb{R}^3$ .

Así, en una ecuacion diferencial estocastica de la forma

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dB_t \tag{2.1}$$

donde  $X_t \in \mathbb{R}^n$ ,  $b(t, x) \in \mathbb{R}^n$ ,  $\sigma(t, x) \in \mathbb{R}^{n \times m}$  y  $B_t$  es un movimiento browniano  $m$ -dimensional, llamaremos a  $b$  el *coeficiente de deriva* y a  $\sigma$  o a  $\frac{1}{2}\sigma\sigma^t$

<sup>1</sup>Se prefiere la interpretación de Itô a la de Stratonovich por ser siempre la integral una martingala.

el coeficiente de difusión. La solución de la ecuación (2.1) puede pensarse como la descripción matemática del movimiento de una partícula en un fluido en movimiento; y por tanto tales procesos estocásticos son llamados *difusiones*.

Retomando la forma integral de la ecuación (2.1), se obtiene

$$X_t = X_0 + \int_0^t b(s, X_s) ds + \int_0^t \sigma(s, X_s) dB_s \quad (2.2)$$

por lo que es de interés definir

$$\int_S^T f(s, \omega) dB_s(\omega), \quad \text{para } 0 \leq S \leq T$$

para un espacio de integrandos.

**Definición 2.2.1.** Sea  $\mathcal{H}$  la clase de funciones

$$f(t, \omega) : [0, \infty) \times \Omega \rightarrow \mathbb{R}$$

tales que

- $(t, \omega) \rightarrow f(t, \omega)$  es  $\mathcal{B} \times \mathcal{F}$ -medible, donde  $\mathcal{B}$  es la  $\sigma$ -álgebra de Borel en  $[0, \infty)$ ,  $\mathcal{F}$  una filtración (ver A.1.2).
- $f(t, \omega)$  es  $\mathcal{F}_t$ -adaptado.
- $\mathbb{E} \left[ \int_S^T f(t, \omega)^2 dt \right] < \infty$ .

La parte b) nos dice que la función  $f$  al tiempo  $t$  está determinada por los valores tomados de la historia de las trayectorias de  $\{B_s\}$  hasta el tiempo  $t$ . Por ejemplo,  $f_1(\omega) = B_{t/2}$  es  $\mathcal{F}_t$ -medible, mientras que  $f_2(\omega) = B_{2t}$  no lo es.

Para funciones  $f$  en  $\mathcal{H}$  se define la integral de Itô

$$\int_S^T f(t, \omega) dB_t(\omega) = \lim_{n \rightarrow \infty} \left[ \sum_{j=1}^n f(t_{j-1}, \omega) \Delta B_j(\omega) \right] \quad (2.3)$$

donde  $B_t$  es un movimiento browniano unidimensional, y  $S = t_0 < t_1 < t_2 < \dots < t_{n-1} < t_n = T$  con  $\Delta B_n = B(t_n) - B(t_{n-1})$ . La integral estocástica es una operación natural asociada a las trayectorias brownianas: una trayectoria se corta en incrementos consecutivos gaussianos, cada incremento se multiplica por una variable aleatoria y estos números se suman de nuevo para reconstruir la integral estocástica. Así, podemos pensar en la integral estocástica como una caminata aleatoria con incrementos que tienen diferentes amplitudes, una especie de caminata aleatoria no homogénea. En este contexto es importante enfatizar el papel de las funciones adaptadas. Considere el incremento  $j$ -ésimo después de multiplicarlo por la variable aleatoria  $f(t_{j-1})$ :

$$f(t_{j-1}) \Delta B_j = f(t_{j-1})(B(t_j) - B(t_{j-1})).$$

Una vez que se conoce la historia de la trayectoria hasta el tiempo  $t_{j-1}$ , el valor de  $f(t_{j-1})$  se conoce tambien. Por tanto, el incremento de la integral estocastica en el siguiente periodo condicionado al pasado hasta el tiempo  $t$  es gaussiano con media cero y varianza  $f(t_{j-1})^2(t_j - t_{j-1})$ . As, las funciones adaptadas son funciones adecuadas para definir una caminata aleatoria continua y no homogenea para la integral estocastica.

La forma de escoger a  $t$  en la particion es la que define que la integral sea de Itô. Si  $t$  se toma como el punto medio de la particion, la integral que se obtiene es la de *Stratonovich*.

Algunas propiedades de la integral de Itô son:

**Teorema 2.2.1.** Sean  $f$  y  $g$  en  $\mathcal{H}$ ,  $0 \leq S < U \leq T$ . Entonces

a) (la isometría de Itô)

$$\mathbb{E} \left[ \left( \int_S^T f(t, \omega) dB_t(\omega) \right)^2 \right] = \mathbb{E} \left[ \int_S^T f^2(t, \omega) dt \right] \quad \text{para todo } f \in \mathcal{H}$$

b)  $\int_S^T f dB_t = \int_S^U f dB_t + \int_U^T f dB_t$  para casi toda  $\omega$

c) (linealidad)  $\int_S^T (cf + g) dB_t = c \int_S^T f dB_t + \int_S^T g dB_t$  ( $c$  constante) para casi toda  $\omega$

d) (media cero)  $\mathbb{E} \left[ \int_S^T f dB_t \right] = 0$

e)  $\int_S^T f dB_t$  es  $\mathcal{F}_T$ -medible

f)  $M_t(\omega) = \int_0^t f(s, \omega) dB_s$  es una martingala con respecto a  $\mathcal{F}_t$ . Ver A.1.3.

**Definicion 2.2.2 (El proceso de Itô unidimensional).** Llamaremos un proceso de Itô a un proceso estocastico  $X_t$  de la forma

$$dX_t = u(t, \omega)dt + v(t, \omega)dB_t.$$

**Teorema 2.2.2 (La formula de Itô unidimensional).** Sea  $X_t$  un proceso de Itô y  $g(t, x) \in C^2([0, \infty) \times \mathbb{R})$  (i.e.,  $g$  es una función dos veces diferenciable en  $[0, \infty) \times \mathbb{R}$ ). Entonces

$$Y_t = g(t, X_t)$$

es también un proceso de Itô, y

$$dY_t = \frac{\partial g(t, X_t)}{\partial t} dt + \frac{\partial g(t, X_t)}{\partial x} dX_t + \frac{1}{2} \frac{\partial^2 g(t, X_t)}{\partial x^2} (dX_t)^2,$$

donde  $(dX_t)^2 = (dX_t) \cdot (dX_t)$  se calcula de acuerdo a la siguiente regla

$$dt \cdot dt = dt \cdot dB_t = dB_t \cdot dt = 0, \quad dB_t \cdot dB_t = dt.$$



**Ejemplo 2.2.1.** Un modelo estocástico usado con frecuencia en finanzas es el llamado modelo lognormal. En este caso se supone que el precio del proceso evoluciona según la ecuación estocástica:

$$dS_t = \mu S_t dt + \sigma S_t dB_t,$$

donde  $\mu$  y  $\sigma$  son constantes reales. Entonces, aplicando la fórmula de Itô a  $g(t, x) = \ln x$ , para  $x > 0$

$$d(\ln S_t) = 0 + \frac{1}{S_t} \cdot dS_t + \frac{1}{2} \left( -\frac{1}{S_t^2} \right) (dS_t)^2.$$

Así,

$$\ln S_t - \ln S_0 = \int_0^t \frac{\mu S_r dr + \sigma S_r dB_r}{S_r} - \frac{1}{2} \int_0^t \frac{S_r^2 \sigma^2}{S_r^2} dr,$$

es decir,

$$\ln \frac{S_t}{S_0} = \int_0^t \left( \mu - \frac{1}{2} \sigma^2 \right) dr + \int_0^t \sigma dB_r = \left( \mu - \frac{1}{2} \sigma^2 \right) t + \sigma B_t.$$

Por tanto,

$$S_t = S_0 e^{(\mu - \frac{1}{2} \sigma^2)t + \sigma B_t}. \quad (2.4)$$

Calculemos ahora  $\mathbb{E}[S_t]$ , sea  $Y_t = e^{\sigma B_t}$  entonces

$$dY_t = 0 \cdot dt + \sigma e^{\sigma B_t} dB_t + \frac{1}{2} \sigma^2 e^{\sigma B_t} dt$$

ó

$$Y_t - Y_0 = \sigma \int_0^t e^{\sigma B_s} dB_s + \frac{1}{2} \sigma^2 \int_0^t Y_s ds.$$

Ahora, por el teorema 2.2.1 d) y si  $\mathbb{E}[Y_0] = y_0$  tenemos que

$$\mathbb{E}[Y_t] = y_0 + \frac{1}{2} \sigma^2 \int_0^t \mathbb{E}[Y_s] ds,$$

de donde se sigue que

$$\frac{d}{dt} \mathbb{E}[Y_t] = \frac{1}{2} \sigma^2 \mathbb{E}[Y_t]$$

así,

$$\mathbb{E}[Y_t] = e^{\frac{1}{2} \sigma^2 t}.$$

Finalmente

$$\mathbb{E}[S_t] = S_0 e^{(\mu - \frac{1}{2} \sigma^2)t + \frac{1}{2} \sigma^2 t} = S_0 e^{\mu t}.$$

**Ejemplo 2.2.2.** Un ejemplo de gran utilidad que incorpora el concepto de correlación finita temporal es la ecuación de Ornstein-Uhlenbeck, cuya solución se conoce como el proceso de Ornstein-Uhlenbeck.

$$dX_t = \mu X_t dt + \sigma dB_t.$$

Multiplicando por  $e^{-\mu t}$ , se obtiene

$$e^{-\mu t} dX_t = e^{-\mu t} \mu X_t dt + e^{-\mu t} \sigma dB_t,$$

entonces

$$d(e^{-\mu t} X_t) = \sigma e^{-\mu t} dB_t,$$

así

$$X_t = e^{\mu t} \left( X_0 + \sigma \int_0^t e^{-\mu s} dB_s \right).$$

Calculemos ahora su primer y segundo momentos. Por el teorema 2.2.1 d),

$$\mathbb{E}[X_t] = X_0 e^{\mu t}.$$

Ahora,

$$\begin{aligned} \text{Var}[X_t] &= \mathbb{E}[X_t - \mathbb{E}(X_t)]^2 \\ &= \mathbb{E} \left[ \sigma e^{\mu t} \int_0^t e^{-\mu s} dB_s \right]^2, \quad \text{y por la isometría de Itô (teorema 2.2.1)} \\ &= \sigma^2 e^{2\mu t} \int_0^t e^{-2\mu s} ds \\ &= \frac{\sigma^2}{2\mu} (e^{2\mu t} - 1). \end{aligned}$$

Multiplicamos  $X_t$  con  $X_{t+\tau}$  ( $\tau > 0$ ), tomamos esperanza y junto con el teorema 2.2.1 llegamos a la función de correlación

$$\begin{aligned} \mathbb{E}[X_t X_{t+\tau}] &= X_0^2 e^{2\mu t + \mu\tau} + e^{\mu\tau} \sigma^2 e^{2\mu t} \int_0^t e^{-2\mu s} ds \\ &= e^{\mu\tau} \left[ X_0^2 e^{2\mu t} + \frac{\sigma^2}{2\mu} (e^{2\mu t} - 1) \right]. \end{aligned}$$

Cuando  $t \rightarrow \infty$  y  $\mu < 0$  obtenemos el proceso de Ornstein-Uhlenbeck estacionario,

$$\begin{aligned} \mathbb{E}[X_t] &= 0, \\ \mathbb{E}[X_t X_{t+\tau}] &= -\frac{\sigma^2}{2\mu} e^{\mu\tau}. \end{aligned}$$

De manera similar, al considerar ahora  $\tau < 0$  y  $\mu < 0$ , obtenemos

$$\mathbb{E}[X_t X_{t+\tau}] = -\frac{\sigma^2}{2\mu} e^{-\mu\tau}.$$

Así,

$$\mathbb{E}[X_t X_{t+\tau}] = -\frac{\sigma^2}{2\mu} e^{\mu|\tau|}.$$

El proceso de Ornstein-Uhlenbeck<sup>2</sup> tiene una correlación temporal diferente de cero, lo que significa que los valores del ruido para diferentes tiempos no son variables aleatorias independientes.

El comportamiento cualitativo para  $\mu > 0$  es completamente diferente para el caso  $\mu < 0$ . Cuando  $\mu < 0$ ,  $X_t$  converge en distribución a  $N(0, -\sigma^2/2\mu)$ , y por supuesto  $X_t$  no converge casi seguramente.

<sup>2</sup>El proceso de Ornstein-Uhlenbeck es el único proceso que es simultáneamente Gaussiano, Markoviano y estacionario.

**Definición 2.2.3 (La formula de Itô multidimensional).** Sea  $B(t, \omega) = (B_1(t, \omega), \dots, B_m(t, \omega))$  un movimiento browniano  $m$ -dimensional. Si cada uno de los procesos  $u_i(t, \omega)$  y  $v_{ij}(t, \omega)$  satisfacen las condiciones de la definición 2.2.2 para  $(1 \leq i \leq n, 1 \leq j \leq m)$  entonces podemos formar el siguiente arreglo de Itô

$$\begin{aligned} dX_1 &= u_1 dt + v_{11} dB_1 + \dots + v_{1m} dB_m \\ &\vdots \\ dX_n &= u_n dt + v_{n1} dB_1 + \dots + v_{nm} dB_m \end{aligned}$$

o en notación matricial simplemente

$$dX(t) = u dt + v dB(t),$$

donde

$$X(t) = \begin{pmatrix} X_1(t) \\ \vdots \\ X_n(t) \end{pmatrix}, u = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}, v = \begin{pmatrix} u_{11} & \dots & v_{1m} \\ \vdots & & \vdots \\ u_{n1} & \dots & v_{nm} \end{pmatrix}, dB(t) = \begin{pmatrix} dB_1(t) \\ \vdots \\ dB_m(t) \end{pmatrix}.$$

Llamaremos al proceso  $X_t$  un proceso de Itô  $n$ -dimensional.

Sea  $g(t, x) = (g_1(t, x), \dots, g_p(t, x))$  una función  $C^2$  de  $[0, \infty) \times \mathbb{R}^n$  a  $\mathbb{R}^p$ . Entonces el proceso

$$Y(t, \omega) = g(t, X(t))$$

es también un proceso de Itô, cuya componente  $Y_k$  está dada por

$$dY_k = \frac{\partial g_k(t, X)}{\partial t} dt + \sum_i \frac{\partial g_k(t, X)}{\partial x_i} dX_i + \frac{1}{2} \sum_{ij} \frac{\partial^2 g_k(t, X)}{\partial x_i \partial x_j} dX_i dX_j$$

donde  $dB_i dB_j = \delta_{ij} dt$ ,  $dB_i dt = dt dB_i = dt dt = 0$ .

Como la teoría de las ecuaciones diferenciales ordinarias, la teoría de las ecuaciones diferenciales estocásticas tiene un teorema estándar de existencia y unicidad.

**Teorema 2.2.3 (Existencia y unicidad para las ecuaciones diferenciales estocásticas).** Sea  $T > 0$  y  $b(\cdot, \cdot) : [0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ ,  $\sigma(\cdot, \cdot) : [0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}^{n \times m}$  funciones medibles que satisfacen

a) (Condición de Crecimiento)

$$|b(t, x)| + |\sigma(t, x)| \leq C(1 + |x|); \quad x \in \mathbb{R}^n, t \in [0, T]$$

para alguna constante  $C$ , (donde  $|\sigma|^2 = \sum |\sigma_{ij}|^2 = \text{tr}(\sigma \sigma^\top)$ ) y tal que

b) (Condición de Lipschitz)

$$|b(t, x) - b(t, y)| + |\sigma(t, x) - \sigma(t, y)| \leq D|x - y|; \quad x, y \in \mathbb{R}^n, t \in [0, T]$$

para alguna constante  $D$ .

Sea  $Z$  una variable aleatoria que es independiente de la  $\sigma$ -álgebra  $\mathcal{F}_\infty^{(m)}$  generada por  $B_s(\cdot)$ ,  $s \geq 0$  y tal que

$$\mathbb{E}[|Z|^2] < \infty.$$

Entonces la ecuación diferencial estocástica (2.1) con  $0 \leq t \leq T$ ,  $X_0 = Z$  tiene una solución única y continua con la propiedad de que  $X_t(\omega)$  es un proceso adaptado a la filtración  $\mathcal{F}_t^Z$  generada por  $Z$  y  $B_s$ ;  $s \leq t$  y

$$\mathbb{E} \left[ \int_0^T |X_t|^2 dt \right] < \infty.$$

Si  $(\Omega, \mathcal{F}, \mathbb{P})$  y  $\{B_t\}_{t \geq 0}$  son dados, entonces la solución de la ecuación (2.1), si existe, es llamada una *solución fuerte*; en cambio, si originalmente tenemos los coeficientes  $b(x, t)$  y  $\sigma(x, t)$  podemos preguntarnos si existen un par de procesos  $(\tilde{X}_t, \tilde{B}_t)$ , y  $\tilde{B}_t$  un movimiento browniano respecto a la filtración  $\mathcal{H}_t$  de nida sobre el espacio de probabilidad  $(\Omega, \mathcal{H}, \mathbb{P})$  tal que la ecuación (2.1) se satisface, entonces a  $(\tilde{X}_t, \tilde{B}_t)$  se le llama *solución débil*. El concepto de solución débil permite dar solución a la ecuación diferencial estocástica cuando no existe una solución fuerte. Las soluciones débiles son soluciones en distribución y pueden ser de nidas en algún otro espacio de probabilidad.

### 2.3. El Generador Infinitesimal

Es fundamental para muchas aplicaciones que se pueda asociar un operador diferencial  $A$  de segundo orden a un proceso de difusión  $X_t$ . La conexión básica entre  $A$  y  $X_t$  es que  $A$  genera al proceso  $X_t$ .

**Definición 2.3.1 (Difusión de Itô).** Un proceso estocástico  $X_t(\omega) = X(t, \omega) : [0, \infty) \times \Omega \rightarrow \mathbb{R}^n$  homogéneo en el tiempo que satisface una ecuación diferencial estocástica de la forma

$$dX_t = b(X_t)dt + \sigma(X_t)dB_t, \quad t \geq s; X_s = x \quad (2.5)$$

donde  $B_t$  es un movimiento browniano  $m$ -dimensional y  $b : \mathbb{R}^n \rightarrow \mathbb{R}^n$ ,  $\sigma : \mathbb{R}^n \rightarrow \mathbb{R}^{n \times m}$  satisfacen la condición de Lipschitz y de crecimiento se llama *difusión de Itô*.

**Definición 2.3.2 (Generador infinitesimal).** Sea  $X_t$  una difusión de Itô homogénea en el tiempo en  $\mathbb{R}^n$ . El generador infinitesimal  $A$  de  $X_t$  se define como

$$Af(x) = \lim_{t \downarrow 0} \frac{\mathbb{E}^x [f(X_t)] - f(x)}{t}; \quad x \in \mathbb{R}^n.$$

El conjunto de funciones  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  tales que el límite existe en  $x$  se denota por  $\mathcal{D}_A(x)$ , mientras que  $\mathcal{D}_A$  denota al conjunto de funciones para las que el límite existe para toda  $x \in \mathbb{R}^n$ .

La relación entre  $A$  y los coeficientes  $b, \sigma$  en la ecuación diferencial estocástica (2.5) la proporciona el siguiente teorema.

**Teorema 2.3.1.** Sea  $X_t$  la difusión de Itô

$$dX_t = b(X_t)dt + \sigma(X_t)dB_t.$$

Si  $f \in C_0^2(\mathbb{R}^n)$  entonces  $f \in \mathcal{D}_A$  y

$$Af(x) = \sum_i b_i(x) \frac{\partial f}{\partial x_i} + \frac{1}{2} \sum_{i,j} (\sigma \sigma^t)_{i,j}(x) \frac{\partial^2 f}{\partial x_i \partial x_j}.$$

**Ejemplo 2.3.1.** El movimiento browniano  $n$ -dimensional es por supuesto la solución de la ecuación diferencial estocástica

$$dX_t = dB_t,$$

es decir, tenemos  $b = 0$  y  $\sigma = I_n$ , la matriz de identidad  $n$ -dimensional. Por tanto el generador de  $B_t$  es

$$Af = \frac{1}{2} \sum \frac{\partial^2 f}{\partial x_i^2}; \quad f = f(x_1, \dots, x_n) \in C_0^2(\mathbb{R}^n),$$

es decir,  $A = \frac{1}{2}\Delta$ , donde  $\Delta$  es el laplaciano. Mientras que para el ejemplo 2.2.1  $b(x) = \mu x$ ,  $\sigma(x) = \sigma x$ , y su generador infinitesimal es

$$Af(x) = \mu x \frac{df}{dx} + \frac{1}{2} \sigma^2 x^2 \frac{d^2 f}{dx^2}.$$

**Teorema 2.3.2 (Formula de Dynkin).** Sea  $f \in C_0^2(\mathbb{R}^n)$ . Suponga que es un tiempo de paro,  $\mathbb{E}^x[\tau] < \infty$ . Entonces

$$\mathbb{E}^x [f(X_\tau)] = f(x) + \mathbb{E}^x \left[ \int_0^\tau Af(X_s) ds \right]. \quad (2.6)$$

**Teorema 2.3.3 (Ecuación regresiva de Kolmogorov).** Sea  $f \in C_0^2(\mathbb{R}^n)$ .

a) Defina

$$u(t, x) = \mathbb{E}^x [f(X_t)]. \quad (2.7)$$

Entonces,  $u(t, \cdot) \in \mathcal{D}_A$  para cada  $t$  y

$$\frac{\partial u}{\partial t} = Au, \quad t > 0, x \in \mathbb{R}^n \quad (2.8)$$

$$u(0, x) = f(x); \quad x \in \mathbb{R}^n \quad (2.9)$$

b) Más aún, si  $w(t, x) \in C^{1,2}(\mathbb{R} \times \mathbb{R}^n)$  es una función acotada que satisface (2.8) y (2.9) entonces  $w(t, x) = u(t, x)$  dada por (2.7).

Con un poco más de trabajo podemos obtener la siguiente generalización de la ecuación regresiva de Kolmogorov:

**Teorema 2.3.4 (La fórmula de Feynman-Kač).** Sea  $f \in C_0^2(\mathbb{R}^n)$  y  $q \in C(\mathbb{R}^n)$ . Suponga que  $q$  está acotada inferiormente.

a) Sea

$$v(t, x) = \mathbb{E}^x \left[ e^{-\int_0^t q(X_s) ds} f(X_t) \right]. \quad (2.10)$$

Entonces

$$\frac{\partial v}{\partial t} = Av - qv; \quad t > 0, x \in \mathbb{R}^n \quad (2.11)$$

$$v(0, x) = f(x); \quad x \in \mathbb{R}^n \quad (2.12)$$

b) Más aún, si  $w(t, x) \in C^{1,2}(\mathbb{R} \times \mathbb{R}^n)$  es una función acotada en  $K \times \mathbb{R}^n$  para cada compacto  $K \subset \mathbb{R}$  y  $w$  satisface (2.11), (2.12), entonces  $w(t, x) = v(t, x)$ , dada por (2.10).

**Ejemplo 2.3.2 (Ecuación regresiva).** Es interesante observar qué pasa en el teorema 2.3.3 cuando tomamos  $f(\cdot) = \delta(\cdot - y)$ :

$$\mathbb{E}^x [f(X_t)] = \mathbb{E} [\delta(X_t - y) | X_0 = x] = \mathbb{P} [X_t = y | X_0 = x] = p_t(x, y), \quad (2.13)$$

entonces

$$\frac{\partial p_t(x, y)}{\partial t} = A_x p_t(x, y),^3 \quad (2.14)$$

es decir, la densidad de  $X_t$  condicionada a que  $X_0 = x$ , satisface la ecuación regresiva de Kolmogorov.

**Ejemplo 2.3.3 (Ecuación de Fokker-Planck).** Ahora, sea  $X_t$  una difusión de Itô en  $\mathbb{R}^n$  con generador

$$Af(y) = \frac{1}{2} \sum_{i,j} a_{ij}(y) \frac{\partial^2 f}{\partial y_i \partial y_j} + \sum_i b_i(y) \frac{\partial f}{\partial y_i}; \quad f \in C_0^2$$

y suponga que la medida de transición de  $X_t$  tiene una densidad  $p_t(x, y)$ , i.e.,

$$\mathbb{E}^x [f(X_t)] = \int_{\mathbb{R}^n} f(y) p_t(x, y) dy; \quad f \in C_0^2.$$

Suponga que  $y \rightarrow p_t(x, y)$  es suave para cada  $t, x$ .

Usando la fórmula de Dynkin, ecuación (2.6), se tiene que

$$\int_{\mathbb{R}^n} f(y) p_t(x, y) dy = \mathbb{E}^x [f(X_t)] = f(x) + \int_0^t \int_{\mathbb{R}^n} A_y f(y) p_s(x, y) dy ds, \quad f \in C_0^2;$$

---

<sup>3</sup>Dado que el proceso  $X_t$  es homogéneo en el tiempo, i.e.,  $p_t^{x,s}(x, y) = p_{t-s}^{x,0}(x, y) = p_\tau(x, y)$ , entonces

$$\frac{\partial p_t^{x,s}(x, y)}{\partial t} = \frac{\partial p_\tau(x, y)}{\partial \tau},$$

y

$$-\frac{\partial p_t^{x,s}(x, y)}{\partial s} = \frac{\partial p_\tau(x, y)}{\partial \tau}.$$

y diferenciando respecto de  $t$ ,

$$\int_{\mathbb{R}^n} f(y) \frac{\partial}{\partial t} p_t(x, y) dy = 0 + \int_{\mathbb{R}^n} A_y f(y) p_t(x, y) dy.$$

$A_y^*$  es el adjunto de  $A$  (ver ejemplo C.1.1) y opera sobre la variable  $y$  y está dado por

$$A_y^* \phi(y) = \frac{1}{2} \sum_{i,j} \frac{\partial^2}{\partial y_i \partial y_j} (a_{ij} \phi) - \sum_i \frac{\partial}{\partial y_i} (b_i \phi); \quad \phi \in C^2, \quad (2.15)$$

entonces

$$\langle A\phi, \psi \rangle = \langle \phi, A^* \psi \rangle \quad \text{para } \phi \in C_0^2, \psi \in C^2,$$

donde  $\langle \cdot, \cdot \rangle$  es el producto interior en  $L^2(dy)$ ; en particular

$$\langle A_y f, p_t \rangle = \langle f, A_y^* p_t \rangle,$$

es decir

$$\int_{\mathbb{R}^n} A_y f(y) p_t(x, y) dy = \int_{\mathbb{R}^n} f(y) A_y^* p_t(x, y) dy,$$

entonces

$$\int_{\mathbb{R}^n} f(y) \frac{\partial}{\partial t} p_t(x, y) dy = \int_{\mathbb{R}^n} f(y) A_y^* p_t(x, y) dy,$$

así,

$$\frac{\partial}{\partial t} p_t(x, y) = A_y^* p_t(x, y) \quad \text{para todo } x, y. \quad (2.16)$$

En muchas aplicaciones, como física, ingeniería y finanzas, la importancia de las difusiones está en su conexión con las ecuaciones diferenciales parciales, y en ocasiones las difusiones se especifican por una ecuación diferencial llamada ecuación de Fokker-Planck, ecuación (2.16), también conocida como ecuación progresiva de Kolmogorov. Aunque estas ecuaciones raramente tienen solución analítica, pueden ser resueltas fácilmente de manera numérica. En la práctica es suficiente verificar que las condiciones de existencia y unicidad se satisfacen y las soluciones se pueden calcular con el grado deseado de precisión.

En seguida se muestra una conexión sorprendente entre las ecuaciones diferenciales estocásticas y la solución de ecuaciones diferenciales parciales de segundo orden. Un ejemplo es la solución de la ecuación de Black-Scholes en la valuación de opciones, la cual es más fácil de resolver por métodos estocásticos que una ecuación diferencial de segundo orden. Al principio parecer a extraño pues un problema es completamente determinístico y el otro entraña ruido blanco<sup>4</sup> en la ecuación diferencial.

Sea  $dX_t = b(X_t)dt + \sigma(X_t)dB_t$  una difusión de Itô en  $\mathbb{R}^n$  con generador  $A$ , y si  $f \in C_0^2(\mathbb{R}^n)$  pongamos  $Z = f(X)$  y para simplificar la notación suprimimos el índice  $t$  y  $X_1, \dots, X_n$  y  $B_1, \dots, B_n$  denotaran las coordenadas de  $X$  y  $B$  respectivamente, entonces aplicando la fórmula de Itô multidimensional (definición 2.2.3)

$$dZ = \sum_i \frac{\partial f(X)}{\partial x_i} dX_i + \frac{1}{2} \sum_{ij} \frac{\partial^2 f(X)}{\partial x_i \partial x_j} dX_i dX_j$$

<sup>4</sup>Se usa ruido blanco como un símil a la luz blanca que está constituida por todas las frecuencias (colores). El ruido blanco se entenderá como la derivada del movimiento browniano.

y como

$$dX_i = b_i dt + \sum_k \sigma_{ik} dB_k = b_i dt + (\sigma dB)_i$$

$$dZ = \sum_i b_i \frac{\partial f}{\partial x_i} dt + \frac{1}{2} \sum_{ij} \frac{\partial^2 f}{\partial x_i \partial x_j} (\sigma dB)_i (\sigma dB)_j + \sum_i \frac{\partial f}{\partial x_i} (\sigma dB)_i$$

y observando que

$$\begin{aligned} (\sigma dB)_i \cdot (\sigma dB)_j &= \left( \sum_k \sigma_{ik} dB_k \right) \left( \sum_n \sigma_{jn} dB_n \right) \\ &= \left( \sum_k \sigma_{ik} \sigma_{jk} dt \right) = (\sigma \sigma^\top)_{ij} dt, \end{aligned}$$

se obtiene

$$\begin{aligned} f(X_t) &= f(X_0) + \int_0^t \left( \sum_i b_i \frac{\partial f}{\partial x_i} + \frac{1}{2} \sum_{ij} (\sigma \sigma^\top)_{ij} \frac{\partial^2 f}{\partial x_i \partial x_j} \right) ds \\ &\quad + \sum_{i,k} \int_0^t \sigma_{ik} \frac{\partial f}{\partial x_i} dB_k. \end{aligned}$$

Aquí es donde se observa la utilidad del generador infinitesimal;

$$f(X_t) = f(x) + \int_0^t Af(X_s) ds + \int_0^t \nabla f^\top(X_s) \sigma(X_s) dB_s.$$

De ninos,

$$M_t = f(X_t) - \int_0^t Af(X_r) dr \quad \left( = f(x) + \int_0^t \nabla f^\top(X_r) \sigma(X_r) dB_r \right).$$

Ahora, como las integrales de Itô son martingalas (Teorema 2.2.1  $f$ ) con respecto a las  $\sigma$ -álgebras  $\{\mathcal{F}_t^{(m)}\}$ , tenemos que para  $t > s$

$$\mathbb{E}^x[M_t | \mathcal{F}_s^{(m)}] = M_s.$$

Luego entonces,

$$\mathbb{E}^x[M_t | \mathcal{M}_s] = \mathbb{E}^x \left[ \mathbb{E}^x[M_t | \mathcal{F}_s^{(m)}] | \mathcal{M}_s \right] = \mathbb{E}^x[M_s | \mathcal{M}_s] = M_s,$$

pues  $M_t$  es  $\mathcal{M}_t$ -medible. Así hemos demostrado:

**Teorema 2.3.5.** Si  $X_t$  es una difusión de Itô en  $\mathbb{R}^n$  con generador  $A$ , entonces para todo  $f \in C_0^2(\mathbb{R}^n)$  el proceso

$$M_t = f(X_t) - \int_0^t Af(X_r) dr$$

es una martingala respecto de  $\{\mathcal{M}_t\}$ .



**Ejemplo 2.3.4.** En conexión con la deducción de la fórmula de Black-Scholes para la valuación de opciones financieras aparece la siguiente ecuación diferencial parcial (ver [1]);

$$\frac{\partial c}{\partial t} = \mu x \frac{\partial c}{\partial x} + \frac{1}{2} \sigma^2 x^2 \frac{\partial^2 c}{\partial x^2} - rc; \quad t > 0, x \in \mathbb{R},$$

$$c(0, x) = (x - K)^+ \doteq \max(x - K, 0);$$

donde  $r > 0, K > 0, \sigma, \mu$ , son constantes. Es decir, la ecuación diferencial parcial la puedo escribir en términos del generador infinitesimal  $A$ , (ver ejemplo 2.3.1) con  $b(x) = \mu x, \sigma(x) = \sigma x$ , y con la siguiente ecuación de difusión de Itô  $dX_t = \mu X_t dt + \sigma X_t dB_t$ , cuya solución es

$$X_t = X_0 e^{(\mu - \frac{1}{2}\sigma^2)t + \sigma B_t},$$

por el ejemplo 2.2.1.

Así,

$$\frac{\partial c}{\partial t} = Ac - rc$$

con  $q(x) = r$  en el teorema 2.3.4 de la fórmula de Feynman Kač. Entonces la solución a  $c$  es;

$$c(t, x) = \mathbb{E}^x \left[ e^{-\int_0^t q(X_s ds)} f(X_t) \right] = e^{-rt} \mathbb{E}^x [f(X_t)] = e^{-rt} \int_{\mathbb{R}} f(x) f_{X_t}(x) dx.$$

Ahora,

$$\begin{aligned} F_{X_t}(x) &= \mathbb{P}(X_t \leq x) = \mathbb{P}(X_0 e^{(\mu - \frac{1}{2}\sigma^2)t + \sigma B_t} \leq x) \\ &= \mathbb{P}\left(B_t \leq \frac{\ln(x/X_0) - (\mu - \frac{1}{2}\sigma^2)t}{\sigma}\right) \\ &= \int_{-\infty}^{\frac{\ln(x/X_0) - (\mu - \frac{1}{2}\sigma^2)t}{\sigma}} \frac{1}{\sqrt{2\pi t}} e^{-y^2/2t} dy, \end{aligned}$$

luego entonces,

$$f_{X_t}(x) = \frac{1}{\sqrt{2\pi t}} e^{-\left(\frac{\ln(x/X_0) - (\mu - \frac{1}{2}\sigma^2)t}{\sigma}\right)^2/2t} \frac{1}{\sigma} \frac{X_0}{x} \frac{1}{X_0}.$$

Reconocemos a  $f_{X_t}(x)$  como la densidad de una lognormal de parámetros  $[(\mu - \frac{1}{2}\sigma^2)t + \log X_0]$  y  $\sigma^2 t$ . Así,

$$\begin{aligned} c(t, x) &= e^{-rt} \int_{\mathbb{R}} c(0, x) f_{X_t}(x) dx \quad (2.17) \\ &= \int_{\mathbb{R}} (x_0 e^{(\mu - \frac{1}{2}\sigma^2)t + \sigma y} - K)^+ \frac{1}{\sqrt{2\pi t} \sigma x} e^{-\left(\frac{\ln(x/X_0) - (\mu - \frac{1}{2}\sigma^2)t}{\sigma}\right)^2/2t} dx. \end{aligned}$$

Pero  $(x_0 e^{(\mu - \frac{1}{2}\sigma^2)t + \sigma y} - K)^+ = x_0 e^{(\mu - \frac{1}{2}\sigma^2)t + \sigma y} - K$  si  $x_0 e^{(\mu - \frac{1}{2}\sigma^2)t + \sigma y} > K$ , entonces  $y > \frac{\ln(K/x_0) - (\mu - \frac{1}{2}\sigma^2)t}{\sigma} = w$ ;  $y$  con el cambio de variable  $y = \frac{\ln(x/X_0) - (\mu - \frac{1}{2}\sigma^2)t}{\sigma}$ ,  $dx = \sigma x dy$ , se tiene que

$$\begin{aligned} c(t, x) &= e^{-rt} \int_{y=w}^{\infty} (x_0 e^{(\mu - \frac{1}{2}\sigma^2)t + \sigma y} - K) \frac{1}{\sqrt{2\pi t}} e^{-y^2/2t} dy \\ &= I_1 - I_2. \end{aligned}$$

Veamos  $I_1$ ,

$$\begin{aligned} I_1 &= e^{-rt} \int_{y=w}^{\infty} x_0 e^{(\mu - \frac{1}{2}\sigma^2)t + \sigma y} \frac{1}{\sqrt{2\pi t}} e^{-y^2/2t} dy \\ &= x_0 e^{-rt + \mu t} \int_{y=w}^{\infty} \frac{1}{\sqrt{2\pi t}} e^{-\frac{1}{2}\sigma^2 t + \sigma y - y^2/2t} dy, \end{aligned}$$

con el cambio de variable  $u = \frac{y}{\sqrt{t}} - \sigma\sqrt{t}$ ,

$$\begin{aligned} I_1 &= x_0 e^{-rt + \mu t} \int_{u=\frac{w}{\sqrt{t}} - \sigma\sqrt{t}}^{\infty} \frac{1}{\sqrt{2\pi t}} e^{-\frac{u^2}{2}} \sqrt{t} du \\ &= x_0 e^{-rt + \mu t} \Phi\left(\sigma\sqrt{t} - \frac{w}{\sqrt{t}}\right) \\ &= x_0 e^{-rt + \mu t} \Phi(d_1), \end{aligned}$$

con  $d_1 = \frac{\ln(x_0/K) + (\mu + \frac{1}{2}\sigma^2)t}{\sigma\sqrt{t}}$ .

Veamos  $I_2$ ,

$$I_2 = \frac{e^{-rt}}{\sqrt{2\pi t}} \int_{y=w}^{\infty} K e^{-y^2/2t} dy,$$

con el cambio de variable  $u = \frac{y}{\sqrt{t}}$ ,

$$\begin{aligned} I_2 &= \frac{K e^{-rt}}{\sqrt{2\pi}} \int_{u=\frac{w}{\sqrt{t}}}^{\infty} e^{-u^2/2} du \\ &= e^{-rt} K \Phi\left(-\frac{w}{\sqrt{t}}\right) \\ &= e^{-rt} K \Phi(d_2) \end{aligned}$$

con  $d_2 = \frac{\ln(x_0/K) + (\mu - \frac{1}{2}\sigma^2)t}{\sigma\sqrt{t}}$ . Entonces,

$$c(t, x) = x_0 \Phi(d_1) - e^{-rt} K \Phi(d_2),$$

pues en la ecuación diferencial de Black-Scholes  $r = \mu$ .

## Capítulo 3

# Aplicación

*Son varios los objetivos que nos hemos planteado en este capítulo. Aproximar el precio de un derivado en  $n$  estados cambiantes regido por un proceso de nacimiento y muerte. Construir un pozo biestable en el que el tiempo de escape de uno de los pozos tenga una distribución exponencial estableciendo de esta forma la matriz de intensidad de proceso de nacimiento y muerte. Proponer el movimiento Ornstein-Uhlenbeck en contraste al movimiento Browniano en la valuación del derivado.*

### 3.1. Opciones

Un producto financiero derivado es aquel cuyo precio se encuentra en función de otro, al que se le llama el activo subyacente. Con esta definición la cantidad de derivados es muy extensa, sin embargo son cuatro los más utilizados: los *forwards*, los futuros, los *swaps* y las opciones. Una característica de todos ellos, es que en vez de ser instrumentos financieros o más precisamente títulos de crédito, son contratos, lo que implica, desde un punto de vista legal, un manejo diferente.

Consideremos los siguientes supuestos importantes en la valuación de algunos instrumentos financieros. Todos los inversionistas tienen acceso a la misma información relevante y pueden comprar o vender a cualquier precio; sin costos de transacción o comisiones, y el valor de los bienes es perfectamente divisible y líquido. No hay restricción alguna en el monto de algún crédito bancario, y la tasa de interés bancaria es la misma para ambas partes.

Las opciones son ejemplos de derivados que cotizan en bolsa, es decir, instrumentos cuyo valor depende de los precios de otros instrumentos más básicos, tales como las acciones o los bonos. Básicamente una *opción call* (una *opción put*, respectivamente) da el derecho de comprar (de vender, respectivamente) el activo subyacente de la opción en una fecha futura a un precio previamente establecido. Las opciones se pueden pensar como un seguro tradicional, es decir, en el que de no utilizarse (dado que no hay siniestro), se pierde la prima paga-

da. Las opciones pueden ser o no estandarizadas, y pueden operarse o no en las bolsas establecidas. Desde luego en caso de que esten en bolsa su volumen de operacion aumenta considerablemente. Las opciones se han comercializado por siglos, sin embargo a habido una expansion del mercado de opciones a partir del surgimiento de las bolsas de opciones en 1973. Las bolsas mas importantes para opciones estandarizadas son la CBOE (*Chicago Board of Trade*) y la CME (*Chicago Mercantile Exchange*), ambas en EUA. En el caso de Mexico, a la fecha, las opciones aun no operan de forma generalizada, aunque algunos intermediarios financieros ofrecen hacer contratos de opciones a la medida, en particular de dolares. No obstante a partir de 1992, las casas de bolsa, principalmente, emitieron un instrumento al que se le llamo Títulos Opcionales, al que tambien se le conocio con el nombre de *warrant*. A pesar de ser un título de credito y no un contrato, este instrumento esta diseñado de manera analoga a una opcion tradicional, y es el primer intento serio de la comunidad financiera para que en nuestro país se ofrezca este derivado al publico en general. Desde luego se espera que el MexDer (Mercado Mexicano de Derivados) permita operar, en breve plazo, este instrumento de manera similar al que se hace en otros países. A la fecha las opciones han cobrado una gran importancia, sin embargo el máximo seguramente se da en 1997, cuando La Real Academia Sueca de Ciencias concede el Premio Nobel de Economía a Robert C. Merton y a Myron S. Scholes. Black, Merton y Scholes sentaron las bases para el rapido crecimiento del mercado de derivados en los últimos 10 años. Sin embargo, su metodo tiene una utilizacion mas general, y se han creado nuevas areas de investigacion, tanto dentro como fuera de la economía financiera. Es interesante observar que las opciones mas comercializadas son las *opciones Americanas*, es decir, aquellas en que el tenedor tiene el derecho de ejercer la opcion en cualquier momento antes de que la opcion expire. De no ser así, es decir, cuando una opcion se puede ejercer solo en la fecha de expiracion, se le conoce como una *opción Europea*. Cuando el inversionista notifica a su agente su intencion de ejercer una opcion, si la opcion es una *call*, el inversionista debe comprar el activo subyacente (la accion) al llamado *precio de ejercicio* (si es una opcion *put*, el inversionista debe vender el activo subyacente al precio de ejercicio).

El precio de ejercicio se denota por  $K$  y la fecha de expiracion por  $T$ . Enfatizamos que una opcion da el derecho de hacer algo; sin embargo, el tenedor de la opcion no esta obligado a ejercer este derecho. Para comprar una opcion, el inversionista necesita pagar el precio de la opcion (la *prima*).

Sea  $S_T$  el precio de la accion (el activo subyacente) a la fecha de expiracion  $T$ . Es natural suponer que se desconoce el precio  $S_T$  al tiempo 0, así pues,  $S_T$  da aleatoriedad a nuestro modelo; y desde la perspectiva del tenedor de la opcion, la ganancia  $g$  en la fecha de expiracion  $T$  de una opcion Europea esta dada por la formula

$$g(S_T) = (S_T - K)^+ \doteq \max\{S_T - K, 0\},$$

Es decir, si en la fecha de expiracion  $T$  el precio de la accion es menor que el precio de ejercicio, el tenedor de la opcion puede comprar la accion directamente en el mercado, pagando menos del valor de  $K$ . En otras palabras, ser irracional

ejercer la opción, cuando menos para un inversionista que pre ere tener mas bienes que menos. Por otra parte, si en la fecha de expiracion el precio de la accion es mayor a  $K$ , el inversionista ejerce su derecho de comprar la accion al precio de ejercicio  $K$  pactado. As , al vender la accion inmediatamente en el mercado, el tenedor de la opción obtiene una ganancia inmediata de  $S_T - K$ . Paralelamente a la opción *call*, la opción *put* da el derecho al tenedor de vender la accion al precio de ejercicio y en la fecha de expiracion. Usando la misma notacion, se llega a la siguiente expresion de la ganancia  $h$  en la fecha de ejercicio  $T$  de una opción *put* Europea

$$h(S_T) = (K - S_T)^+ \doteq \max\{K - S_T, 0\}.$$

De aqu se sigue inmediatamente que las ganancias de la *call* y la *put* satisfacen la siguiente igualdad, la llamada paridad *put-call*

$$g(S_T) - h(S_T) = (S_T - K)^+ - (K - S_T)^+ = S_T - K.$$

Basicamente, la paridad *put-call* signi ca que el precio de una opción *put* Europea se determina del precio de una opción *call* Europea con el mismo precio de ejercicio y fecha de expiracion.

## 3.2. Futuros

Otros derivados importantes son los futuros y las opciones sobre futuros. Los futuros se ejercen sobre una amplia gama de bienes (azucar, trigo, naranja, oro, etcetera) e instrumentos nancieros (divisas, bonos, ndices, etcetera). Un punto interesante es considerar al futuro como la suma de dos opciones, una *call* y una *put*, con la misma prima y el mismo precio de ejercicio. Adoptar la posicion larga en futuros es comprar una *call* y emitir una *put*, corta en futuros es emitir una *call* y comprar una *put*. En particular en nuestro pa s, a la fecha, en el MexDer, se operan futuros sobre el IPC (Indice de Precios y Cotizaciones), sobre el dolar, sobre el Cete, sobre la TIIIE (Tasa de Inters Interbancaria de Equilibrio) y sobre acciones (Cemex, Femsa, Gcarso, Gfbb y Telmex). Para comercializar con futuros, las casas de bolsa especi can ciertas normas en el contrato. Los precios se reportan en la prensa nanciera, y se determinan por la oferta y la demanda. Si hay mas inversionistas que pre eran comprar futuros que venderlos, el precio se eleva; si ocurre lo contrario, el precio baja. Las ganancias y perdidas se rigen a traves de un procedimiento espec co diario conocido como *marking to market*. Un deposito inicial del inversionista, conocido como *margin inicial*, se ajusta diariamente para reflejar las ganancias y perdidas debidas a los movimientos en el precio de los futuros. Imaginemos por ejemplo a un inversionista que compra futuros (posicion larga) y que por tanto tiene el compromiso de comprar. Si hay un decremento en el precio del futuro, su cuenta de margen se reduce, su corredor tiene que pagar esta suma a la casa de bolsa y la casa de bolsa pasa este dinero al corredor de la parte que asumio la posicion corta (quien vendio los futuros). De manera similar,

si el precio de los futuros se eleva, los corredores de las partes con posiciones cortas pagan a la casa de bolsa, y los corredores de las partes con posición larga reciben el dinero de la casa de bolsa. De esta manera, el mercado es regulado diariamente al cierre de cada día (*marked to market*). Finalmente, al llegar el día de vencimiento la entrega se hace de la parte con posición corta. El precio recibido es generalmente el precio del futuro en el último momento en que se ajustó la cuenta de margen.

### 3.3. Contratos Adelantados

Un contrato adelantado (*forward*) es un compromiso para comprar o vender un bien en un tiempo futuro específico y a un precio pactado. Posiblemente este sea el derivado más antiguo que existe, y que en su origen se empleaba para asegurar un precio de las cosechas agrícolas. Una de las partes de un contrato *forward* asume una posición larga y se compromete a comprar el activo subyacente en una fecha futura específica al *precio de entrega*; la otra parte asume una posición corta y se compromete a vender el bien en la misma fecha y al mismo precio. Al momento en que se realiza el contrato, se determina el precio de entrega para que el valor del contrato *forward* para ambas partes sea cero. Por tanto algunos de los aspectos de los *forwards* se parecen a los de los futuros; sin embargo, los *forwards* no se cotizan en bolsa como los futuros, que además son estandarizados. Las bolsas más importantes para los futuros son la CBOT y el CME ambas en EUA. En el caso de México la bolsa correspondiente es el MexDer.

Los *swaps* son el intercambio de flujos entre dos participantes, si bien algunos lo pueden entender como una serie de *forwards*. No obstante, en sí, el *swap* se puede visualizar como una representación de cualquier instrumento o derivado, toda vez que lo que buscan los participantes es precisamente eso, intercambiar flujos (y en el que cada quien asume el riesgo de que se de o no ese flujo). Estos a su vez pueden o no ser estandarizados. Aquí lo más común es que no lo sean y que cada uno de ellos sea “un traje a la medida”. Un aspecto de los *swaps*, es que se puede determinar mejor la época de inicio de este derivado: precisamente después del colapso de los Acuerdos de *Bretton Woods*. De 1944 y hasta 1971, las finanzas en el mundo occidental, estuvieron basadas en los llamados Acuerdos de Bretton Woods, mediante los cuales había nacido el Fondo Monetario Internacional (FMI). Los países miembros se habían comprometido a fijar sus respectivas monedas en relación al oro y al dólar estadounidense (35 dólares = 1 onza de oro). Una parte relevante del acuerdo fue que EUA, cuya moneda, el dólar, se iba a utilizar para transacciones internacionales, no podía emitir dinero nuevo sin el respaldo pertinente en reservas de oro. Para fines prácticos, los acuerdos significaron que las principales variables financieras permanecieran fijas, i.e., la paridad de las monedas era fija y la tasa de interés también era fija. Dicho de otra forma, en ese lapso de tiempo, el riesgo financiero preocupaba muy someramente a las personas y, en consecuencia, a las empresas. En dicha época, el papel de las finanzas fue más bien secundario, y la generación de ri-

queza estaba mas orientada a la forma de producir y de distribuir. A principios de la decada de 1970, en particular en EUA, se tuvieron indicios de que no iba a ser posible una transicion tranquila: para EUA esto signi caba que su dolar cada vez valiera menos y al no vislumbrarse medidas, que no fueran drasticas, se empezo a alentar la especulacion. Por ejemplo, se compraban dolares a una paridad fija mediante el pago de una moneda europea, y se cambiaban al gobierno de EUA por onzas de oro (pagando 35 dolares por cada una), el cual se vendia en el propio EUA o en otros paises a un precio mucho mayor, o incluso se podian guardar previendo un incremento del precio del oro. El dolar empezo a no ser conveniente como reserva internacional para los bancos centrales europeos, en su lugar empezaron a preferir las reservas constituidas en oro, lo que provoco que las reservas de oro de EUA empezaran a bajar. A principios de 1971 se da un desprendimiento general a gran escala de dolares, lo que se convierte rapidamente en panico. Como resultado de esta crisis, se suprime la paridad fija, y quedan flotando, con respecto al dolar, el marco aleman, el dolar canadiense y el florin holandes, es decir que el tipo de cambio de dichas monedas respecto al dolar estadounidense podia variar de un minuto a otro. Esto era muy novedoso para la epoca. No obstante, los movimientos especulativos en torno a la moneda estadounidense continuaron y se incrementaron a tal grado que obligaron, el 15 de agosto de 1971, al presidente de EUA, Richard M. Nixon, en el marco de un gran deficit presupuestario, a anunciar un plan historico de medidas que afectaban tanto al interior como al exterior, y que, en sintesis, era desvincular el dolar del oro, y con ello terminar con la esencia de los Acuerdos de Bretton Woods. En la practica esto signi co la devaluacion del dolar frente al oro (a 38 dolares la onza), la confirmacion de que la paridad de las monedas ya no iba a ser fija, y ademas, que las tasas de interes empezaran a variar de manera mas frecuente, y en algunos casos tampoco ser fija. Un efecto complementario que se genero, y que despues ha cobrado gran relevancia, fue la introduccion de plazos cada vez menores, a fin de atraer o retener las inversiones. A la par de esta situacion, se producen dos fenomenos tecnologicos: el desarrollo de la computacion y de las telecomunicaciones, cuya notoria evolucion permiten que el dinero se pueda mover a una velocidad asombrosa. A tal grado que, en la actualidad, un pais, que no cuenta con estas tecnologas, este expuesto a ser saqueado en pocos minutos. Los eventos anteriores han dado como resultado el mundo financiero tal como lo conocemos en la actualidad, y que se puede resumir en una palabra: la volatilidad. Ver [2] y [1].

### 3.4. La Fórmula de Samuelson

No obstante el gran numero de formas de arribar a la formula de Black-Scholes, una forma peculiar de llegar a ella es a traves de la formula de Samuelson.

Consideremos un modelo financiero de un periodo en el que hay una accion con valor  $S_T$  al final del periodo y valor inicial  $S_0$ . Un bono de precio unitario con ganancia  $e^{rT}$  y sea  $g(S_T)$  el valor de un derivado del tipo Europeo al tiempo

de ejercicio.

**Teorema 3.4.1.** *El precio  $p(g)$  de un derivado  $g(S_T)$  está dado por*

$$p(g) = e^{-rT} \mathbb{E}_{\mathbb{P}} \left[ g \left( \frac{S_T}{\mathbb{E}S_T} S_0 e^{rT} \right) \right]. \quad (3.1)$$

En donde  $\mathbb{E}_{\mathbb{P}}$  es la esperanza respecto de la medida física: la distribución de  $S_T$  es cualquiera. La verificación del teorema se encuentra en [12]. Como caso particular podemos obtener la fórmula de Black-Scholes si suponemos el modelo lognormal, ecuación (2.4) del ejemplo 2.2.1.

$$\begin{aligned} p(g) &= e^{-rT} \mathbb{E}_{\mathbb{P}} \left[ g \left( \frac{S_0 e^{(\mu - \frac{1}{2}\sigma^2)T + \sigma B_T}}{S_0 e^{\mu T}} S_0 e^{rT} \right) \right] \\ &= e^{-rT} \mathbb{E} \left[ g \left( S_0 e^{(r - \frac{\sigma^2}{2})T + \sigma B_T} \right) \right] \end{aligned}$$

que es la ecuación (2.17) del ejemplo 2.3.4. Así,

$$p(g) = S_0 (d_1) - K e^{-rT} (d_2). \quad (3.2)$$

### 3.5. El Modelo Modulado de Markov

El objetivo fundamental del modelo modulado de Markov es la valuación de derivados en ambientes cambiantes. Sea  $X = X(t)$  un proceso de Markov homogéneo que toma valores en  $E = \{0, 1, 2, \dots, p\}$ , cuya dinámica está gobernada por la matriz de intensidad  $\Lambda = (\lambda_{jk})$ . Así, decimos que el proceso  $X$  de  $n$  ambientes. Suponemos que  $X$  es irreducible, y como el espacio de estados es finito, existe una única distribución inicial estacionaria de  $X$  denotada por  $\pi = (\pi_1, \dots, \pi_p)$ . Observe que los valores de  $X$  pueden ser constantes por pedazos. El modelo modulado de Markov, ver [13], en tiempo continuo es un modelo donde  $\{J_t\}$  es el ambiente subyacente que dicta las condiciones del mercado y en un estado particular el modelo se comporta como el modelo de Black-Scholes con parámetros que dependen del estado subyacente particular. Cuando el estado cambia, otro proceso de Black-Scholes con parámetros nuevos determina la evolución. Es decir, sean  $\mu_i$ ,  $\sigma_i$  y  $r_i$  el rendimiento compuesto continuo e instantánea, la volatilidad y la tasa libre de riesgo cuando el ambiente dictado por  $\{J_t\}$  está en el estado  $i$ .

Sea  $f(t_1, \dots, t_p; T)$  la función de densidad de los tiempos de ocupación en los estados  $0, 1, 2, \dots, p$ . Calcular la densidad no es fácil, pero en el caso de dos estados se hizo en el ejemplo 1.2.4.

Dado que el proceso pasa un tiempo  $t_i$  en el estado  $i$ ,  $i = 1, \dots, p$  y  $t_1 + \dots + t_p = T$ , entonces el precio  $p(g|t_1, \dots, t_p)$  de un derivado de la forma  $g(S_T)$



esta dado por (ver [13] para la obtencion de la ecuacion),

$$\begin{aligned} p(g|t_1, \dots, t_p) &= \left( \prod_{i=1}^p e^{-r_i t_i} \right) \int_{\mathbb{R}^p} g \left[ S_0 \prod_{i=1}^p e^{(r_i - \frac{1}{2}\sigma_i^2)t_i + \sigma_i y_i} \right] dF_{t_i}(y_i) \\ &= \left( \prod_{i=1}^p e^{-r_i t_i} \right) \mathbb{E} \left( g \left[ S_0 \prod_{i=1}^p e^{(r_i - \frac{1}{2}\sigma_i^2)t_i + \sigma_i Z_i} \right] \right) \end{aligned}$$

donde  $F_t$  es la funcion de distribucion de una normal con media cero y varianza  $t$  y  $Z_i$  son variables aleatorias normales e independientes con media cero y varianza  $t_i$ . Si de nimos

$$\begin{aligned} r^* &= \frac{1}{T} \sum_{i=1}^p r_i t_i, \\ (\sigma^*)^2 &= \frac{1}{T} \sum_{i=1}^p \sigma_i^2 t_i, \end{aligned}$$

y

$$\sigma^* B_T \sim N(0, (\sigma^*)^2 T)$$

entonces

$$p(g|t_1, \dots, t_p) = \left( e^{-r^* T} \right) \mathbb{E} \left( g \left[ S_0 e^{(r^* - \frac{1}{2}(\sigma^*)^2)T + \sigma^* B_T} \right] \right).$$

Identifiquemos esta ecuacion como la formula de Black-Scholes. Sea  $BS(g, S_0, r, \sigma^2, T)$  el precio del derivado  $g(S_T)$  segun la formula de Black-Scholes. Entonces

$$\begin{aligned} p(g|t_1, \dots, t_p) &= BS(g, S_0, r^*, (\sigma^*)^2, T) \\ &= S_0 (d_1) - K e^{-\sum_{i=1}^p r_i t_i} (d_2) \end{aligned} \quad (3.3)$$

con

$$\begin{aligned} d_1 &= \frac{\ln(S_0/K) + (r^* + \frac{1}{2}(\sigma^*)^2)T}{\sigma^* \sqrt{T}}, \\ d_2 &= \frac{\ln(S_0/K) + (r^* - \frac{1}{2}(\sigma^*)^2)T}{\sigma^* \sqrt{T}} \end{aligned}$$

Llamaremos a la ecuacion (3.3) la formula de Black-Scholes *ponderada*.

Si no se conocen las trayectorias, lo que comunmente sucede, entonces el precio del derivado  $p(g)$  estara dado por el promedio de los precios sobre todas las posibles trayectorias. Y como solo importa el tiempo que se pasa en cada estado, obtenemos la siguiente expresion para el precio del derivado

$$\begin{aligned} p(g) &= \int_{D(T)} p(g|t_1, \dots, t_p) f(t_1, \dots, t_p; T) dt_1 \dots dt_p \\ &= \mathbb{E}_{\mathbb{T}_{t_1, \dots, t_p}} [BS(g, S_0, r^*, (\sigma^*)^2, T)], \end{aligned} \quad (3.4)$$

donde  $D(T) = \{(t_1, \dots, t_p) | t_i \geq 0, i = 1, \dots, p, t_1 + \dots + t_p = T\}$ . Llamaremos a la ecuacion (3.4) la formula de Black-Scholes con *densidad de tiempos de ocupación*. Así, el precio de un derivado en el modelo modulado de Markov se puede descomponer en precios dados por la formula de Black-Scholes.

### 3.6. Valuación del derivado para 2 estados

En [12] se presenta un ejemplo para calcular el precio del derivado en dos estados con tres diferentes matrices de intensidad, haciendo uso de la ecuación (3.4), con función de densidad dada por (1.16); y con los siguientes datos. Precio inicial de la acción  $S_0 = 100$ , tasa de interés<sup>1</sup> por periodo 5% en el estado uno, 15% en el estado dos, volatilidades  $\sigma_1 = 0.15$  y  $\sigma_2 = 0.25$ . El precio de ejercicio es  $K$ . Los valores son para la opción *call* Europea. *Exacto*, el valor de la *call* en dos estados; *BS-uno*, el valor de la *call* en el estado uno; *BS-dos*, el valor de la *call* en el estado dos.

La ecuación (3.4) en dos estados se puede fácilmente calcular en MATLAB creando primeramente el código para la función de los tiempos de ocupación, ecuación (1.16), en el archivo `densh.m`.

```
% Densidad del tiempo de ocupacion para el proceso en dos estados.
function y= densh(x,q1,q2,T)
% q1 y q2 son los valores de la matriz de intensidad.
% T es el tiempo.

y=exp(-q1.*x-q2.*(T-x)).*...
  (sqrt((q1*q2.*x)/(T-x)).*besseli(1,2*sqrt(q1*q2.*x.*(T-x)))+...
  q1*besseli(0,2*sqrt(q1*q2.*x.*(T-x)))));
  %besseli es la funcion de Bessel
```

El archivo `bsm.m` guardara los valores de la tasa de rendimiento y de la volatilidad en cada estado.

```
function y= bsm(x,K,q1,q2,T)
% K, precio de ejercicio
% Datos del ejemplo
S=100;
r1=log(1+0.05);
r2=log(1+0.15);
sig1=log(1+0.15);
sig2=log(1+0.25);

% Ponderacion
t1=x;
t2=T-x;
r=(r1.*t1+r2.*t2)./T;
sig=sqrt((sig1.^2.*t1+sig2.^2.*t2)./T);
% blsprice calcula el precio de la call y de la put
% segun la formula de Black-Scholes
[c,p]=blsprice(S,K,r,T,sig,0);
y=c.*densh(x,q1,q2,T);
```

<sup>1</sup>La tasa libre de riesgo es la tasa  $r_t = \log(1 + R_t) = \log S_t/S_{t-1}$ ; con  $R_t = S_t/S_{t-1} - 1$

El valor de la *call* con  $(\lambda_1, \lambda_2, 1) = (1, 1, 1)$ , precio de ejercicio de 110 y precio del subyacente de 100 se obtiene integrando la funcion `bsm`:

```
c=quad('bsm(x,110,1,1,1)',0,1)
```

La probabilidad de permanecer en el estado uno la calculamos tecleando:

```
c1=blsprice(100,110,log(1+0.05),1,log(1+0.15),0)*exp(-1*1)
```

Al sumar `c` con `c1` se obtiene el resultado 5.7443.

$(\lambda_1, \lambda_2, t) = (1, 1, 1)$			
$K$	Exacto	BS-uno	BS-dos
110	5.75	3.64	11.04
100	10.57	8.15	16.39
90	17.50	15.17	23.09

$(\lambda_1, \lambda_2, t) = (1, 2, 1)$			
$K$	Exacto	BS-uno	BS-dos
110	5.33	3.64	11.04
100	10.11	8.15	16.39
90	17.05	15.17	23.09

$(\lambda_1, \lambda_2, t) = (2, 1, 1)$			
$K$	Exacto	BS-uno	BS-dos
110	7.03	3.64	11.04
100	12.03	8.15	16.39
90	18.89	15.17	23.09

$(\lambda_1, \lambda_2, t) = (1, 1, 10)$			
$K$	Exacto	BS-uno	BS-dos
110	57.14	36.14	73.26
100	60.68	40.93	75.58
90	64.35	46.09	77.94

$(\lambda_1, \lambda_2, t) = (1, 2, 10)$			
$K$	Exacto	BS-uno	BS-dos
110	51.23	36.14	73.26
100	55.17	40.93	75.58
90	59.29	46.09	77.94

$(\lambda_1, \lambda_2, t) = (2, 1, 10)$			
$K$	Exacto	BS-uno	BS-dos
110	63.16	36.14	73.26
100	66.27	40.93	75.58
90	69.47	46.09	77.94

### 3.7. Aproximación del derivado para $n$ estados

En un proceso  $X(t)$  de nacimiento y muerte, cada partícula vive una cantidad de tiempo aleatoria al final de la cual se divide o muere. Si hay  $i$  ( $i \in \mathbb{N} \cup \{0\}$ ) individuos en la población, del estado  $i$  el proceso salta a  $i + 1$  si ocurre una división o a  $i - 1$  si la partícula muere sin dividirse. Entonces, el tiempo de vida del individuo tiene una distribución exponencial con parámetro  $\lambda_i$ ; y el proceso brinca a  $i + 1$  con probabilidad de transición  $p_{i,i+1}$  y con probabilidad complementaria  $p_{i,i-1}$  brinca a  $i - 1$ .

Usualmente, el proceso de nacimiento y muerte se describe en términos de sus tasas de nacimiento y muerte; en una población de tamaño  $i$ , un individuo nace con una tasa  $\beta_i$  y muere con una tasa  $\delta_i$ ; *i.e.*,

$$\mathbb{P}(X(t+h) = i+1 | X(t) = i) = \beta_i h + \mathcal{O}(h),$$

$$\mathbb{P}(X(t+h) = i-1 | X(t) = i) = \delta_i h + \mathcal{O}(h),$$

y con probabilidad complementaria (cuando no ocurren ni nacimientos ni muertes en  $(t, t+h)$ ) el proceso permanece sin cambios

$$\mathbb{P}(X(t+h) = i | X(t) = i) = 1 - (\beta_i + \delta_i)h + \mathcal{O}(h).$$

Es decir, hablamos de las probabilidades del generador infinitesimal  $\mathbf{A} = \{q_{ij}\}_{i,j \in E}$ , ver teorema 1.2.4. Entonces,

$$q_{i,i+1} = \beta_i, \quad q_{i,i-1} = \delta_i I_{\{i \geq 1\}},$$

$$q_{ij} = 0$$

si  $j \notin \{i-1, i, i+1\}$  y,

$$q_i = \beta_i + \delta_i$$

por conservatividad, de acuerdo con el teorema 1.2.5. Así,

$$\mathbf{A} = \begin{pmatrix} -\beta_0 & \beta_0 & 0 & 0 & \dots & \\ \delta_1 & -(\beta_1 + \delta_1) & \beta_1 & 0 & 0 & \dots \\ 0 & \delta_2 & -(\beta_2 + \delta_2) & \beta_2 & 0 & \dots \\ 0 & 0 & \delta_3 & -(\beta_3 + \delta_3) & \beta_3 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \quad (3.5)$$

Cada vez que el proceso entra en el estado  $i$ , la cantidad de tiempo que pasa en ese estado antes de hacer una transición a un estado diferente tiene una

distribucion exponencial con parametro de intensidad  $q_{ii} = q_i$  y con media  $1/q_i$ ; como puede verificarse en el teorema 1.2.10 b). De igual forma, por el teorema 1.2.10 a) encontramos que la matriz de transicion  $\mathbf{P}$  viene dada por la estructura regenerativa  $p_{i,i+1} = q_{i,i+1}/q_i = \beta_i/(\beta_i + \delta_i)$ ,  $p_{i,i-1} = q_{i,i-1}/q_i = \delta_i/(\beta_i + \delta_i)$  y  $p_{0,1} = 1$ , es decir,

$$\mathbf{P} = \begin{pmatrix} 0 & 1 & 0 & 0 & \dots \\ \frac{\delta_1}{\beta_1 + \delta_1} & 0 & \frac{\beta_1}{\beta_1 + \delta_1} & 0 & \dots \\ 0 & \frac{\delta_2}{\beta_2 + \delta_2} & 0 & \frac{\beta_2}{\beta_2 + \delta_2} & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

Sea  $\mu_{ij}(t)$  la media de los tiempos de ocupacion que el proceso pasa en el estado  $j$  durante  $(0, t]$  habiendo iniciado en  $X(0) = i$  ( $i, j = 1, 2, \dots, p$ ). Definimos el proceso

$$I_{ij}(t) = \begin{cases} 1, & \text{si } \{X(t) = j | X(0) = i\} \\ 0, & \text{en otro caso.} \end{cases}$$

Claramente,

$$\mathbb{P}(I_{ij}(t) = 1) = p_{ij}(t)$$

y

$$\mathbb{E}[I_{ij}(t)] = p_{ij}(t),$$

entonces,

$$\begin{aligned} \mu_{ij}(t) &= \mathbb{E} \int_0^t I_{ij}(\cdot) d\cdot \\ &= \int_0^t \mathbb{E}[I_{ij}(\cdot)] d\cdot \\ &= \int_0^t p_{ij}(\cdot) d\cdot. \end{aligned} \quad (3.6)$$

Cuando  $t \rightarrow \infty$ , obtenemos por la version continua del teorema ergodico (teorema 1.1.3)

$$\lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t p_{ij}(\cdot) d\cdot = \frac{1}{\mu_j} = \pi(j). \quad (3.7)$$

La aproximacion al valor de la *call* para  $n$  estados consiste entonces en encontrar el semigrupo de transicion  $\{\mathbf{P}(t)\}_{t \geq 0}$  para la matriz de intensidad dada en (3.5). Suponiendo que  $n$  es finito  $\mathbf{P}(t)$  se puede encontrar como solucion de la ecuacion regresiva  $\frac{d}{dt} \mathbf{P}(t) = \mathbf{\Lambda} \mathbf{P}(t)$  o de la ecuacion progresiva (1.11) con condicion inicial  $\mathbf{P}(0) = \mathbf{I}$ . Recordemos que la solucion es

$$\mathbf{P}(t) = e^{t\mathbf{\Lambda}},$$

donde la exponencial de la matriz  $t\mathbf{\Lambda}$  esta definida por

$$e^{t\mathbf{\Lambda}} \doteq \sum_{n=0}^{\infty} \frac{t^n}{n!} \mathbf{\Lambda}^n = \mathbf{I} + \sum_{n=1}^{\infty} \frac{t^n}{n!} \mathbf{\Lambda}^n. \quad (3.8)$$

Cuando  $\Lambda$  es una matriz finita, la serie en (3.8) es convergente y diagonalizable, por tanto podemos escribir a  $\Lambda$  como

$$\Lambda = \mathbf{V}\mathbf{D}\mathbf{V}^{-1},$$

donde  $\mathbf{D}$  es una matriz diagonal con diferentes eigenvalores  $\{\lambda_1, \lambda_2, \dots, \lambda_n\}$  de  $\Lambda$  como elementos. As, el semigrupo de transicion es

$$\begin{aligned} \mathbf{P}(t) = e^{t\mathbf{V}\mathbf{D}\mathbf{V}^{-1}} &= \mathbf{I} + \sum_{n=1}^{\infty} \frac{\mathbf{V}(\mathbf{D}t)\mathbf{V}^{-1}}{n!} \\ &= \mathbf{V} \left[ \mathbf{I} + \sum_{n=1}^{\infty} \frac{(\mathbf{D}t)^n}{n!} \right] \mathbf{V}^{-1} \\ &= \mathbf{V}e^{\mathbf{D}t}\mathbf{V}^{-1}, \end{aligned} \tag{3.9}$$

donde

$$e^{\mathbf{D}t} = \begin{pmatrix} e^{\lambda_1 t} & & & 0 \\ & e^{\lambda_2 t} & & \\ & & \ddots & \\ 0 & & & e^{\lambda_n t} \end{pmatrix}.$$

### 3.7.1. La desigualdad de Jensen

La graficas 3.1 y 3.2 sugieren que  $p(g|t_1, \dots, t_p)$  puede ser una funcion concava o convexa, de no ser este el caso podemos entonces aproximarla por una funcion lineal. La desigualdad de Jensen nos ayuda a encontrar una cota para el precio del derivado condicionado a comenzar en el estado  $i$ .

Una funcion  $\varphi(x)$  es convexa si  $\varphi(\lambda x + (1 - \lambda)y) \leq \lambda\varphi(x) + (1 - \lambda)\varphi(y)$  para todo  $x, y$  en  $\mathbb{R}$ , y  $0 < \lambda < 1$ . La funcion  $\varphi(x)$  es concava si  $-\varphi(x)$  es convexa. De manera mas formal, las funciones convexas yacen debajo de las rectas que conectan cualesquiera dos puntos de su grafica: mientras  $\lambda$  va de 0 a 1,  $\lambda\varphi(x_1) + (1 - \lambda)\varphi(x_2)$  define una recta que conecta a  $\varphi(x_1)$  con  $\varphi(x_2)$ . Esta recta yace encima de  $\varphi(x)$  si  $\varphi(x)$  es convexa. Mas aun, una funcion convexa yace encima de todas sus rectas tangentes y este hecho es la base de la desigualdad de Jensen.

**Teorema 3.7.1 (Desigualdad de Jensen).** *Para cualquier vector aleatorio  $(X_1, X_2, \dots, X_p)$ , si  $\varphi(x_1, x_2, \dots, x_p) : \mathbb{R}^p \rightarrow \mathbb{R}$  es una función convexa entonces*

$$\varphi(\mathbb{E}X_1, \mathbb{E}X_2, \dots, \mathbb{E}X_p) \leq \mathbb{E}\varphi(X_1, X_2, \dots, X_p).$$

En efecto, sea  $h(x_1, x_2, \dots, x_p)$  el hiperplano tangente a  $\varphi(x_1, x_2, \dots, x_p)$  en el punto  $\varphi(\mathbb{E}X_1, \mathbb{E}X_2, \dots, \mathbb{E}X_p)$ . Escribamos  $h(x_1, x_2, \dots, x_p) = a_0 + a_1x_1 + a_2x_2 + \dots + a_px_p$  para ciertas constantes  $a_0, a_1, a_2, \dots, a_p$ . Ahora, por la convexidad de  $\varphi$  tenemos que  $\varphi(x_1, x_2, \dots, x_p) \geq a_0 + a_1x_1 + a_2x_2 + \dots + a_px_p$  y como la esperanza preserva las desigualdades

$$\begin{aligned}
\mathbb{E}\varphi(X_1, X_2, \dots, X_p) &\geq \mathbb{E}(a_0 + a_1X_1 + a_2X_2 + \dots + a_pX_p) \\
&= a_0 + a_1\mathbb{E}X_1 + a_2\mathbb{E}X_2 + \dots + a_p\mathbb{E}X_p \\
&= h(\mathbb{E}X_1, \mathbb{E}X_2, \dots, \mathbb{E}X_p) \\
&= \varphi(\mathbb{E}X_1, \mathbb{E}X_2, \dots, \mathbb{E}X_p),
\end{aligned}$$

puesto que  $h$  es tangente en el punto  $(\mathbb{E}X_1, \mathbb{E}X_2, \dots, \mathbb{E}X_p)$ .

De esta forma, si  $p(g|t_1, \dots, t_p)$  es concava, digamos, entonces

$$\begin{aligned}
\mathbb{E}_{\mathbb{T}_{t_1, \dots, t_p}}[BS(g, S_0, r^*, (\sigma^*)^2, T)|J_0 = i] &\leq \\
BS\left(g, S_0, \mathbb{E}_{\mathbb{T}_{t_1, \dots, t_p}}[r^*|J_0 = i], \mathbb{E}_{\mathbb{T}_{t_1, \dots, t_p}}[(\sigma^*)^2|J_0 = i], T\right).
\end{aligned}$$

Ahora,

$$\begin{aligned}
\mathbb{E}_{\mathbb{T}_{t_1, \dots, t_p}}[r^*|J_0 = i] &= \mathbb{E}_{\mathbb{T}_{t_1, \dots, t_p}}\left[\frac{1}{T} \sum_{k=1}^p r_k t_k | J_0 = i\right] \\
&= \frac{1}{T} \sum_{k=1}^p r_k \mathbb{E}_{\mathbb{T}_{t_1, \dots, t_p}}[t_k | J_0 = i] \\
&= \frac{1}{T} \sum_{k=1}^p r_k \mu_{ik}(T) = \bar{r}_i^*,
\end{aligned}$$

de igual manera

$$\mathbb{E}_{\mathbb{T}_{t_1, \dots, t_p}}[(\sigma^*)^2|J_0 = i] = \frac{1}{T} \sum_{k=1}^p \sigma_k^2 \mu_{ik}(T) = (\bar{\sigma}^*)_i^2;$$

as

$$\mathbb{E}_{\mathbb{T}_{t_1, \dots, t_p}}[BS(g, S_0, r^*, (\sigma^*)^2, T)|J_0 = i] \leq BS(g, S_0, \bar{r}_i^*, (\bar{\sigma}^*)_i^2, T)$$

Razon por la cual  $BS(g, S_0, \bar{r}_i^*, (\bar{\sigma}^*)_i^2, T)$  siempre sobrevaluara el precio del derivado cuando sea concava.

### 3.7.2. La griega $\Theta$

La de un derivado financiero es la tasa de cambio del derivado respecto del tiempo manteniendo constante todo lo demas. Para la opcion *call*, al derivar la formula de Black-Scholes respecto del tiempo obtenemos

$$= -\frac{S_0 \varphi(d_1) \sigma}{2\sqrt{T-t}} - rK e^{-r(T-t)} (d_2), \quad (3.10)$$

con  $(\cdot)$  la distribucion y  $\varphi(\cdot)$  la densidad de la normal estandar.

El analisis de convexidad para la formula de Black-Scholes ponderada lo realizaremos a traves de la derivada de .

Observando la condicion  $T = t_1 + \dots + t_p$  en la formula de Black-Scholes ponderada, ecuacion (3.3), podemos escribir

$$\begin{aligned} r^*T &= r_p T + (r_1 - r_p)t_1 + (r_2 - r_p)t_2 + \dots + (r_{p-1} - r_p)t_{p-1} \\ &= R_p, \end{aligned}$$

de forma similar

$$\begin{aligned} (\sigma^*)^2 T &= \sigma_p^2 T - (\sigma_p^2 - \sigma_1^2)t_1 - \dots - (\sigma_p^2 - \sigma_{p-1}^2)t_{p-1} \\ &= T_\sigma. \end{aligned}$$

Entonces

$$\begin{aligned} (r^* + \frac{1}{2}(\sigma^*)^2)T &= (r_p + \frac{1}{2}\sigma_p^2)T \\ &\quad - [(r_p - r_1) + \frac{1}{2}(\sigma_p^2 - \sigma_1^2)]t_1 - [(r_p - r_2) + \frac{1}{2}(\sigma_p^2 - \sigma_2^2)]t_2 \\ &\quad - \dots - [(r_p - r_{p-1}) + \frac{1}{2}(\sigma_p^2 - \sigma_{p-1}^2)]t_{p-1} \\ &= m_p T - m_{p1}t_1 - m_{p2}t_2 - \dots - m_{pp-1}t_{p-1} \\ &= T_r, \end{aligned}$$

y

$$\begin{aligned} (r^* - \frac{1}{2}(\sigma^*)^2)T &= (r_p - \frac{1}{2}\sigma_p^2)T \\ &\quad - [(r_p - r_1) - \frac{1}{2}(\sigma_p^2 - \sigma_1^2)]t_1 - [(r_p - r_2) - \frac{1}{2}(\sigma_p^2 - \sigma_2^2)]t_2 \\ &\quad - \dots - [(r_p - r_{p-1}) - \frac{1}{2}(\sigma_p^2 - \sigma_{p-1}^2)]t_{p-1} \\ &= m_p^- T - m_{p1}^- t_1 - m_{p2}^- t_2 - \dots - m_{pp-1}^- t_{p-1} \\ &= T_r^-. \end{aligned}$$

La formula de Black-Scholes ponderada la podemos entonces escribir como

$$p(g|t_1, \dots, t_p) = S_0 (d_1) - K e^{-R_p} (d_2), \quad (3.11)$$

donde

$$\begin{aligned} d_1(t_1, \dots, t_p) &= \frac{\ln(S_0/K) + T_r}{\sqrt{T_\sigma}}, \\ d_2(t_1, \dots, t_p) &= \frac{\ln(S_0/K) + T_r^-}{\sqrt{T_\sigma}}. \end{aligned}$$



Ahora la para la formula de Black-Scholes ponderada en dos estados la podemos calcular si observamos lo siguiente:

$$\begin{aligned} d_1(t_1, t_2) &= \frac{\ln(S_0/K) + m_2 T}{\sqrt{T_\sigma}} - \frac{m_{21} t_1}{\sqrt{T_\sigma}} \\ \frac{dd_1(t_1, t_2)}{dt_1} &= \frac{(\sigma_2^2 - \sigma_1^2)}{2T_\sigma} \left[ d_1(t_1, t_2) - \frac{2T_\sigma m_{21}}{\sqrt{T_\sigma}(\sigma_2^2 - \sigma_1^2)} \right], \end{aligned}$$

de igual forma para  $d_2(t_1, t_2)$

$$\begin{aligned} d_2(t_1, t_2) &= \frac{\ln(S_0/K) + m_2^- T}{\sqrt{T_\sigma}} - \frac{m_{21}^- t_1}{\sqrt{T_\sigma}} \\ \frac{dd_2(t_1, t_2)}{dt_1} &= \frac{(\sigma_2^2 - \sigma_1^2)}{2T_\sigma} \left[ d_2(t_1, t_2) - \frac{2T_\sigma m_{21}^-}{\sqrt{T_\sigma}(\sigma_2^2 - \sigma_1^2)} \right]; \end{aligned}$$

las diferencias

$$d_1(t_1, t_2) - d_2(t_1, t_2) = \frac{T_\sigma}{\sqrt{T_\sigma}}$$

y

$$\frac{dd_1(t_1, t_2)}{dt_1} - \frac{dd_2(t_1, t_2)}{dt_1} = -\frac{(\sigma_2^2 - \sigma_1^2)}{2\sqrt{T_\sigma}};$$

la relacion entre  $d_1(t_1, t_2)$  y  $d_2(t_1, t_2)$  con la densidad de la normal estandar

$$\varphi(d_2(t_1, t_2)) = \varphi(d_1(t_1, t_2)) \frac{S_0}{K} e^{R_2}.$$

De esta forma

$$\begin{aligned} &= \frac{dp(g|t_1, t_2)}{dt_1} \\ &= S_0 \varphi(d_1(t_1, t_2)) \frac{dd_1(t_1, t_2)}{dt_1} - K e^{-R_2} \varphi(d_2(t_1, t_2)) \frac{dd_2(t_1, t_2)}{dt_1} \\ &\quad + (r_1 - r_2) K (d_2(t_1, t_2)) e^{-R_2} \\ &= -S_0 \varphi(d_1(t_1, t_2)) \frac{(\sigma_2^2 - \sigma_1^2)}{2\sqrt{T_\sigma}} - (r_2 - r_1) K e^{-R_2} (d_2(t_1, t_2)), \end{aligned}$$

y para la derivada de respecto de  $t_1$  obtenemos

$$\begin{aligned} ' &= \frac{S_0(\sigma_2^2 - \sigma_1^2)\varphi(d_1(t_1, t_2))}{2\sqrt{T_\sigma}} \left[ d_1(t_1, t_2) \frac{dd_1(t_1, t_2)}{dt_1} - \frac{(\sigma_2^2 - \sigma_1^2)}{2T_\sigma} \right] \\ &\quad - S_0(r_2 - r_1)\varphi(d_1(t_1, t_2)) \frac{dd_2(t_1, t_2)}{dt_1} - (r_2 - r_1)^2 K (d_2(t_1, t_2)) e^{-R_2}. \end{aligned}$$

La matriz Hessiana  $H_t = (q_{ij}(t))$ ,  $q_{ij} = \frac{\partial^2 f}{\partial t_i \partial t_j}(t_1, t_2, \dots, t_p)$ , de la formula de Black-Scholes en  $n (< p)$  estados se obtiene siguiendo un procedimiento paralelo al desarrollado:

$$\begin{aligned} a_{ii} &= \frac{\partial^2 p(g|t_1, \dots, t_p)}{\partial^2 t_i} \\ &= \frac{S_0(\sigma_p^2 - \sigma_i^2)\varphi(d_1(t_1, \dots, t_p))}{2\sqrt{T_\sigma}} \left[ d_1(t_1, \dots, t_p) \frac{\partial d_1(t_1, \dots, t_p)}{\partial t_i} \right. \\ &\quad \left. - \frac{(\sigma_p^2 - \sigma_i^2)}{2T_\sigma} \right] \\ &\quad - (r_p - r_i)S_0\varphi(d_1(t_1, \dots, t_p)) \frac{\partial d_2(t_1, \dots, t_p)}{\partial t_i} \\ &\quad - (r_p - r_i)^2 K (d_2(t_1, \dots, t_p))e^{-R_p}. \end{aligned}$$

Notamos que,  $\frac{\partial^2 p(g|t_1, \dots, t_p)}{\partial t_i \partial t_j}$  y  $\frac{\partial^2 p(g|t_1, \dots, t_p)}{\partial t_j \partial t_i}$  son funciones continuas en  $D(T)$  y por tanto iguales.

$$\begin{aligned} a_{ji} &= \frac{\partial^2 p(g|t_1, \dots, t_p)}{\partial t_j \partial t_i} \\ &= \frac{S_0(\sigma_p^2 - \sigma_i^2)\varphi(d_1(t_1, \dots, t_p))d_1(t_1, \dots, t_p)}{2\sqrt{T_\sigma}} \frac{\partial d_1(t_1, \dots, t_p)}{\partial t_j} \\ &\quad - \frac{S_0(\sigma_p^2 - \sigma_i^2)(\sigma_p^2 - \sigma_j^2)\varphi(d_1(t_1, \dots, t_p))}{4\sqrt{T_\sigma T_\sigma}} \\ &\quad - (r_p - r_i)S_0\varphi(d_1(t_1, \dots, t_p)) \frac{\partial d_2(t_1, \dots, t_p)}{\partial t_j} \\ &\quad - (r_p - r_i)(r_p - r_j)K (d_2(t_1, \dots, t_p))e^{-R_p}. \end{aligned}$$

Algunos ejemplos clasicos de convexidad en la linea real se obtienen del siguiente teorema.

**Teorema 3.7.2.** *Sea  $f$  una función real con segunda derivada continua en un intervalo abierto  $(a, b)$ . Entonces  $f$  es convexa si y sólo si su segunda derivada  $f''$  es no negativa en  $(a, b)$ .*

En efecto, supongamos que  $f''$  es no negativa en  $(a, b)$ . Entonces  $f'$  es no decreciente en  $(a, b)$ . Para  $a < x < y < b$ ,  $0 < \lambda < 1$  y  $z = (1 - \lambda)x + \lambda y$ , tenemos

$$\begin{aligned} f(z) - f(x) &= \int_x^z f'(t)dt \leq f'(z)(z - x), \\ f(y) - f(z) &= \int_z^y f'(t)dt \geq f'(z)(y - z). \end{aligned}$$

Como  $z - x = \lambda(y - x)$  y  $y - z = (1 - \lambda)(y - x)$ , tenemos

$$\begin{aligned} f(z) &\leq f(x) + \lambda f'(z)(y-x), \\ f(z) &\leq f(y) - (1-\lambda)f'(z)(y-x). \end{aligned}$$

Multiplicando las dos desigualdades por  $(1-\lambda)$  y  $\lambda$  respectivamente y sumandolas, obtenemos

$$(1-\lambda)f(z) + \lambda f(z) \leq (1-\lambda)f(x) + \lambda f(y).$$

El lado izquierdo es simplemente  $f(z) = f((1-\lambda)x + \lambda y)$ , por tanto se tiene la convexidad de  $f$  en  $(a, b)$ .

Si no fuese cierto que  $f''$  es no negativa en  $(a, b)$ , entonces  $f''$  ser a negativa en un subintervalo  $(a', b')$  por continuidad. Por argumentos similares a los ya dados, en  $(a', b')$  tendríamos que

$$\begin{aligned} f(z) - f(x) &> f'(z)(z-x), \\ f(y) - f(z) &< f'(z)(y-z), \end{aligned}$$

y en consecuencia

$$f((1-\lambda)x + \lambda y) > (1-\lambda)f(x) + \lambda f(y).$$

De esta forma  $f$  no ser a convexa en  $(a, b)$ .

El siguiente teorema es una versión multidimensional del teorema 3.7.2

**Teorema 3.7.3.** *Sea  $f$  una función real con diferencial segunda continua en un intervalo abierto y convexo  $C$  de  $\mathbb{R}^n$ . Entonces  $f$  es convexa en  $C$  si y sólo si su matriz Hessiana*

$$Q_x = (q_{ij}(x)), \quad q_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}(x_1, x_2, \dots, x_n),$$

es positiva definida para toda  $x \in C$ . Es decir, si

$$z^\top Q_x z \geq 0 \quad \text{para todo } z \in \mathbb{R}^n.$$

La convexidad de  $f$  en  $C$  es equivalente a la convexidad de la restricción de  $f$  a cada segmento de línea en  $C$ . Es decir a la convexidad de la función  $g(\lambda) = f(y + \lambda z)$  en el intervalo abierto  $\{\lambda | y + \lambda z \in C\}$  para cada  $y \in C$  y  $z \in \mathbb{R}^n$ . Observemos que

$$g''(\lambda) = z^\top Q_x z, \quad x = y + \lambda z.$$

Entonces por el teorema 3.7.2,  $g$  es convexa para cada  $y \in C$  y  $z \in \mathbb{R}^n$  si y sólo si  $z^\top Q_x z \geq 0$  para todo  $z \in \mathbb{R}^n$  y  $x \in C$ .

**Lema 3.7.4.** *Sea  $B = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$  y  $H = (h_1 \ h_2) \begin{pmatrix} a & b \\ b & c \end{pmatrix} \begin{pmatrix} h_1 \\ h_2 \end{pmatrix}$ . Entonces  $H$  es positiva definida si y sólo si  $a > 0$  y  $\det(B) = ac - b^2 > 0$ .*

Observamos que la matriz

$$H = (h_1 \quad h_2) \begin{pmatrix} a & b \\ b & c \end{pmatrix} \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} = ah_1^2 + 2bh_1h_2 + ch_2^2,$$

la podemos reescribir como

$$H = a\left(h_1 + \frac{b}{a}h_2\right)^2 + \left(c - \frac{b^2}{a}\right)h_2^2$$

al completar cuadrados. Supongamos que  $H$  es positiva de nida. Haciendo  $h_2 = 0$  vemos que  $a > 0$ . Haciendo  $h_1 = -\frac{b}{a}h_2$  obtenemos que  $ac - b^2 > 0$ . Si ahora suponemos que  $a > 0$  y  $ac - b^2 > 0$ , observamos que  $H$  es una suma de cuadrados, de donde  $H > 0$ . Si  $H = 0$ , entonces cada cuadrado debe ser cero, lo que implica que tanto  $h_1$  como  $h_2$  deben ser cero. Así,  $H$  es positiva de nida.

De manera semejante,  $H$  es negativa de nida si y solo si  $a < 0$  y  $ac - b^2 > 0$ . Existen criterios similares para matrices simétricas  $n \times n$ . Considerando las  $n$  submatrices cuadradas que se pueden obtener a lo largo de la diagonal, la función cuadrática asociada a la matriz  $B$  es positiva de nida si y solo si los determinantes de estas submatrices diagonales son todas positivas. Para la matriz negativa de nida debemos alternar los signos comenzado con el signo menos.

**Ejemplo 3.7.1 (2 estados).** En el ejemplo 1.2.1 se encontró el siguiente resultado para dos estados

$$\mathbf{P}(t) = \frac{1}{\lambda_1 + \lambda_2} \begin{pmatrix} \lambda_2 + \lambda_1 e^{\lambda t} & \lambda_1 - \lambda_1 e^{\lambda t} \\ \lambda_2 - \lambda_2 e^{\lambda t} & \lambda_1 + \lambda_2 e^{\lambda t} \end{pmatrix}, \quad (3.12)$$

con  $\lambda = -(\lambda_1 + \lambda_2) \neq 0$ , que se obtuvo a través de la ecuación (1.14). Ahora, usando la ecuación (3.6) obtenemos para  $i, j = 1, 2$ ,

$$\begin{aligned} \mu_{11}(t) &= \frac{\lambda_2 t}{\lambda_1 + \lambda_2} + \frac{\lambda_1}{(\lambda_1 + \lambda_2)^2} [1 - e^{-(\lambda_1 + \lambda_2)t}] \\ \mu_{21}(t) &= \frac{\lambda_2 t}{\lambda_1 + \lambda_2} - \frac{\lambda_2}{(\lambda_1 + \lambda_2)^2} [1 - e^{-(\lambda_1 + \lambda_2)t}] \\ \mu_{12}(t) &= \frac{\lambda_1 t}{\lambda_1 + \lambda_2} - \frac{\lambda_1}{(\lambda_1 + \lambda_2)^2} [1 - e^{-(\lambda_1 + \lambda_2)t}] \\ \mu_{22}(t) &= \frac{\lambda_1 t}{\lambda_1 + \lambda_2} + \frac{\lambda_2}{(\lambda_1 + \lambda_2)^2} [1 - e^{-(\lambda_1 + \lambda_2)t}] \end{aligned}$$

y, por la ecuación (3.7)

$$\lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t p_{11}(\tau) d\tau = \lim_{t \rightarrow \infty} \frac{1}{t} \mu_{11}(t) = \lim_{t \rightarrow \infty} \frac{1}{t} \mu_{21}(t) = \frac{\lambda_2}{\lambda_1 + \lambda_2} = \pi(1) \quad (3.13)$$

y,

$$\lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t p_{12}(\tau) d\tau = \lim_{t \rightarrow \infty} \frac{1}{t} \mu_{12}(t) = \lim_{t \rightarrow \infty} \frac{1}{t} \mu_{22}(t) = \frac{\lambda_1}{\lambda_1 + \lambda_2} = \pi(2). \quad (3.14)$$

Queremos aproximar los resultados obtenidos en [12] para  $(\lambda_1, \lambda_2, t)$ . Como el proceso comienza en el estado uno, usamos  $\mu_{11}(t)$ . En el primer caso  $(\lambda_1, \lambda_2, t) = (1, 1, 1)$  tenemos que

$$\mu_{11}(1) = 0.7162$$

y, aplicando la fórmula de Black-Scholes ponderada (ecuación (3.3)) cuando  $K = 110$ ,  $\bar{r}_1^* = 0.0746$ ,  $(\bar{\sigma}^*)_1^2 = 0.1677^2$  llegamos a que el precio de la call en dos estados es 5.7574. La gráfica 3.1 muestra que la derivada de  $\Theta$  es negativa en el intervalo  $(0, 0.8207)$ , por tanto  $p(g|t_1, t_2)$  es cóncava; y positiva en  $(0.8207, 1)$ , donde es convexa. El tiempo 0.7162 se encuentra en el primer intervalo, luego entonces la desigualdad de Jensen nos proporciona una cota superior para el valor verdadero del derivado. Encontramos que la diferencia con [12] (5.75) no es significativa. Presentamos las tablas de los cálculos realizados.

$(\lambda_1, \lambda_2, t)$	$(1, 1, 1)$	$(1, 2, 1)$	$(2, 1, 1)$
$\mu_{11}(t)$	0.7162	0.7722	0.5445
$K$	Call	Call	Call
110	5.7574	5.3385	7.0383
100	10.6620	10.1806	12.1021
90	17.5764	17.1135	18.9629

$(\lambda_1, \lambda_2, t)$	$(1, 1, 10)$	$(1, 2, 10)$	$(2, 1, 10)$
$\mu_{11}(t)$	5.2500	6.7778	3.5556
$K$	Call	Call	Call
110	57.4869	51.4551	63.3543
100	61.0169	55.3949	66.4556
90	64.6655	59.5040	69.6379

Observamos lo siguiente

1. Los valores encontrados para la call en dos ambientes difieren de manera no significativa respecto de lo obtenido en [12].
2. Como el proceso comienza en el estado uno sólo podemos hacer uso de  $\mu_{11}(t)$ .
3. En [12] la función de densidad (1.16) es incompleta; como se observa en la gráfica 1.2 para valores pequeños de  $t$ , por lo que el valor de la call debe ser ajustado por la probabilidad de permanecer en el estado uno,  $e^{-\lambda_1 t}$ . En nuestro caso el cálculo es directo sin importar el valor de  $t$ .
4. El precio de la call en nuestra aproximación siempre es ligeramente mayor a lo encontrado en [12]. Esto es claro pues,  $p(g|t_1, t_2)$  en todos los demás casos es cóncava.

Presentamos, ahora, el valor de la call para dos ambientes comenzando en el estado dos.

$(\lambda_1, \lambda_2, t)$	$(1,1,1)$	$(1,2,1)$	$(2,1,1)$
$\mu_{22}(t)$	0.7162	0.5445	0.7722
$K$	Call	Call	Call
110	8.9663	7.6991	9.3772
100	14.2024	12.8297	14.6420
90	20.9854	19.6638	21.4085

$(\lambda_1, \lambda_2, t)$	$(1,1,10)$	$(1,2,10)$	$(2,1,10)$
$\mu_{22}(t)$	5.2500	3.5556	6.7778
$K$	Call	Call	Call
110	59.3044	52.8350	64.4146
100	62.7045	56.6841	67.4358
90	66.2104	60.6897	70.5321

**Ejemplo 3.7.2 (3 estados).** Supondremos los mismos datos que se usaron para el cálculo de la call en dos estados, y agregaremos los valores de la tasa y la volatilidad para el tercer estado;  $r_3 = 10\%$ ,  $\sigma_3 = 10\%$ , pero con matriz de intensidad

$$\Lambda = \begin{pmatrix} -2 & 2 & 0 \\ 2 & -(1+2) & 1 \\ 0 & 1 & -1 \end{pmatrix},$$

es decir,  $\beta_0 = 2$ ,  $\delta_1 = 2$ ,  $\beta_1 = 1$  y  $\delta_2 = 1$ . Con esta matriz de intensidad se obtiene el siguiente semigrupo de transición  $\{\mathbf{P}(t)\}_{t \geq 0}$

$$\begin{aligned} p_{11}(t) &= -2 \frac{1 + e^{(-3+\sqrt{3})t} + e^{-(3+\sqrt{3})t}}{(3+\sqrt{3})(-3+\sqrt{3})} \\ p_{12}(t) &= \frac{\sqrt{3}e^{-(3+\sqrt{3})t} + e^{-(3+\sqrt{3})t} - \sqrt{3}e^{(-3+\sqrt{3})t} + e^{(-3+\sqrt{3})t} - 2}{(3+\sqrt{3})(-3+\sqrt{3})} \\ p_{13}(t) &= -\frac{\sqrt{3}e^{-(3+\sqrt{3})t} - e^{-(3+\sqrt{3})t} - \sqrt{3}e^{(-3+\sqrt{3})t} - e^{(-3+\sqrt{3})t} + 2}{(3+\sqrt{3})(-3+\sqrt{3})} \\ p_{21}(t) &= \frac{\sqrt{3}e^{-(3+\sqrt{3})t} + e^{-(3+\sqrt{3})t} - \sqrt{3}e^{(-3+\sqrt{3})t} + e^{(-3+\sqrt{3})t} - 2}{(3+\sqrt{3})(-3+\sqrt{3})} \\ p_{22}(t) &= -\frac{2 + \sqrt{3}e^{-(3+\sqrt{3})t} + 2e^{-(3+\sqrt{3})t} - \sqrt{3}e^{(-3+\sqrt{3})t} + 2e^{(-3+\sqrt{3})t}}{(3+\sqrt{3})(-3+\sqrt{3})} \\ p_{23}(t) &= \frac{e^{-(3+\sqrt{3})t} + e^{(-3+\sqrt{3})t} - 2}{(3+\sqrt{3})(-3+\sqrt{3})} \\ p_{31}(t) &= -\frac{\sqrt{3}e^{-(3+\sqrt{3})t} - e^{-(3+\sqrt{3})t} - \sqrt{3}e^{(-3+\sqrt{3})t} - e^{(-3+\sqrt{3})t} + 2}{(3+\sqrt{3})(-3+\sqrt{3})} \\ p_{32}(t) &= \frac{e^{-(3+\sqrt{3})t} + e^{(-3+\sqrt{3})t} - 2}{(3+\sqrt{3})(-3+\sqrt{3})} \\ p_{33}(t) &= \frac{-2 + \sqrt{3}e^{-(3+\sqrt{3})t} - 2e^{-(3+\sqrt{3})t} - \sqrt{3}e^{(-3+\sqrt{3})t} - 2e^{(-3+\sqrt{3})t}}{(3+\sqrt{3})(-3+\sqrt{3})} \end{aligned}$$

Integrando  $\mathbf{P}(t)$  para el tiempo  $t = 2$  calculamos la media de los tiempos de ocupación comenzando en el estado uno. Así:  $\mu_{11}(2) = 0.9792$ ,  $\mu_{12}(2) = 0.6591$ ,  $\mu_{13}(2) = 0.3618$ .

Aplicando la fórmula de Black-Scholes ponderada (ecuación (3.3)) cuando  $K = 110$ ,  $T = 2$ ,  $\bar{r}_1^* = 0.0872$ ,  $(\bar{\sigma}^*)_1^2 = 0.1662^2$  llegamos a que el precio de la call en tres estados es 13.3016. Hemos calculado numéricamente las regiones de concavidad y convexidad para  $p(g|t_1, t_2, t_3)$ , es decir aquella que cumplen con las condiciones del lema 3.7.4, las regiones son vacías. Por lo que sólo podemos afirmar que el valor encontrado de 13.3016 es una buena aproximación al valor verdadero del derivado. Lo mismo sucede para los valores de  $K = 100, 90$ . La siguiente tabla resume los cálculos realizados.

Comenzando en el edo. uno	
$\mu_{11}(2)$	0.9792
$\mu_{12}(2)$	0.6591
$\mu_{13}(2)$	0.3618
$K$	Call
110	13.3016
100	18.8551
90	25.5657

Comenzando en el edo. dos	
$\mu_{21}(2)$	0.6591
$\mu_{22}(2)$	0.8305
$\mu_{23}(2)$	0.5104
$K$	Call
110	14.8070
100	20.4496
90	27.1430

Comenzando en el edo. tres	
$\mu_{31}(2)$	0.3618
$\mu_{32}(2)$	0.5104
$\mu_{33}(2)$	1.1278
$K$	Call
110	13.5625
100	19.5460
90	26.6220

Hemos observado que si los valores en el estado uno y dos son iguales,  $p(g|t_1, t_2, t_3)$  puede tener regiones tanto de concavidad y convexidad como regiones que no lo son. De aquí que sea importante realizar el análisis de convexidad en los puntos de interés.

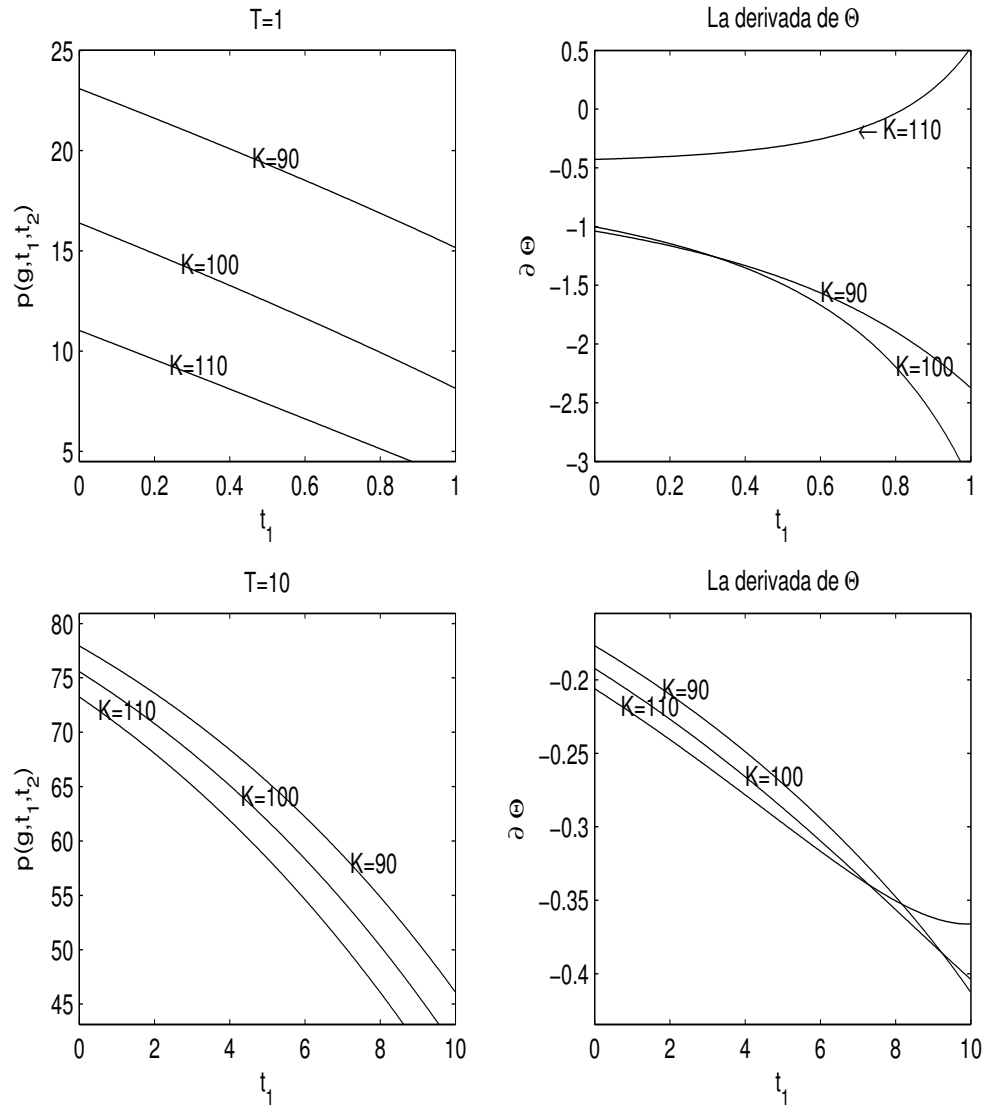


Figura 3.1: Valores de la *call* en dos estados calculadas con  $p(g|t_1, t_2)$  junto con la derivada de  $\Theta$  del ejemplo 3.7.1 para diferentes valores de  $K$ .



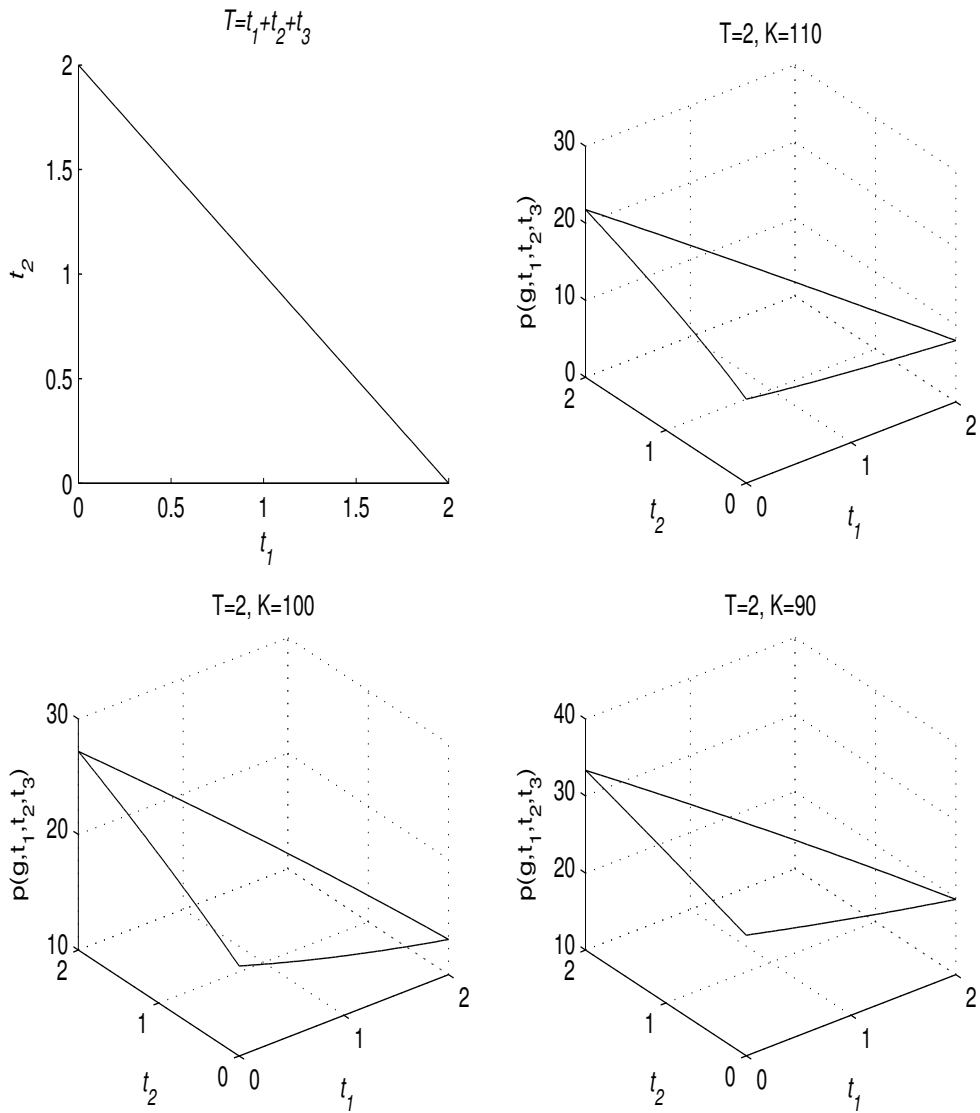


Figura 3.2: Valor de la *call* en tres estados calculadas con  $p(g|t_1, t_2, t_3)$  del ejemplo 3.7.2 para diferentes valores de  $K$ .

### 3.8. Correlación temporal en el sistema

Todos los sistemas macroscopicos estan sujetos a perturbaciones irregulares que no pueden ser modelados de manera determinística; llamamos a tales perturbaciones ruido. El ruido ha tenido un papel cambiante en la historia de la ciencia. Hasta el siglo XIX el ruido se consideraba una molestia que deb a ser evitada o eliminada; y es a principios del siglo XX cuando fue posible obtener informacion acerca de los sistemas físicos debido a estas fluctuaciones.

Supondremos que la evolucion de un sistema dinamico puede ser modelada por la siguiente ecuacion de Langevin:

$$\dot{x} = f(x) + g(x)\eta(t), \quad (3.15)$$

en donde el  $\{\cdot\}$  denota la derivada respecto del tiempo,  $\eta(t)$  el termino de ruido coloreado (correlacion temporal finita) con distribucion no Gaussiana; teniendo-se, por tanto, una ecuacion diferencial estocastica no Markoviana. Con mayor precision,  $\eta(t)$  es un proceso Markoviano generado por la siguiente ecuacion

$$\dot{\eta} = h(\eta) + \varepsilon\xi(t). \quad (3.16)$$

En donde los momentos del ruido blanco  $\xi(t)$  (ver ejemplo 3.8.1) son

$$\mathbb{E}[\xi(t)] = 0$$

$$\mathbb{E}[\xi(t)\xi(s)] = 2\delta(t-s).$$

Es decir, se tiene el siguiente sistema;

$$\begin{cases} \dot{x} = f(x) + g(x)\eta(t) \\ \dot{\eta} = h(\eta) + \varepsilon\xi(t), \quad \mathbb{E}[\xi(t)\xi(s)] = 2\delta(t-s). \end{cases} \quad (3.17)$$

Donde  $x$  denota el estado del sistema al tiempo  $t$ ,  $\eta$  es la cantidad que fluctua y describe el efecto del medio sobre el sistema; es el acoplamiento del sistema y el medio. La funcion  $f$  es la derivada del potencial  $U(x)$  (ver figura 3.3 (a)),  $f(x) = -U'(x)$ .

Observamos que el planteamiento del sistema (3.17) es bastante general, por lo que estudiaremos un caso particular:

$$\begin{cases} \dot{x} = f(x) + g(x)\eta(t) \\ \dot{\eta} = -\frac{1}{2}V'_q(\eta) + \frac{\sqrt{D}}{D}\xi(t), \quad \mathbb{E}[\xi(t)\xi(s)] = 2\delta(t-s) \end{cases} \quad (3.18)$$

donde

$$V_q(\eta) = \frac{D}{(q-1)} \ln \left[ 1 + \frac{(q-1)\eta^2}{2D} \right]. \quad (3.19)$$

Este planteamiento particular del ruido  $\eta$  nos permite facilmente alejarnos del comportamiento gaussiano variando el parametro  $q$ , ver [7], [8], [9] y [10] para ampliar estos temas.  $D$  y  $\varepsilon$  son parametros de ruido relacionados a la intensidad y tiempo de correlacion respectivamente.

**Ejemplo 3.8.1 (Eliminacion Adiabatica).** Si una partícula pequeña de masa  $m$  es sumergida en un fluido, una fuerza de fricción actuará en la partícula. La expresión más simple para tal fuerza de fricción la da la ley de Stokes

$$F_c = -\alpha v.$$

Entonces, la ecuación de movimiento para la partícula en ausencia de otras fuerzas adicionales es

$$m\dot{v} + \alpha v = 0 \quad (3.20)$$

ó,

$$\dot{v} + \gamma v = 0; \quad \gamma = \alpha/m = 1/\tau.$$

Así, la velocidad  $v(t)$  decrece a cero con tiempo de relajación  $\tau = 1/\gamma$  según

$$v(t) = v(0)e^{t/\tau} = v(0)e^{-\gamma t};$$

esto es, llamamos a  $\tau$  el tiempo de relajación pues determina el tiempo  $t$  a partir del cual  $v(t) \rightarrow 0$ .

La ecuación (3.20) es válida solamente si la masa de la partícula es grande a fin de que la velocidad debida a las fluctuaciones termal sea despreciable. Una masa más pequeña  $m$  hará que la velocidad termal sea observable y por tanto la velocidad de esta partícula ya no puede ser descrita de manera correcta por la ecuación (3.20). La modificación de (3.20) que lleva a una ecuación termal adecuada es la adición de una fuerza  $F_f(t)$  que fluctúa

$$F(t) = F_c(t) + F_f(t) = -\alpha v(t) + F_f(t). \quad (3.21)$$

La fuerza  $F_f(t)$  es una fuerza estocástica. De no ser así, deberíamos ser capaces de resolver el sistema de ecuaciones del movimiento para todas las moléculas del fluido y de la partícula. Debido a que el número de moléculas en el fluido es grande (del orden de  $10^{23}$ ) no podemos resolver este sistema de ecuaciones. Más aún, como no conocemos los valores iniciales de todas las moléculas del fluido, no podemos calcular el movimiento exacto de la partícula inmersa en el fluido.

Insertando (3.21) en (3.20) y dividiendo entre la masa obtenemos las ecuaciones del movimiento

$$\dot{x} = v \quad (3.22a)$$

$$\dot{v} + \gamma v = \xi(t) \quad (3.22b)$$

En donde definimos la fuerza de fluctuación por unidad de masa

$$\xi(t) = F_f(t)/m,$$

que es llamada fuerza de Langevin. Algunas propiedades de la fuerza de Langevin son:

1.  $\mathbb{E}[\xi(t)] = 0$ ; el efecto promedio de las colisiones es nulo.
2.  $\mathbb{E}[\xi(t)\xi(s)] = q\delta(t-s)$ ; para diferentes tiempos las  $\xi$ 's son independientes. La  $\delta$  de Dirac aparece pues de otra manera la energía promedio de la partícula no sería finita, violando el principio de equipartición de energía que afirma que la partícula entra en equilibrio termodinámico cuando  $t \rightarrow \infty$ . El valor de  $q$  es calculado de esta manera, encontrándose que  $q = 2\gamma kT/m$ ; con temperatura  $T$  y constante de Boltzmann  $k$ .

3. Con ayuda de la identidad  $\int \varphi(s)\delta(s-t)ds = \varphi(t)$  dada en [15] vemos que la densidad espectral  $S(\omega)$  de la fuerza de Langevin  $\xi(t)$  es la transformada de Fourier de la función de correlación y que es independiente de  $\omega$ :

$$S(\omega) = 2 \int_{-\infty}^{\infty} e^{-i\omega\tau} q\delta(\tau)d\tau = 2q.$$

Es por esta razón que la fuerza  $\xi$  es llamada ruido blanco y si no tiene una correlación dependiente de  $\delta$ , i.e., si la densidad espectral depende de la frecuencia  $\omega$ , entonces se usa el término ruido coloreado.

4. Queremos resolver (3.22b) para la condición inicial  $v_0$  al tiempo 0. La solución es como en el ejemplo 2.2.2,

$$v(t) = v_0 e^{-\gamma t} + \int_0^t e^{-\gamma(t-s)} \xi(s) ds. \quad (3.23)$$

Consideremos la situación en que el coeficiente de fricción  $\gamma$  es grande, entonces para  $t \gg \tau$  la ecuación (3.23) se relajará a

$$v(t) \rightarrow \int_0^t e^{-\gamma(t-s)} \xi(s) ds. \quad (3.24)$$

De esta forma podemos escribir a la ecuación (3.22a) como

$$\begin{aligned} \dot{x} &= \int_{s=0}^t e^{-\gamma(t-s)} \xi(s) ds \\ &= \frac{2}{\gamma} \int_{s=0}^t \frac{1}{2\tau} e^{-\gamma(t-s)} \xi(s) ds \\ &= \frac{1}{\gamma} \int_{s=-t}^t \frac{1}{2\tau} e^{-|t-s|/\tau} \xi(s) ds, \end{aligned} \quad (3.25)$$

y si  $\tau \rightarrow 0$

$$\begin{aligned} \dot{x} &= \frac{1}{\gamma} \int_{s=-t}^t \delta(t-s) \xi(s) ds \\ &= \frac{1}{\gamma} \xi(t). \end{aligned}$$

Así<sup>2</sup>, hemos eliminado a la variable  $\dot{v}$  en (3.22b). Físicamente significa que valores grandes de la fricción  $\gamma$  (o, en general, tiempos pequeños de relajación), obligan a  $\dot{v}(t)$  a relajarse. Este procedimiento es el prototipo de todas las eliminaciones adiabáticas.<sup>3</sup>

Ahora, la función de correlación para la velocidad es

$$\begin{aligned} \mathbb{E}[v(t_1)v(t_2)] &= v_0 e^{-\gamma(t_1+t_2)} + \int_0^{t_1} \int_0^{t_2} e^{-\gamma(t_1+t_2-s_1-s_2)} q\delta(s_1-s_2) ds_1 ds_2 \\ &= v_0 e^{-\gamma(t_1+t_2)} + \frac{q}{2\gamma} (e^{-\gamma|t_1-t_2|} - e^{-\gamma(t_1+t_2)}), \end{aligned}$$

<sup>2</sup>La densidad de la exponencial doble,  $f(x) = \frac{1}{2\sigma} e^{-|x-\mu|/\sigma}$  con  $x \in \mathbb{R}$ , tiene media  $\mu$  y varianza  $2\sigma^2$ .

<sup>3</sup>Un proceso, que en termodinámica, pasa sin perder o ganar calor.

usando que  $\mathbb{E}[\xi(t)] = 0$  y la propiedad de  $\delta(\cdot)$  utilizada recientemente al calcular la densidad espectral. Siendo  $\gamma > 0$  si  $t_1, t_2 \rightarrow \infty$  con diferencia constante  $|t_1 - t_2|$ , la función de correlación se vuelve estacionaria e independiente de la velocidad inicial  $v_0$

$$\mathbb{E}[v(t_1)v(t_2)] = \frac{q}{2\gamma} e^{-\gamma|t_1 - t_2|}. \quad (3.26)$$

La ecuación diferencial estocástica (ecuación de Langevin) de una partícula de masa  $m$  que se mueve bajo la influencia de una fuerza externa  $F = mU'(x)$ , donde  $mU(x)$  se le conoce como el potencial; y con constante de fricción  $\gamma$  está dada por

$$\begin{aligned} \dot{x} &= v \\ \dot{v} &= -\gamma v + F(x)/m + q^{1/2}\xi(t), \quad \mathbb{E}[\xi(t)\xi(s)] = q\delta(t-s), \end{aligned} \quad (3.27)$$

La función de densidad  $p_t(x, v)$  satisface la ecuación de Fokker-Planck, ver ecuación (2.16), mejor conocida como ecuación de Klein-Kramers; que es una ecuación del movimiento para la densidad de la posición y la velocidad que describe el movimiento Browniano de una partícula en un campo externo. En el caso unidimensional la ecuación de Fokker-Planck correspondiente es

$$\frac{\partial p}{\partial t} = -\frac{\partial}{\partial x}(vp) + \frac{\partial}{\partial v} \left( \gamma vp - \frac{F(x)}{m} p \right) + \frac{\gamma kT}{m} \frac{\partial^2 p}{\partial v^2}.$$

La ecuación (3.27) se puede escribir como la ecuación de movimiento

$$m\ddot{x} + m\gamma\dot{x} = F(x) + mq^{1/2}\xi(t).$$

Si consideramos valores grandes de la fricción  $\gamma$  podemos relajar a  $\ddot{x} = \dot{v}$  (eliminación adiabática). Entonces, suponemos que

$$\dot{x} = \frac{F(x)}{m\gamma} + q^{1/2}\xi(t)/\gamma$$

con la siguiente ecuación de Fokker-Planck correspondiente

$$\frac{\partial p}{\partial t} = -\frac{1}{m\gamma} \frac{\partial}{\partial x}(F(x)p) + \frac{kT}{m\gamma} \frac{\partial^2 p}{\partial x^2}.$$

Esta ecuación se conoce como ecuación de Smoluchowski.

### 3.8.1. Aproximación unificada de ruido coloreado

La solución al problema no Markoviano inicialmente planteado (ecuación (3.15)) se ha transformado en un problema Markoviano extendiendo el número de variables y ecuaciones (ecuaciones (3.17)), de esta manera podemos resolver la ecuación asociada de Fokker-Planck para obtener la función de densidad. Sin embargo, el planteamiento original del problema está en función de una ecuación diferencial estocástica no Markoviana, por lo que podemos usar el procedimiento de eliminación *adiabática* (ver ejemplo 3.8.1) que permite reducir el problema extendido a una aproximación Markoviana efectiva en términos de la variable original, ver [17]. Lo anterior conduce a una ecuación aproximada de Fokker-Planck en una variable.

Claramente, cuando  $q \rightarrow 1$  en la ecuacion (3.19), se tiene que

$$\lim_{q \rightarrow 1} V_q(\eta) = \lim_{q \rightarrow 1} \ln \left[ 1 + \beta(q-1) \frac{\eta^2}{2} \right]^{\frac{1}{\beta(q-1)}} = \frac{\eta^2}{2},$$

donde  $\beta = 1/D$ , y en este caso se obtiene el sistema

$$\begin{cases} \dot{x} = f(x) + g(x)\eta(t) \\ \dot{\eta} = -\frac{1}{\tau}\eta + \frac{\sqrt{D}}{\tau}\xi(t), \quad \mathbb{E}[\xi(t)\xi(s)] = 2\delta(t-s) \end{cases} \quad (3.28)$$

en el que restringiremos el analisis. La segunda ecuacion del sistema (3.28) tiene la siguiente funcion de correlacion calculada de manera similar que en (3.26) del ejemplo 3.8.1.

$$\mathbb{E}[\eta(t)\eta(s)] = \frac{D}{\tau} e^{-|t-s|/\tau} \rightarrow 2D\delta(t-s), \quad \text{cuando } \tau \rightarrow 0. \quad (3.29)$$

Sea ahora  $B_t$  un movimiento browniano estandar. Calculemos la siguiente funcion de correlacion

$$\begin{aligned} \mathbb{E}[\sqrt{2D}\dot{B}_t\sqrt{2D}\dot{B}_s] &= 2D \frac{\partial^2}{\partial t \partial s} \text{m n}(t,s) = 2D \frac{\partial}{\partial t} \mathcal{H}(t-s) \\ &= 2D\delta(t-s), \end{aligned} \quad (3.30)$$

donde  $\mathcal{H}$  es la funcion Heaviside o funcion escalonada, considerada como funcion generalizada. Luego entonces  $\eta(t) = \sqrt{2D}\dot{B}_t$ . De esta forma el sistema (3.28) se reduce a la siguiente ecuacion de difusion

$$dx = f(x)dt + g(x)\sqrt{2D}dB_t. \quad (3.31)$$

El sistema (3.28) tiene una estructura de ruido multiplicativo, la cual se puede transformar en ruido aditivo con la transformacion  $y = \int^x g^{-1}(s)ds$ : hacemos  $x = u(y)$ , entonces  $\dot{y} = \dot{x}g^{-1}(u(y)) = \dot{x}\tilde{g}^{-1}(y)$ , y (3.28) queda

$$\dot{y}\tilde{g}(y) = f(u(y)) + \tilde{g}(y)\eta(t)$$

$$\begin{cases} \dot{y} = \frac{\tilde{f}(y)}{\tilde{g}(y)} + \eta(t) = \tilde{h}(y) + \eta(t), \\ \dot{\eta} = -\frac{1}{\tau}\eta + \frac{\sqrt{D}}{\tau}\xi(t), \quad \mathbb{E}[\xi(t)\xi(t')] = 2\delta(t-t'). \end{cases}$$

Tomando la segunda derivada temporal de  $y$  y sustituyendo los valores de  $\dot{\eta}(t)$  y  $\eta(t)$  obtenemos

$$\ddot{y} + \dot{y}[-\frac{1}{\tau} - \tilde{h}'(y)] - \frac{\tilde{h}(y)}{\tilde{g}(y)} = \frac{D^{1/2}}{\tau}\xi(t). \quad (3.32)$$

Con el cambio de escala temporal  $t = \sqrt{s}$ , la ecuacion (3.32) se convierte en

$$\begin{aligned} s : \quad & \frac{1}{\sqrt{s}} \frac{1}{\sqrt{s}} \ddot{y} + \frac{\dot{y}}{\sqrt{s}} \left[ -\frac{1}{2} - \tilde{h}'(y) \right] - \frac{\tilde{h}(y)}{\sqrt{s}} = \frac{D^{1/2}}{\sqrt{s}} \xi(s), \\ s : \quad & \ddot{y} + \dot{y} \left[ -\frac{1}{2} - \frac{1}{2} \tilde{h}'(y) \right] - \tilde{h}(y) = \frac{D^{1/2}}{\sqrt{s}} \xi(s), \\ s : \quad & \ddot{y} + \tilde{\gamma}(y, \dot{y}) \dot{y} - \tilde{h}(y) = \frac{D^{1/2}}{\sqrt{s}} \xi(s), \end{aligned}$$

con  $\tilde{\gamma}(y, \dot{y}) = -\frac{1}{2} - \frac{1}{2} \tilde{h}'(y)$ , pues

$$\begin{aligned} \mathbb{E}[D^{1/2} \xi(t) D^{1/2} \xi(t')] &= D \mathbb{E}[\xi(t) \xi(t')] = D 2\delta(t - t') = D 2\delta(\sqrt{s} - \sqrt{s'}) \\ &= \frac{D}{\sqrt{s}} 2\delta(s - s') = \frac{D}{\sqrt{s}} \mathbb{E}[\xi(s) \xi(s')] \\ &= \mathbb{E} \left[ \sqrt{\frac{D}{s}} \xi(s) \sqrt{\frac{D}{s'}} \xi(s') \right], \end{aligned}$$

usando propiedades de  $\delta$  (ver [15]) y de la esperanza. Entonces la ecuacion (3.32) en la nueva escala temporal es

$$s : \quad \ddot{y} + \tilde{\gamma}(y, \dot{y}) \dot{y} - \tilde{h}(y) = \frac{D^{1/2}}{\sqrt{s}} \xi_s, \quad \mathbb{E}[\xi(s) \xi(s')] = 2\delta(s - s'), \quad (3.33)$$

con  $\tilde{\gamma}(y, \dot{y})$  como coeficiente de friccion. Si el coeficiente de friccion es grande, podemos aplicar la cancelacion adiabatica del ejemplo 3.8.1. Con  $\dot{y} = 0$  obtenemos una aproximacion markoviana efectiva del sistema (3.28):

$$s : \quad \dot{y} \left[ -\frac{1}{2} - \frac{1}{2} \tilde{h}'(y) \right] - \tilde{h}(y) = \frac{D^{1/2}}{\sqrt{s}} \xi(s), \quad (3.34)$$

y como

$$\begin{aligned} \tilde{h}'(y) &= \frac{\tilde{g}(y) f'(u(y)) u'(y) - \tilde{f}(y) g'(u(y)) u'(y)}{\tilde{g}^2(y)} \\ &= \tilde{f}'(y) - \tilde{f}(y) \frac{\tilde{g}'(y)}{\tilde{g}(y)}, \text{ pues } \dot{x} = u'(y) \dot{y} = \tilde{g}(y) \dot{y}, \end{aligned}$$

la ecuacion (3.34) es

$$\begin{aligned} s : \quad & \dot{y} \left[ -\frac{1}{2} - \frac{1}{2} \left( \tilde{f}'(y) - \tilde{f}(y) \frac{\tilde{g}'(y)}{\tilde{g}(y)} \right) \right] - \tilde{h}(y) = \frac{D^{1/2}}{\sqrt{s}} \xi(s), \\ s : \quad & \dot{y} \tilde{\gamma}_1(y, \dot{y}) - \tilde{h}(y) = \frac{D^{1/2}}{\sqrt{s}} \xi(s). \end{aligned}$$

que en el tiempo original  $t = \sqrt{s}$ , y con ruido multiplicativo  $g(x)\xi(t)$  se lee

$$\begin{aligned} \frac{\dot{x} \sqrt{s}}{g(x)} \left[ -\frac{1}{2} - \frac{1}{2} \left( f'(x) - f(x) \frac{g'(x)}{g(x)} \right) \right] - \frac{f(x)}{g(x)} &= D^{1/2} \xi(t), \\ \frac{\dot{x}}{g(x)} \left[ 1 - \left( f'(x) - f(x) \frac{g'(x)}{g(x)} \right) \right] - \frac{f(x)}{g(x)} &= D^{1/2} \xi(t), \end{aligned}$$

con  $\gamma_1(x, \cdot) = 1 - [f'(x) - f(x)g'(x)/g(x)]$ ; entonces

$$\dot{x} = \frac{f(x)}{\gamma_1(x, \cdot)} + \frac{g(x)D^{1/2}}{\gamma_1(x, \cdot)}\xi(t), \quad \mathbb{E}[\xi(t)\xi(s)] = 2\delta(t-s). \quad (3.35)$$

La ecuacion (3.35) es el resultado principal del cual facilmente obtenemos la ecuacion de Fokker-Planck asociada. Observemos que cuando  $D \rightarrow 0$ , la ecuacion (3.35) se convierte en

$$\dot{x} = f(x) + g(x)\sqrt{2D}\dot{B}_t, \quad (3.36)$$

que es lo que se esperaba.

### 3.9. Escape de un pozo biestable

Consideremos la ecuacion de difusion (3.36) con  $f(x) = -U'(x)$  y  $\sqrt{b(x)} = g(x)\sqrt{2D}$

$$dx = f(x)dt + \sqrt{b(x)}dB_t. \quad (3.37)$$

Nos preguntamos cuanto tiempo una part cula cuya posicion esta descrita por la ecuacion (3.37) permanece en una cierta region de  $x$ .

Supongamos que la part cula se encuentra inicialmente en  $x$  al tiempo  $t$  y que pertenece al intervalo abierto  $(a, b)$ . Diremos que la part cula deja el sistema cuando alcanza el extremo  $a$  o  $b$  (estos puntos seran las condiciones de frontera de la ecuacion diferencial).

Bajo estas condiciones, la probabilidad de que al tiempo  $t$  la part cula se encuentre en el intervalo  $(a, b)$  es

$$\int_a^b p_t(x, y)dy \doteq G(x, t). \quad (3.38)$$

Sea  $T(x)$  el tiempo en el que la part cula  $x$  deja  $(a, b)$ . Para un proceso con estados absorbentes en la frontera la probabilidad de no salir de  $(a, b)$  al tiempo  $t$  es identica a la probabilidad de encontrarlo en el tiempo  $t$  en un punto del intervalo,

$$\mathbb{P}(T \geq t) = \int_a^b p_t(x, y)dy$$

es decir que  $G(x, t)$  es  $\mathbb{P}(T \geq t)$ .

Ahora, la ecuacion regresiva asociada a esta probabilidad es (ver ejemplo 2.3.2),

$$\frac{\partial p_t(x, y)}{\partial t} = f(x)\frac{\partial p_t(x, y)}{\partial x} + \frac{1}{2}b(x)\frac{\partial^2 p_t(x, y)}{\partial x^2},$$

con condicion inicial

$$p_0(x, y) = \delta(x - y)$$



y por tanto,  $G(x, t)$  obedece la ecuacion

$$\frac{\partial G(x, t)}{\partial t} = f(x) \frac{\partial G(x, t)}{\partial x} + \frac{1}{2} b(x) \frac{\partial G(x, t)}{\partial x^2}. \quad (3.39)$$

Dado que  $\mathbb{P}(T < t) + \mathbb{P}(T \geq t) = 1$  la densidad de  $T(x)$  es  $f_T(t) = -\frac{\partial G(x, t)}{\partial t}$ . Entonces, el  $n$ -esimo momento de  $T(x)$  es

$$\mathbb{E}[T^n(x)] = M_n(x) = - \int_{t=0}^{\infty} t^n \frac{\partial G(x, t)}{\partial t} dt.$$

El primer momento de  $T(x)$  es

$$\begin{aligned} M(x) &= - \int_{t=0}^{\infty} t \frac{\partial G(x, t)}{\partial t} dt = -tG(x, t)|_{t=0}^{\infty} + \int_{t=0}^{\infty} G(x, t) dt \\ &= \int_{t=0}^{\infty} G(x, t) dt, \end{aligned} \quad (3.40)$$

despues de aplicar integracion por partes (si de nimos  $\int_{t=0}^{\infty} p_t(x, y) dt = p(x, y)$  llegamos a que  $M(x) = \int_{t=0}^{\infty} \int_a^b p_t(x, y) dy = \int_a^b p(x, y) dy$ ). Similarmente, encontramos que el  $n$ -esimo momento de  $T(x)$  es

$$M_n(x) = n \int_{t=0}^{\infty} t^{n-1} G(x, t) dt.$$

Podemos, ahora, encontrar una ecuacion diferencial para el primer momento  $M(x)$  usando (3.40) e integrando (3.39) de cero a infinito. Observando que

$$\int_{t=0}^{\infty} \frac{\partial G(x, t)}{\partial t} dt = G(x, \infty) - G(x, 0) = -1,$$

llegamos a que

$$-1 = f(x)M'(x) + \frac{1}{2}b(x)M''(x), \quad (3.41)$$

con condiciones de frontera que dependeran de que los extremos del intervalo  $(a, b)$  sean absorbentes o reflejantes.

De igual forma (multiplicando por  $nt^{n-1}$  la ecuacion (3.39) e integrando de cero a infinito), vemos que la ecuacion diferencial para  $M_n(x)$  es

$$-nM_{n-1}(x) = f(x)M'_n(x) + \frac{1}{2}b(x)M''_n(x). \quad (3.42)$$

### 3.9.1. El tiempo de escape es exponencial

Sea  $U(x)$  un potencial biestable (ver figura 3.3 (a)), con  $u_{\min(-)}$  y  $u_{\min(+)}$  sus minimos y  $u_{\max}$  el maximo local.

Supongamos que  $V(x)$  es otro potencial con los mismos puntos extremos de  $U(x)$  y que  $b'(u_{\max}) = 0$ . En el intervalo  $(a, b)$  digamos que  $a$  es un punto

reflejante ( $a \approx -\infty$ ) y que  $b$  es absorbente ( $b = u_{\text{máx}}$ ; la partícula desaparece al alcanzar este punto). Las condiciones de frontera son entonces

$$M(u_{\text{máx}}) = 0 \quad (3.43a)$$

$$M'(-\infty) = 0. \quad (3.43b)$$

Multiplicaremos la ecuación (3.42) por  $\exp(\int_{z=u_{\text{máx}}}^x \frac{f(z)}{0,5b(z)} dz) = e^{-V(x)}$ . Observamos que  $f(u_{\text{máx}}) = 0$  pues  $f(x) = -U'(x)$ .

$$M_n''(x)e^{-V(x)} + \frac{f(x)}{0,5b(x)}M_n'(x)e^{-V(x)} = -\frac{nM_{n-1}(x)}{0,5b(x)}e^{-V(x)},$$

agrupando

$$\left(M_n'(x)e^{-V(x)}\right)' = -\frac{nM_{n-1}(x)}{0,5b(x)}e^{-V(x)},$$

integrando y observando la condición de frontera  $M'(-\infty) = 0$ , escribimos

$$M_n'(x) = e^{V(x)} \int_{y_n=-\infty}^x -\frac{nM_{n-1}(y_n)}{0,5b(y_n)} e^{-V(y_n)} dy_n.$$

Así

$$M_n(x) = n \int_{x_n=x}^{u_{\text{máx}}} e^{V(x_n)} \int_{y_n=-\infty}^{x_n} M_{n-1}(y_n) \frac{e^{-V(y_n)}}{0,5b(y_n)} dy_n dx_n. \quad (3.44)$$

Lo que significa que podemos escribir los momentos  $M_n$  de forma recursiva:

$$\begin{aligned} M_{n-1}(x) &= (n-1) \int_{x_{n-1}=x}^{u_{\text{máx}}} e^{V(x_{n-1})} \int_{y_{n-1}=-\infty}^{x_{n-1}} M_{n-2}(y_{n-1}) \\ &\quad \frac{e^{-V(y_{n-1})}}{0,5b(y_{n-1})} dy_{n-1} dx_{n-1} \\ &\quad \vdots \\ M(x) &= \int_{x_1=x}^{u_{\text{máx}}} e^{V(x_1)} \int_{y_1=-\infty}^{x_1} \frac{e^{-V(y_1)}}{0,5b(y_1)} dy_1 dx_1, \end{aligned} \quad (3.45)$$

La función  $H(x) = \frac{e^{-V(x)}}{0,5b(x)}$  en la ecuación (3.45) tiene un mínimo local en  $u_{\text{máx}}$  ( $H'(u_{\text{máx}}) = 0$ ) con máximos locales próximos a  $u_{\text{mín}(-)}$  y  $u_{\text{mín}(+)}$  (si no es que iguales) y cuando  $x \rightarrow \pm\infty$ ,  $H(x)$  tiende a cero (como en la gráfica 3.3 (b)). De esta forma  $\int_{y=-\infty}^x H(y) dy$  hace su máxima contribución cerca de  $x = u_{\text{máx}}$  (método de Laplace<sup>4</sup>), en cuyo caso  $\int_{y=-\infty}^x H(y) dy$  será casi constante para

<sup>4</sup>Se basa en la idea de que la mayor contribución del valor de una integral proviene de la vecindad de los puntos en el intervalo  $[a, b]$  en donde el integrando tiene su valor máximo.

aquellos valores de  $x$  que hagan de  $\int_{x_1=x}^{u_{\text{máx}}} e^{V(x_1)} dx_1$  signi cativamente diferente de cero. Por tanto, podemos aproximar (3.45) como

$$M(x) = \int_{y_1=-\infty}^{u_{\text{máx}}} \frac{e^{-V(y_1)}}{0,5b(y_1)} dy_1 \int_{x_1=u_{\text{mín}(-)}}^c e^{V(x_1)} dx_1 = \frac{1}{\lambda},$$

con  $u_{\text{máx}} < c$ .

Debemos observar que el tiempo que la part cula pasa en una vecindad de  $u_{\text{máx}}$  es insigni cante con respecto al tiempo que pasa en una vecindad inmediata de  $u_{\text{mín}}$ . Si tambien observamos que la probabilidad de superar la barrera en  $u_{\text{max}}$  en cualquier direccion es un medio, el tiempo de escape efectivo,  $M_{\text{efect}}$ , sera  $\frac{1}{2}M = \frac{1}{2\lambda}$ .

Ahora, escribiendo recursivamente los momentos  $M_n$  llegamos a una sucesion de integrales anidadas

$$M_n(x) = n! \int_{x_n=x}^{u_{\text{máx}}} e^{V(x_n)} \int_{y_n=-\infty}^{x_n} \int_{x_{n-1}=y_n}^{u_{\text{máx}}} e^{V(x_{n-1})} \int_{y_{n-1}=-\infty}^{x_{n-1}} \dots \int_{x_1=y}^{u_{\text{máx}}} e^{V(x_1)} \int_{y_1=-\infty}^{x_1} \frac{e^{-V(y_1)}}{0,5b(y_1)} \dots \frac{e^{-V(y_n)}}{0,5b(y_n)} dy_1 dx_1 \dots dy_n dx_n,$$

y usando el mismo argumento que usamos para separar la integral en (3.45) obtenemos

$$\mathbb{E}[T^n(x)] = M_n(x) = n!(1/2\lambda)^n, \quad (3.46)$$

con lo que concluimos que  $T(x)$  tiene una distribucion exponencial<sup>5</sup>. El mismo resultado se encuentra para el pozo derecho.

Por consiguiente, una part cula que se encuentre en cualesquiera de los dos pozos escapara con una intensidad  $\lambda_1$  de uno de ellos y con otra intensidad  $\lambda_2$  del otro, fluctuando entre pozo y pozo. Consecuentemente observamos una similitud en los tiempos de permanencia entre pozos y estados.

<sup>5</sup>La función generatriz de momentos de la densidad exponencial,  $f(x) = \lambda e^{-\lambda x}$ , es

$$\mathcal{M}_x(t) = \frac{\lambda}{\lambda - t}.$$

Entonces

$$\begin{aligned} \mathcal{M}'_x(t)|_{t=0} &= \frac{\lambda}{\lambda^2} \\ \mathcal{M}''_x(t)|_{t=0} &= \frac{2\lambda}{\lambda^3} \\ &\vdots \\ \mathcal{M}_x^{(r)}(t)|_{t=0} &= \frac{r!\lambda}{\lambda^{r+1}}. \end{aligned}$$

**Ejemplo 3.9.1 (Correlacion en el derivado).** Sea  $U$  el siguiente potencial biestable

$$U(x) = \begin{cases} -\frac{a_1 x^2}{2} + \frac{b_1 x^4}{4}, & \text{si } x \geq 0 \\ -\frac{a_2 x^2}{2} + \frac{b_2 x^4}{4}, & \text{en otro caso.} \end{cases}$$

Esta forma del potencial asimétrico nos permite variar la curvatura del potencial modificando solamente sus coeficientes. Su máximo local está en  $u_{\text{máx}} = 0$  (que consideraremos un punto absorbente) y en  $u_{\text{mín}} = \pm\sqrt{a/b}$  se encuentran sus dos mínimos. La altura entre el punto máximo y el punto mínimo es  $E = a^2/4b$ . Sea  $g(x) = 1$ , consecuentemente  $b(x) = 2D$ . Digamos que  $D = 1/2$ . La ecuación de difusión (3.37) es entonces

$$dx = -U'(x)dt + dB_t, \quad (3.47)$$

con ecuación diferencial de tiempos de ocupación asociada

$$-1 = -U'(x)M'(x) + \frac{1}{2}M''(x). \quad (3.48)$$

Calculemos  $M(\sqrt{a/b})$  cuando  $b = a^2$  para el potencial  $U(x)$ . De esta forma tenemos que  $E = 1/4$ . MATLAB resuelve de manera sencilla la ecuación (3.48) para los datos proporcionados a través de su función `bvp4c`.

El primer paso consiste en reescribir la ecuación diferencial (3.41) como un sistema de dos ecuaciones diferenciales ordinarias. Con  $y_1(x) = M(x)$  y  $y_2(x) = M'(x)$ , obtenemos el siguiente sistema de ecuaciones diferenciales

$$\begin{cases} y_1'(x) = y_2(x) \\ y_2'(x) = -\frac{2}{b(x)} - \frac{2f(x)}{b(x)}y_2(x) \end{cases} \quad (3.49)$$

Este sistema de ecuaciones diferenciales ordinarias lo codificamos en MATLAB en el archivo `mfpt_ode.m`.

```
function dy = mfpt_ode(x,y,D,tao)
% Sistema de ecuaciones diferenciales para la media de los
% tiempos de ocupacion.
% Nota. Aunque las constantes D y tao no se usen aqui, la necesitara
% bvp4c para cambiar el valor de estas constantes.

b=2*D;
aa=5.04;      % las constantes del potencial
bb=aa^2;

f=aa*x-bb*x^3; % EL POTENCIAL ESTANDAR SIN RUIDO (tao = 0)

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%CON RUIDO (tao > 0), SE DEBE MODIFICAR f, b Y AGREGAR
%EL VALOR DE gamma

%% gamma=1+tao*(-aa+3*bb*x^2);
```

```

%% b=2*D./gamma^2;
%% f=(aa*x-bb*x^3)./gamma; %El potencial modificado por el ruido
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
dy=[y(2) % el sistema
    -2/b - ((2*f)/b)*y(2)];

```

Las condiciones de frontera (3.43a) y (3.43b) de la ecuación diferencial las escribimos en el archivo `mfpt_bc.m`.

```

function res = mfpt_bc(ya,yb,D,tao)
% Se definen las condiciones de frontera para mfpt_ode.
% Aunque la D y la tao no se usen aqui, la necesitara bvp4c
% para poder variar estas constantes.
res = [ yb(1) % (1) se refiere a las condiciones iniciales en y
        ya(2) ] ; % (2) se refiere a las condiciones iniciales en y'

```

Finalmente, tecleamos el siguiente código para obtener la solución y la gráfica para el punto  $a = 5.04$ .

```

vdpsolinit = bvpinit(linspace(-4,0,4),[0 1]);
% [0,1] es una primera aproximacion
D=0.5; tao=0;
sol=bvp4c(@mfpt_ode,@mfpt_bc,vdpsolinit,[ ],D,tao);
plot(sol.x,sol.y), grid,
a=5.04; b=a^2; x=-sqrt(a/b), M=deval(sol,x)

```

Encontramos para  $a = 5.04$  un valor de  $M(\pm 0.4454) = 0.5003$  y para  $a = 2.52$  un valor de  $M(\pm 0.6299) = 1.0007$  (ver figuras 3.3 (c) y (d)). Por lo que podemos decir que la media de los tiempos de escape de los pozos del potencial  $U(x)$  tiene asociada las intensidades  $\lambda = 1$  ó  $\lambda = 2$  según consideremos sus coeficientes; proponiendo de esta manera una matriz de intensidad para un proceso en dos ambientes cambiantes.

Al existir correlación ( $\tau > 0$ ) vemos que la ecuación diferencial para la media de los tiempos de ocupación debe corregirse con el coeficiente de fricción  $\gamma_1(x, \tau) = 1 - \tau [f'(x) - f(x)g'(x)/g(x)] = 1 + \tau[U''(x)]$  dado en la ecuación (3.35), así la ecuación de difusión (3.47) se lee ahora

$$dx = -\frac{U'(x)}{\gamma_1} dt + \frac{1}{\gamma_1} dB_t,$$

con ecuación diferencial de tiempos de ocupación asociada

$$-1 = -\frac{U'(x)}{\gamma_1} M'(x) + \frac{1}{2\gamma_1^2} M''(x).$$

El mismo código que se usó para calcular la media de los tiempos de ocupación sin correlación se puede utilizar, pero modificando ahora los correspondientes valores de  $f$  y  $b$ .

Notemos que este potencial modificado mantiene su máximo en cero y que cuando  $\tau \rightarrow 0$  recuperamos el potencial original.

Ahora, con los datos del ejemplo 3.7.1 y aplicando la fórmula (3.4) de Black-Scholes con densidad de tiempos de ocupación para  $K$  y  $T$ , obtenemos los siguientes precios para la opción call.

$K = 110, T = 1$			
$\tau$	$\lambda_1$	$\lambda_2$	$c$
$0$	$1$	$1$	$5.7443$
$0.03$	$0.7728$	$0.7728$	$5.4618$
$0.05$	$0.6662$	$0.6662$	$5.3$

$K = 100, T = 1$			
$\tau$	$\lambda_1$	$\lambda_2$	$c$
$0$	$1$	$1$	$10.5721$
$0.03$	$0.7728$	$0.7728$	$10.2422$
$0.05$	$0.6662$	$0.6662$	$10.0556$

$K = 90, T = 1$			
$\tau$	$\lambda_1$	$\lambda_2$	$c$
$0$	$1$	$1$	$17.4921$
$0.03$	$0.7728$	$0.7728$	$17.1765$
$0.05$	$0.6662$	$0.6662$	$16.9985$

$K = 110, T = 1$			
$\tau$	$\lambda_1$	$\lambda_2$	$c$
$0$	$1$	$2$	$5.3303$
$0.03$	$0.7728$	$1.2437$	$5.2609$
$0.05$	$0.6662$	$0.9714$	$5.1732$

$K = 100, T = 1$			
$\tau$	$\lambda_1$	$\lambda_2$	$c$
$0$	$1$	$2$	$10.1079$
$0.03$	$0.7728$	$1.2437$	$10.0188$
$0.05$	$0.6662$	$0.9714$	$9.9150$

$K = 90, T = 1$			
$\tau$	$\lambda_1$	$\lambda_2$	$c$
$0$	$1$	$2$	$17.0483$
$0.03$	$0.7728$	$1.2437$	$17.9632$
$0.05$	$0.6662$	$0.9714$	$16.8621$

$K = 110, T = 1$			
$\tau$	$\lambda_1$	$\lambda_2$	$c$
0	2	1	7.0248
0.03	1.2437	0.7728	6.2520
0.05	0.9714	0.6662	5.8782

$K = 100, T = 1$			
$\tau$	$\lambda_1$	$\lambda_2$	$c$
0	2	1	12.0216
0.03	1.2437	0.7728	11.1418
0.05	0.9714	0.6662	10.7154

$K = 90, T = 1$			
$\tau$	$\lambda_1$	$\lambda_2$	$c$
0	2	1	18.8860
0.03	1.2437	0.7728	18.0415
0.05	0.9714	0.6662	17.6319

$K = 110, T = 10$			
$\tau$	$\lambda_1$	$\lambda_2$	$c$
0	1	1	57.1429
0.03	0.7728	0.7728	56.7681
0.05	0.6662	0.6662	56.5109

$K = 100, T = 10$			
$\tau$	$\lambda_1$	$\lambda_2$	$c$
0	1	1	60.6819
0.03	0.7728	0.7728	60.3305
0.05	0.6662	0.6662	60.0888

$K = 90, T = 10$			
$\tau$	$\lambda_1$	$\lambda_2$	$c$
0	1	1	64.3401
0.03	0.7728	0.7728	64.0207
0.05	0.6662	0.6662	63.7996

$K = 110, T = 10$			
$\tau$	$\lambda_1$	$\lambda_2$	$c$
0	1	2	51.2247
0.03	0.7728	1.2437	52.8393
0.05	0.6662	0.9714	53.4772

$K = 100, T = 10$			
$\tau$	$\lambda_1$	$\lambda_2$	$c$
0	1	2	55.1676
0.03	0.7728	1.2437	56.6704
0.05	0.6662	0.9714	57.2617

$K = 90, T = 10$			
$\tau$	$\lambda_1$	$\lambda_2$	$c$
0	1	2	59.2857
0.03	0.7728	1.2437	60.6657
0.05	0.6662	0.9714	61.2051

$K = 110, T = 10$			
$\tau$	$\lambda_1$	$\lambda_2$	$c$
0	2	1	63.1588
0.03	1.2437	0.7728	61.1178
0.05	0.9714	0.6662	60.0355

$K = 100, T = 10$			
$\tau$	$\lambda_1$	$\lambda_2$	$c$
0	2	1	66.2691
0.03	1.2437	0.7728	64.3754
0.05	0.9714	0.6662	63.3684

$K = 90, T = 10$			
$\tau$	$\lambda_1$	$\lambda_2$	$c$
0	2	1	69.4641
0.03	1.2437	0.7728	67.7309
0.05	0.9714	0.6662	66.8077

*El efecto de la correlación en el precio del derivado es una depreciación y en otros casos una apreciación.*



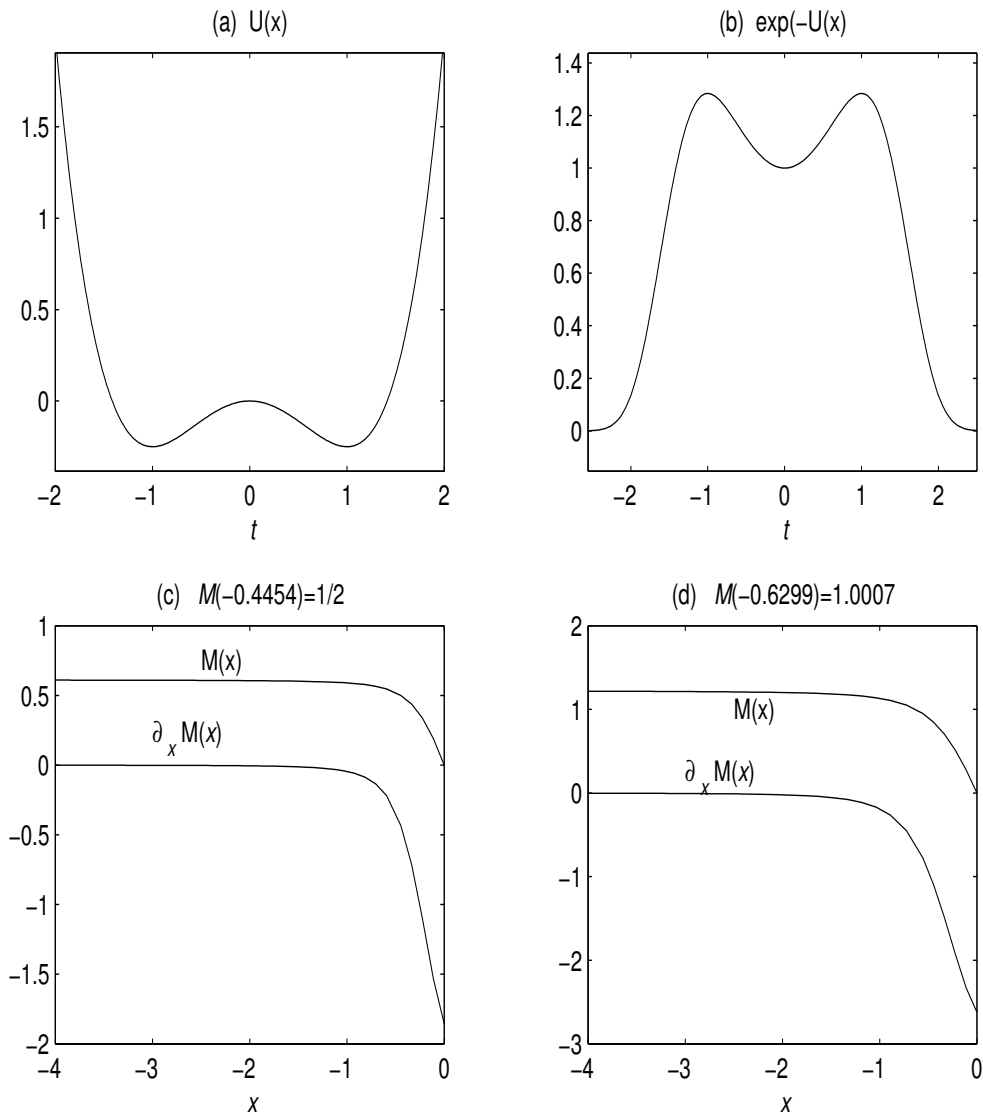


Figura 3.3: Un potencial biestable, (a). (b), la exponencial de un potencial biestable. (c) y (d) gráficas de la media de los tiempos de ocupación del ejemplo 3.9.1 junto con su derivada.

## Conclusion

El modelo modulado de Markov presenta una solución exacta al precio de la opción en dos estados cambiantes. Claramente el valor del derivado es una ponderación del valor de la opción en el primer y segundo estado a través del tiempo de permanencia en cada estado. Las aproximaciones encontradas para dos ambientes son bastante precisas para el tiempo  $T = 1$  por lo que pensamos que el sobreprecio de la opción para tiempos superiores a 1 y mayor número de estados no debe ser considerablemente grande (la concavidad no es muy pronunciada) sirviendo de futura referencia en cálculos que podrán hacerse por simulación.

Cuando existe una pequeña correlación en el modelo ( $\rho > 0$ ), los cambios en la media de los tiempos de ocupación se presentan de manera significativa, modificando consecuentemente el precio de la opción. Dado que la correlación implica que el proceso tiene memoria se espera una subvaluación en el precio del derivado (si se recuerda cuando ocurrió un cambio de estado y se sabe el tiempo promedio para que se vuelva a dar, entonces uno puede cubrirse y por tanto el riesgo se reduce). Sin embargo, con los datos encontrados no podemos afirmar en general que cuando exista correlación temporal el precio del derivado disminuya. Simplemente se modifica su precio. Visto, ahora, como el tiempo de permanencia en los pozos, la media de los tiempos de ocupación en cada uno de ellos se modifica cargándose el tiempo de permanencia hacia uno de ellos.

Los resultados encontrados para el precio de la *call* en dos estados pueden ahora extenderse. La matriz de intensidad se convierte en la matriz de intensidad de un proceso de nacimiento y muerte, en donde el tiempo de permanencia en los estados es el tiempo de permanencia en los pozos y en donde, por consiguiente, solo se permiten transiciones entre pozos adyacentes. El proceso de nacimiento y muerte es bastante general, como procesos más simples encontramos los procesos puros de nacimiento o de muerte y el proceso Poisson.

Cuando la información financiera pueda presentarse en forma de bloques, cada uno de ellos con su respectiva volatilidad y tasa de rendimiento, la aproximación del precio del derivado para  $n$  estados será de gran ayuda en estas circunstancias.

# Apendice A

## A.1. Espacios de Probabilidad

**Definición A.1.1 ( $\sigma$ -Algebra).** Sea  $\Omega$  un conjunto, una  $\sigma$ -algebra  $\mathcal{F}$  sobre  $\Omega$  es una familia  $\mathcal{F}$  de subconjuntos de  $\Omega$  con las siguientes propiedades:

- (a)  $\emptyset \in \mathcal{F}$
- (b)  $F \in \mathcal{F} \Rightarrow F^C \in \mathcal{F}$
- (c)  $A_1, A_2, \dots \in \mathcal{F} \Rightarrow A := \bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$ .

La razón para definir a la  $\sigma$ -algebra es que a menudo es imposible definir probabilidades interesantes en todos los subconjuntos de  $\Omega$ , así que solo se puede avanzar si se definen probabilidades en ciertos subconjuntos de  $\Omega$ .

Una medida de probabilidad  $\mathbb{P}$  en un espacio de medida  $(\Omega, \mathcal{F})$  es una función  $\mathbb{P} : \mathcal{F} \rightarrow [0, 1]$  tal que

- (a)  $\mathbb{P}(\emptyset) = 0, \mathbb{P}(\Omega) = 1$
- (b) Si  $A_1, A_2, \dots \in \mathcal{F}$  y  $\{A_i\}_{i=1}^{\infty}$  son conjuntos disjuntos entonces

$$\mathbb{P}\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mathbb{P}(A_i)$$

A la terna  $(\Omega, \mathcal{F}, \mathbb{P})$  se le conoce como espacio de probabilidad.

**Definición A.1.2 (Filtración).** Una filtración  $\mathbb{F}$  es una colección de  $\sigma$ -algebras  $\{\mathcal{F}_t\}_{t \geq 0}$  tal que

$$0 \leq s < t \Rightarrow \mathcal{F}_s \subset \mathcal{F}_t.$$

La filtración natural del proceso  $\{X_t\}_{t \geq 0}$  es  $\mathcal{F}_t^X \equiv \sigma(X_s : s \leq t)$ . Se usa a la filtración  $\mathbb{F}$  para modelar un flujo de información. Mientras pasa el tiempo un observador sabe más y más aspectos detallados, es decir, particiones más finas de  $\Omega$ . Para el precio de una acción,  $\mathbb{F}$  describe como la información de los precios se muestra ante los inversionistas.

Sea  $\mathcal{B}$  la  $\sigma$ -álgebra de Borel en  $\mathbb{R}$ . Una función  $X : \Omega \rightarrow \mathbb{R}$  es una *variable aleatoria* si  $X$  es una función  $\mathcal{F}$ -medible (i.e.,  $X^{-1}(B) \in \mathcal{F}$  para  $B \in \mathcal{B}$ ). Un proceso estocástico  $\{X_t\}_{t \geq 0}$  es *adaptado* a una filtración  $\mathbb{F}$  si  $X_t$  es  $\mathcal{F}_t$ -medible para cada  $t \geq 0$ . La filtración más pequeña a la cual  $X$  es adaptado es la filtración natural.

**Definición A.1.3 (Martingala).** Un proceso estocástico  $\{M_t\}_{t \geq 0}$  en  $(\Omega, \mathcal{F}, \mathbb{P})$  es una *martingala* con respecto a una filtración  $\{\mathcal{M}_t\}_{t \geq 0}$  (y con respecto a  $\mathbb{P}$  también) si

- a)  $M_t$  es  $\mathcal{M}_t$ -medible para toda  $t$ ,
- b)  $\mathbb{E}[|M_t|] < \infty$  para toda  $t$ , y
- c)  $\mathbb{E}[M_t | \mathcal{M}_s] = M_s$  para toda  $t \geq s$ .

**Definición A.1.4 (Tiempo de paro).** Sea  $\{\mathcal{F}_t\}$  una familia creciente de  $\sigma$ -álgebras. Una función  $\tau : \Omega \rightarrow [0, \infty]$  se llama tiempo de paro con respecto a  $\{\mathcal{F}_t\}$  si

$$\{\omega; (\tau \leq t)\} \in \mathcal{F}_t,$$

para todo  $t \geq 0$ .

Es decir, es posible decidir si  $\tau \leq t$  ha ocurrido o no con base en  $\{\mathcal{F}_t\}$ .

# Apendice B

## B.1. Espacios de Banach y de Hilbert

**Definición B.1.1 (Espacio normado).** Sea  $X$  un espacio lineal sobre  $\mathbb{K}$ . Entonces,  $X$  es llamado un *espacio normado* sobre  $\mathbb{K}$  si y solo si existe una norma  $\|\cdot\|$  sobre  $X$ , i.e., para todo  $u, v \in X$  y  $\alpha \in \mathbb{K}$ , se cumple lo siguiente:

- (a)  $\|u\| \geq 0$  (i.e.,  $\|u\|$  es un numero real no negativo).
- (b)  $\|u\| = 0$  si y solo si  $u = 0$ .
- (c)  $\|\alpha u\| = |\alpha| \|u\|$ .
- (d)  $\|u + v\| \leq \|u\| + \|v\|$  (desigualdad del triangulo).

**Definición B.1.2 (Sucesion de Cauchy).** La sucesion  $u_n$  en el espacio normado  $X$  se llama una *sucesión de Cauchy* si y solo si , para todo  $\epsilon > 0$ , existe un numero  $n_0(\epsilon)$  tal que

$$\|u_n - u_m\| < \epsilon \quad \text{para todo } n, m \geq n_0(\epsilon).$$

**Definición B.1.3 (Espacio de Banach).** Se llama al espacio normado  $X$  un *espacio de Banach* si y solo si cada sucesion de Cauchy es convergente.

Los espacios de Banach tambien se conocen como *espacios normados completos*.

**Definición B.1.4 (Espacio Dual).** Sea  $X$  un espacio normado sobre  $\mathbb{K}$ . Se entiende por una *funcional continua y lineal* a un operador continuo y lineal

$$f : X \rightarrow \mathbb{K}.$$

Al conjunto de todas las funcionales continuas y lineales sobre  $X$  se le conoce como *espacio dual*  $X^*$  de  $X$ ; y se le denota  $X^* = L(X, \mathbb{K})$ .

**Definición B.1.5 (Forma bilineal).** Una forma bilineal  $b(\cdot, \cdot)$ , en un espacio lineal  $X$  es un mapeo  $b : X \times X \rightarrow \mathbb{K}$  tal que cada uno de los mapeos  $v \mapsto b(v, w)$  y  $w \mapsto b(v, w)$  es una forma lineal en  $X$ .

$$b(\alpha u + \beta w, v) = \alpha b(u, v) + \beta b(w, v), \quad b(u, \alpha v + \beta w) = \alpha b(u, v) + \beta b(u, w)$$

con  $u, v, w \in X$  y  $\alpha, \beta \in \mathbb{K}$ . Es simétrica si  $b(v, w) = b(w, v)$  para todo  $v, w \in X$ . Un producto interior, denotado por  $(\cdot, \cdot)$ , es una forma bilineal simétrica en un espacio lineal  $X$  que satisface

- (a)  $(u, u) \geq 0$  y  $(u, u) = 0$  si y solo si  $u = 0$ ;
- (b)  $(u, \alpha v + \beta w) = \alpha(u, v) + \beta(u, w)$ ;
- (c)  $\overline{(u, v)} = (v, u)$

con  $u, v, w \in X$  y  $\alpha, \beta \in \mathbb{K}$ .

**Definición B.1.6 (Espacio de Hilbert).** Un espacio lineal  $X$  sobre  $\mathbb{K}$  es un *espacio de Hilbert* si y solo si se cumple lo siguiente:

- (a) existe un producto interior en  $X$ , y
- (b) cada sucesión de Cauchy con respecto a la norma

$$\|u\| := (u, u)^{1/2}$$

es convergente.

Sea  $\Omega \subset \mathbb{R}^n$  un dominio abierto y acotado,  $L_p(\Omega)$ ,  $1 \leq p \leq \infty$  denota al espacio de funciones  $u$  de nidas en  $\Omega$  cuyos valores absolutos tienen potencia  $p$  y que son Lebesgue integrables en  $\Omega$ . La norma en  $L_p(\Omega)$  esta dada por

$$\|u\|_{L_p(\Omega)} = \left( \int_{\Omega} |u|^p d\Omega \right)^{1/p}.$$

Por tanto  $L_p(\Omega) = \{u : \|u\|_{L_p(\Omega)} < \infty\}$ .

El espacio de Lebesgue  $L_p(\Omega)$  es un espacio de Banach,  $L_1(\Omega)$  no es un espacio de Hilbert;  $L_2(\Omega)$  es un espacio de Hilbert, pues  $L_2(\Omega)$  es el espacio de funciones cuadrado integrables de nidas en  $\Omega$  y dotadas con producto interno

$$(u, v) = \int_{\Omega} uv d\Omega \quad u, v \in L_2(\Omega)$$

y norma

$$\|u\|_0 = (u, u)^{1/2} \quad u \in L_2(\Omega).$$

Ahora, para cualquier entero no negativo  $k$ , definimos el espacio de **Sobolev** como:

$$H^k(\Omega) = \{u \in L_2(\Omega) : D^\alpha u \in L_2(\Omega) \text{ para } |\alpha| \leq k\}$$

donde usamos la notación de índices múltiples,

$$D^\alpha u = \frac{\partial^{|\alpha|} u}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} \dots \partial x_n^{\alpha_n}},$$

$$|\alpha| = |(\alpha_1, \dots, \alpha_n)| = \alpha_1 + \alpha_2 + \dots + \alpha_n.$$

Así,  $H^k(\Omega)$  consta de funciones cuya derivada hasta el orden  $k$  son cuadrado integrable.  $H^k(\Omega)$  está equipado con la norma

$$\|u\|_k = \left( \|u\|_0^2 + \sum_{|\alpha| \leq k} \|D^{|\alpha|}u\|_0^2 \right)^{1/2}$$

Observamos el hecho de que  $H^k(\Omega)$  es también un espacio de Hilbert. Claramente  $H^2(\Omega) = \{u \in L^2(\Omega) : \frac{\partial u}{\partial x_i} \in L^2(\Omega), \frac{\partial^2 u}{\partial x_i \partial x_j} \in L^2(\Omega) \text{ para } i, j = 1, \dots, n\}$ ,  $H^1(\Omega) = \{u \in L^2(\Omega) : \frac{\partial u}{\partial x_i} \in L^2(\Omega) \text{ para } i = 1, \dots, n\}$ , y  $H^0(\Omega) = L_2(\Omega)$ .

**Definición B.1.7 (Operador Adjunto).** Sea  $A : D(A) \subset X \rightarrow X$  un operador lineal, donde el dominio de  $A$ ,  $D(A)$ , es denso en el espacio de Hilbert  $X$  sobre  $\mathbb{K}$ . Definimos,

$$v \in D(A^*)$$

si y solo si existe un elemento  $w \in X$  tal que

$$(Au, v) = (u, w) \quad \text{para todo } u \in D(A);$$

y escribimos  $A^*v := w$ . De esta manera se obtiene el *operador adjunto*

$$A^* : D(A) \subseteq X \rightarrow X.$$

Como consecuencia se puede demostrar que el operador adjunto  $A^*$  es lineal y que para cada  $\alpha \in \mathbb{K}$ ,  $(\alpha A)^* = \bar{\alpha}A^*$ .  $A$  es llamado *auto-adjunto* si y solo si  $A = A^*$ .





# Apendice C

## C.1. Fórmulas de Green

Sea  $G \subset \mathbb{R}^n$  y  $\partial G$  su frontera.

**Definición C.1.1.** Si  $\partial G$  es  $C^1(G)$  y  $u \in C^1(G)$ , entonces a lo largo de  $\partial G$  se define la derivada normal de  $u$  como

$$\frac{\partial u}{\partial n} \doteq \mathbf{n} \cdot \nabla u$$

con  $\mathbf{n}$  un vector unitario normal.

**Teorema C.1.1 (Teorema de Gauss-Green).** Suponga que  $u \in C^1(G)$ , entonces

$$\int_G \nabla u \, dV = \int_{\partial G} u \cdot \mathbf{n} \, dS$$

es decir,

$$\int_G u_{x_i} \, dV = \int_{\partial G} u n_i \, dS \quad (i = 1, \dots, n).$$

El término  $\nabla u \, dV$  se interpreta como la tasa neta a la cual un fluido sale de un elemento de  $G$  donde  $dV$  es el volumen del elemento, el cual es fijo en el espacio. Si se conserva la masa, entonces esto también debe ser la tasa a la cual el fluido pasa a través de la frontera  $\partial G$ . Si  $dS$  es un elemento de área de esta frontera con dirección normal unitaria  $\mathbf{n}$ , la tasa del fluido a través de  $dS$  es  $u \cdot \mathbf{n} \, dS$ . La importancia del teorema es que transforma una integral de volumen en una integral de superficie y *vice versa*. Este resultado puede presentarse de diferentes formas:

**Teorema C.1.2 (Fórmula de integración por partes).** Sean  $u, v \in C^1(G)$ , entonces

$$\int_G u_{x_i} v \, dV = - \int_G u v_{x_i} \, dV + \int_{\partial G} u v n_i \, dS \quad (i = 1, \dots, n).$$

El teorema se verifica aplicando el teorema C.1.1 a  $uv$ .

**Teorema C.1.3 (Formulas de Green).** Sean  $u, v \in C^2(G)$ , entonces

a.  $\int_G \Delta u v = \int_{\partial G} \frac{\partial u}{\partial n} dS,$

b. (Primera identidad de Green)

$$\int_G \nabla u \cdot \nabla v dV = - \int_G u \Delta v dV + \int_{\partial G} u \frac{\partial v}{\partial n} dS,$$

c. (Segunda identidad de Green)

$$\int_G u \Delta v - v \Delta u dV = \int_{\partial G} u \frac{\partial v}{\partial n} - v \frac{\partial u}{\partial n} dS.$$

Usando la formula de integracion por partes con  $u_{x_i}$  en lugar de  $u$  y  $v = 1$ , vemos que

$$\int_G u_{x_i x_i} dV = \int_{\partial G} u_{x_i} n_i dS.$$

Sumando  $i = 1, \dots, n$  se obtiene (a). Para verificar (b), usamos C.1.2 con  $v = v_{x_i}$ . Escribir (b) con  $u$  y  $v$  intercambiados; restar para obtener (c).

**Ejemplo C.1.1.** Calculemos el adjunto del operador

$$Af(y) = \frac{1}{2} \sum_{i,j} a_{ij}(y) \frac{\partial^2 f}{\partial y_i \partial y_j} + \sum_i b_i(y) \frac{\partial f}{\partial y_i}; \quad f \in C_0^2$$

con  $\langle \cdot, \cdot \rangle$  el producto interior en  $L^2(dy)$ ;

$$\langle A_y f, p_t \rangle = \langle f, A_y^* p_t \rangle.$$

Así

$$\int_{\mathbb{R}^n} A_y f(y) p_t(x, y) dy = \int_{\mathbb{R}^n} f(y) A_y^* p_t(x, y) dy,$$

entonces

$$\int_{\mathbb{R}^n} A_y f(y) p_t(x, y) dy = \int_{\mathbb{R}^n} \left( \frac{1}{2} \sum_{i,j} a_{ij}(y) \frac{\partial^2 f(y)}{\partial y_i \partial y_j} + \sum_i b_i(y) \frac{\partial f(y)}{\partial y_i} \right) p_t(x, y) dy.$$

Ahora, aplicando integración por partes en el lado derecho de la igualdad: dos veces en el primer sumando, una sola vez en el segundo y observando que la integral en la frontera de los términos que involucran a  $p_t(x, y)$  es cero; obtenemos

$$\begin{aligned} & \int_{\mathbb{R}^n} A_y f(y) p_t(x, y) dy \\ &= \int_{\mathbb{R}^n} f(y) \left[ \frac{1}{2} \sum_{i,j} \frac{\partial^2}{\partial y_i \partial y_j} (a_{ij}(y) p_t(x, y)) - \sum_i \frac{\partial}{\partial y_i} (b_i(y) p_t(x, y)) \right] dy \\ &= \int_{\mathbb{R}^n} f(y) A_y^* p_t(x, y) dy. \end{aligned}$$

# Bibliografía

- [1] Mendez R os Enrique. *Métodos Analíticos y Numéricos en Matemáticas Financieras*. Tesis de Licenciatura. UNAM, 2000.
- [2] Mart nez Huerta Adolfo. *Valuación y análisis de opciones sobre acciones mediante una fórmula puntual y la construcción de escenarios*. Tesis de Doctorado, proxima a publicarse. Universidad la Salle, 2005.
- [3] Bachelier, L. *Théorie de la Speculation*, Annales de l'Ecole Normale Supérieure Series 3 17, 21-86, 1900. English translation in [4].
- [4] P.H. Cootner, ed. *The Random Character of Stock Market Prices*, MIT Press, Cambridge, MA, 1964.
- [5] Çinlar, Erhan. *Introduction to Stochastic Processes*. Englewood Cliffs, Prentice Hall, 1975.
- [6] Bremaud, Pierre. *Markov chains: Gibbs Fields, Monte Carlo Simulation and Queues*. New York, Springer, 1999.
- [7] M. A. Fuentes, Horacio S. Wio, and Raul Toral. *Effective Markovian approximation for NonGaussian Noises: a path integral approach*. Physica A, 303, 91-104, 2002.
- [8] H. S. Wio, P. Colet, L. Pesquera, M. A. Rodr guez and M. San Miguel. *Path-integral formulation for stochastic processes driven by colored noise*. Phys. Rev. A 40, 7312, 1989.
- [9] F. Castro, H. S. Wio and G. Abramson. *Colored-noise problem: A Markovian interpolation procedure*. Phys. Rev. E 52, 159, 1995.
- [10] F. Castro, A. Sanchez and H. S. Wio. *Reentrance Phenomena in Noise Induced Transitions*. Phys. Rev. Lett. 75, 1691, 1995.
- [11] A.T. Bharucha-Reid, *Elements of the Theory of Markov Processes and Their Applications*. McGraw-Hill, 1960.
- [12] Mogens Bladt and Pablo Padilla Longoria. *Nonlinear Financial Models: Finite Markov Modulation and its Limits. Quantitative Analysis in Financial Markets*. Collected papers of the New York University Mathematical

- Finance Seminar, Volume III. Editor Marco Avellaneda. World Scientific, 2001.
- [13] Mogens Bladt, *A Markov Modulated Financial Model*, J. Comm. Stats.: Stoch. Models 14, 225-240, 1998.
- [14] H. Risken. *The Fokker-Planck Equation: Methods of Solutions and Applications. Vol 18*. Springer Series in Synergetics, Berlin, 1984.
- [15] Papoulis Athanasios, *Signal Analysis*. McGraw-Hill, 1985.
- [16] William Feller. *Introducción a la teoría de las probabilidades y sus aplicaciones. Vol 1*. Limusa, 1980.
- [17] Peter Jung and Peter Hänggi, *A unified colored-noise approximation*, Phys. Rev. A 35: 4464, 1987.
- [18] Maxi San Miguel and Raul Toral. *Stochastic Effects in Physical Systems, in Instabilities and Nonequilibrium Structures VI*. Eds. E. Tirapegui, J. Martínez and R. Tiemann. Kluwer Academic Publishers, 2000.
- [19] Bernt Øksendal. *Stochastic Differential Equations. An Introduction with Applications*. Springer, 2000.
- [20] Gardiner, C. W. *Handbook of Stochastic Methods: for Physics, Chemistry and the Natural Science*, 2nd ed. Berlin, Springer-Verlag, 1985.
- [21] L. Pontryagin, A. Andronov and A. Vitt, 1933. *On the statistical treatment of dynamical systems*, as reproduced in *Noise in nonlinear dynamical systems*, Vol. 1. F. Moss and P. V. E. McClintock, eds. Cambridge Univ. Press, 1989.
- [22] Marek Musiela and Marek Rutkowski. *Martingale methods in financial modelling*. Berlin, Springer, 1997.