



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

POSGRADO EN CIENCIAS MATEMÁTICAS

FACULTAD DE CIENCIAS

“SOBRE LA CONSTRUCCIÓN Y ESTIMACIÓN DE
PROCESOS DE DIFUSIÓN ESTACIONARIOS”

QUE PARA OBTENER EL GRADO ACADÉMICO DE

MAESTRA EN CIENCIAS

P R E S E N T A

ADRIANA ROLDÁN RODRÍGUEZ

DIRECTOR DE LA TESIS: DR. RAMSÉS HUMBERTO MENA CHÁVEZ

MÉXICO, D.F.

NOVIEMBRE, 2011



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.

A mis papás, Carmen y Roberto

Agradecimientos

Me alegra mucho agradecer a todas las personas que me apoyaron en el proceso de elaborar ésta tesis, aun cuando con algunas ya no comparto caminos o espacios les agradezco su confianza, amistad, amor, porras, compañía e incluso llamadas de atención durante este proceso.

Ramsés, muchísimas gracias por toda la instrucción, el apoyo y la guía, y sobre todo por aceptar ser mi director de tesis.

A la Dra. Ana Meda, la Dra. Eliane Rodrigues, la Dra. María del Rocio Elizondo y al Dr. Luis Antonio Rincón les agradezco la minuciosa revisión y todas las recomendaciones en beneficio de este trabajo. También le agradezco al Dr. Juan Carlos Martínez el apoyo en la parte final de este proceso.

Gracias papá y mamá por su grandioso amor, confianza y por todo. Besos, los amo.

Beto, eres el mejor hermano. Gracias y besos de tu mejor hermana, te quiero mucho y ¡salud!

A mamá Chata, mamá Rebeca y abuelito Jaime les agradezco su cariño y sus consejos. También a papá Daniel le doy las gracias, pues llevo presentes todas sus enseñanzas dadas con el ejemplo.

A mis tíos, tías, primos y primas les agradezco su cariño.

A todos mis amigos y compañeros les agradezco precisamente su amistad, compañía y presencia en alguna etapa de mi vida. Los llevo en mente y corazón. Especialmente a Chio, Andrés, Yesenia, Iván, Hypatia, Harim, Benja, Laura, Eze, Amanda y Adriana.

Por tu compañía y por tu buena vibra, gracias Kika.

Finalmente, te agradezco Dios por todo lo que hay y hubo en y detrás de este trabajo, por todos los momentos dentro y fuera de la UNAM. Gracias.

Índice general

Introducción	1
1. Procesos estocásticos	3
1.1. Introducción y definiciones	3
1.2. Estacionariedad e incrementos independientes	6
1.3. Procesos de Markov	8
1.4. Movimiento browniano	10
1.5. Conclusiones	14
2. Procesos de difusión	15
2.1. Definiciones	15
2.2. Ejemplos de procesos de difusión	21
2.3. Propiedades generales	24
2.4. Conclusiones	33
3. Procesos de difusión con densidad marginal y función de auto-correlación dadas	35
3.1. Introducción	35
3.2. Construcción de procesos de difusión	36
3.3. Expresiones generales de procesos de difusión	42
3.4. Modelos de difusión explícitos	46
3.5. Conclusiones	63
4. Construcción alternativa de un proceso de difusión	65
4.1. Construcción de difusiones con probabilidades de transición	65
4.2. El proceso de difusión Gamma-Poisson	69
4.3. Resultados	71
4.4. Conclusiones	81
A. Método de Broyden-Fletcher-Goldfarb-Shanno	83
B. Prueba aumentada de Dickey Fuller	85

Introducción

El motivo principal de este trabajo es hacer un análisis de los modelos de difusión con tendencia lineal con una distribución marginal conocida y preestablecida, pero arbitraria. Asimismo también se presenta una construcción de modelos de difusión basándose en probabilidades de transición.

Los *procesos de difusión* resultan ser una buena herramienta para obtener un modelo aproximado de variables continuas en tiempo continuo. Es posible para fenómenos físicos, biológicos, econométricos y sociales encontrar una aproximación por medio de procesos de difusión, tal es el caso de los movimientos moleculares de partículas sujetas a interacción entre ellas, las fluctuaciones de los precios en un mercado perfecto, la variación del crecimiento poblacional o la sustitución de genes en el desarrollo evolutivo, entre otros.

El capítulo 1 de este trabajo es una recopilación de definiciones y de resultados en la teoría de procesos estocásticos necesarios para establecer la terminología y la notación de los procesos de difusión. Por su parte, en el capítulo 2 se establecen las definiciones requeridas para la definición de los procesos de difusión, como es el caso de la media y varianza infinitesimales, y también se incluyen algunos resultados y ejemplos referentes a los mismos. En el capítulo 3 podemos encontrar el motivo primordial de este trabajo, donde se presentan varios resultados interesantes sobre los procesos de difusión que tienen tendencia lineal, una distribución marginal conocida, pero arbitraria. Además, se muestran los cálculos para obtener los coeficientes de difusión explícitamente, que dependen de la distribución. Finalmente, el capítulo 4, también relevante en este trabajo, lleva como objetivo el estudiar la densidad de transición de un proceso de difusión específico por medio del uso de variables latentes, y también aplicar un derivado de esta técnica a datos reales para determinar los parámetros correspondientes al usar ciertas condiciones distribucionales.

Capítulo 1

Procesos estocásticos

La teoría de procesos estocásticos estudia familias infinitas de variables aleatorias. Estas variables aleatorias pueden depender de un parámetro entero o de un parámetro real, por ejemplo el tiempo. Típicamente, un proceso estocástico se considera un estado de cierto sistema que varía aleatoriamente con el tiempo. Si los estados de este sistema son observados en momentos de tiempo discreto entonces los relacionamos con una sucesión aleatoria o con un *proceso estocástico a tiempo discreto*. Sin embargo, cuando los estados se observan en intervalos de tiempo, o se piensa provienen de un fenómeno que se da en tiempo continuo, los relacionamos con *procesos estocásticos a tiempo continuo*. Los procesos de difusión son una clase de los procesos estocásticos a tiempo continuo.

Este capítulo es una compilación de definiciones y resultados tomados de varios libros para establecer las bases en la definición de los procesos de difusión que son motivo del siguiente capítulo, entre ellos están Mikosch [11], Karlin y Taylor [6], Skorokhod [16], Skorokhod [17] y Billingsley [2].

1.1. Introducción y definiciones

Supongamos que existe un espacio de probabilidad común $(\Omega, \mathcal{A}, \mathbf{P})$ donde están definidas las familias de variables aleatorias. Denotemos por \mathfrak{R} a los números reales y por \mathfrak{C} a los números complejos.

Definición 1 *Un proceso estocástico X*

$$X = \{X_t, t \in \mathcal{T}\} = \{X(t, \omega), t \in \mathcal{T}, \omega \in \Omega\}$$

*es una función de dos variables $X : \mathcal{T} \times \Omega \rightarrow \mathfrak{R}$, donde $X_t = X_t(\cdot) = X(t, \cdot)$ es una variable aleatoria para cada $t \in \mathcal{T}$. Para cada $\omega \in \Omega$ llamamos a $X(\cdot, \omega) : \mathcal{T} \rightarrow \mathfrak{R}$ una **realización**, una **curva muestral** o una **trayectoria** del proceso estocástico X .*

A menudo \mathcal{T} es un intervalo, por ejemplo $\mathcal{T} = [a, b]$, $[a, b)$ ó $[a, \infty)$, donde $a < b$. En estos casos llamamos a X un *proceso estocástico en tiempo continuo*.

Ahora, vamos a mencionar las características no aleatorias de un proceso estocástico tales como su distribución, esperanza, etc.

Un proceso estocástico no trivial $X = \{X_t, t \in \mathcal{T}\}$ con conjunto de índices infinito \mathcal{T} es un objeto infinito-dimensional, es decir, la colección de variables aleatorias $X_t, t \in \mathcal{T}$. Una manera alterna de ver un proceso estocástico es acomodarlo de modo que cada variable aleatoria sea la entrada de un vector con n entradas, $n = 1, 2, \dots$

Definición 2 Las *distribuciones finito-dimensionales (fidis)* del proceso estocástico X son las distribuciones de los vectores finito-dimensionales $(X_{t_1}, \dots, X_{t_n})$, para todas las posibles elecciones de tiempos $t_1, \dots, t_n \in \mathcal{T}$ y para cada $n \geq 1$, es decir,

$$\{F_{X_{t_1} X_{t_2} \dots X_{t_n}} : t_n = 1, 2, \dots, \quad y \quad n = 1, 2, \dots\}.$$

Desde el punto de vista distribucional, la característica básica de probabilidad de un proceso estocástico es su familia de distribuciones finito-dimensionales. Algunas propiedades de una familia de distribuciones finito-dimensionales son:

1. $F_{t_1, \dots, t_n}(A_1, \dots, A_{k-1}, \cdot, A_{k+1}, \dots, A_n)$, $1 \leq k \leq n$, es una medida en \mathcal{A} , donde \mathcal{A} es la σ -álgebra de Ω .
2. $F_{t_1, \dots, t_n}(A_1, \dots, A_n) = F_{t_{i_1}, \dots, t_{i_n}}(A_{i_1}, \dots, A_{i_n})$, si i_1, \dots, i_n es una permutación de los números $1, \dots, n$.
3. $F_{t_1, \dots, t_{n-1}, t_n}(A_1, \dots, A_{n-1}, \Omega) = F_{t_1, \dots, t_{n-1}}(A_1, \dots, A_{n-1})$.

Las propiedades (1)-(3) son llamadas las *condiciones de consistencia para las distribuciones finito-dimensionales*, aunque no siempre se satisfacen. Estas condiciones permiten que las fidis determinen la distribución de X . En este caso, haremos referencia a la colección de las fidis como la *distribución del proceso estocástico*.

Entonces, con el siguiente teorema, conocido como el teorema de existencia de Kolmogorov, podemos afirmar la existencia de un proceso estocástico, $X = \{X_t(\omega); t \in \mathcal{T}, \omega \in \Omega\}$, asociado a una familia de distribuciones finito-dimensionales, $F_{t_1, \dots, t_n}(A_1, \dots, A_n)$, que satisfaga las condiciones de consistencia tal que

$$\mathbf{P}[X_{t_1}(\omega) \in A_1, \dots, X_{t_n}(\omega) \in A_n] = F_{t_1, \dots, t_n}(A_1, \dots, A_n). \quad (1.1)$$

Teorema 1 Supongamos que $\Omega = \mathfrak{R}^m$. Si una familia de funciones $F_{t_1, \dots, t_n}(A_1, \dots, A_n)$ posee las condiciones de consistencia, entonces existe una función aleatoria $X(t, \omega)$ tal que la igualdad (1.1) se cumple para toda n y para toda $t_1, \dots, t_n \in \mathfrak{R}$, y $A_1, \dots, A_n \in \mathfrak{R}^m$.

Demostración Ver el Teorema de consistencia de Kolmogorov en la página 8 de Skorokhod [17]. \square

Para ambos conjuntos de tiempo \mathcal{T} , continuos y discretos, es útil distinguir varias clases de procesos estocásticos de acuerdo con su relación con el tiempo. Primero recordemos que para un vector aleatorio $X = (X_1, \dots, X_n)$ definimos su esperanza como $\mu_X = (\mathbf{E}[X_1], \dots, \mathbf{E}[X_n])$ y su matriz de covarianzas como $\Sigma_X = (\mathbf{Cov}(X_i, X_j))_{i,j=1,\dots,n}$.

Entonces para un proceso estocástico $X = \{X_t, t \in \mathcal{T}\}$ visto como una colección de vectores finito-dimensionales $(X_{t_1}, \dots, X_{t_n})$, $t_1, \dots, t_n \in \mathcal{T}$ y $n \geq 1$, se calculan la esperanza y la matriz de covarianzas para cada vector como se indica en el párrafo anterior.

Definición 3 La *función esperanza* de X está definida para cada X_t por

$$\mu(t) = \mu_{X_t} = \mathbf{E}[X_t], \quad t \in \mathcal{T}.$$

La *función covarianza* de X está definida para cada par X_t, X_s por

$$C(t, s) = \mathbf{Cov}(X_t, X_s) = \mathbf{E}[(X_t - \mu_{X_t})(X_s - \mu_{X_s})], \quad t, s \in \mathcal{T}.$$

La *función varianza* de X está definida para cada X_t por

$$\sigma^2(t) = C(t, t) = \mathbf{Var}(X_t), \quad t \in \mathcal{T}.$$

Estas expresiones dan información sobre la variabilidad en el tiempo de un proceso estocástico. La función esperanza $\mu(t)$ es una cantidad determinista en la que las trayectorias de X se concentran en su alrededor. La función covarianza $C(t, s)$ es una medida determinista de dependencia en el proceso X . La función varianza $\sigma^2(t)$ se puede considerar como una medida determinista de expansión de las trayectorias de X alrededor de $\mu(t)$.

Ejemplo 1 La definición de una densidad gaussiana n -dimensional está dada por

$$f_X(x) = \frac{1}{(2\pi)^n (\det(\Sigma))^{1/2}} \exp \left\{ -\frac{1}{2} (x - \mu) \Sigma^{-1} (x - \mu)' \right\}; \quad \mu, x \in \mathbb{R}^n.$$

Un proceso estocástico es llamado *gausiano* con parámetros μ y Σ si todas sus fidsis son distribuciones gaussianas multivariadas con los parámetros μ y Σ como su esperanza y matriz de covarianzas, respectivamente. La distribución de un proceso estocástico gausiano está determinada sólo por la colección de las esperanzas y matrices de covarianzas de las fidsis.

Un proceso gausiano simple en $\mathcal{T} = [0, 1]$ consiste de las variables aleatorias independientes e idénticamente distribuidas normal con media cero y varianza 1, es decir, $N(0, 1)$. En este caso las fidsis son caracterizadas por las funciones de distribución:

$$\begin{aligned} \mathbf{P}[X_{t_1} \leq x_1, \dots, X_{t_n} \leq x_n] \\ = \mathbf{P}[X_{t_1} \leq x_1] \cdots \mathbf{P}[X_{t_n} \leq x_n] \end{aligned}$$

$$= \Phi(x_1) \cdots \Phi(x_n),$$

para $0 \leq t_1 \leq \cdots \leq t_n \leq 1$, $(x_1, \dots, x_n) \in \mathfrak{R}^n$. Las trayectorias de este proceso son muy irregulares. Este proceso tiene como esperanzas y covarianzas:

$$\mu(t) = 0$$

y

$$C(t, s) = \begin{cases} 1, & \text{si } t = s, \\ 0, & \text{si } t \neq s. \end{cases}$$

1.2. Estacionariedad e incrementos independientes

Los procesos estocásticos se pueden clasificar en *procesos estacionarios* y *procesos no estacionarios*. Los procesos estacionarios son una forma de clasificación en que los instantes en que se estudian los procesos no son importantes.

Definición 4 Decimos que un proceso $X = \{X_t, t \in \mathcal{T}\}$ es **estacionario en el sentido estricto** si sus fids son invariantes bajo desplazamientos de tiempo, es decir, si

$$F_{t_1+h, t_2+h, \dots, t_n+h} = F_{t_1, t_2, \dots, t_n}, \quad (1.2)$$

para todas $t_i, t_i + h \in \mathcal{T}$, donde $h \in \mathfrak{R}$, $i = 1, 2, \dots, n$ y $n = 1, 2, 3, \dots$

Observación 1 Son equivalentes las notaciones:

$$F_{t_1, t_2, \dots, t_j} = F_{X_{i_1}, X_{i_2}, \dots, X_{i_j}}.$$

En particular, X es estacionario en sentido estricto si X_t tiene la misma distribución para toda $t \in \mathcal{T}$. Un ejemplo de un proceso estacionario en el sentido estricto es una sucesión de variables aleatorias independientes e idénticamente distribuidas (i.i.d.).

Definición 5 Si hay una constante μ y una función $c : \mathfrak{R} \rightarrow \mathfrak{R}$ tal que las medias, varianzas y covarianzas de un proceso estocástico satisfacen:

$$\begin{aligned} \mu(t) &= \mu, \\ \sigma^2(t) &= c(0) \text{ y} \\ C(s, t) &= c(|t - s|), \end{aligned} \quad (1.3)$$

para toda $s, t \in \mathcal{T}$, entonces al proceso lo llamamos **estacionario en el sentido amplio** o **estacionario de segundo orden**. Es decir, el proceso es estacionario sólo con respecto a los primeros y segundos momentos.

Ejemplo 2 Un proceso Poisson homogéneo o proceso Poisson con parámetro de intensidad $\lambda > 0$ es un proceso estocástico $X = \{X_t, t > 0\}$ tal que

- i. $X_0 = 0$, inicia en cero casi seguramente,
- ii. $X_t - X_s$ es una variable aleatoria que se distribuye Poisson con parámetro $\lambda(t-s)$ para toda $0 \leq s < t$,
- iii. los incrementos $X_{t_2} - X_{t_1}$ y $X_{t_4} - X_{t_3}$ son independientes y estacionarios, para toda $0 \leq t_1 < t_2 \leq t_3 < t_4$.

Sus medias, varianzas y covarianzas son $\mu(t) = \lambda t$, $\sigma^2(t) = \lambda t$ y $C(s, t) = \lambda \min\{s, t\}$, respectivamente, para toda $s, t > 0$.

Un proceso estacionario en el sentido estricto es un proceso estacionario en el sentido amplio si sus medias, varianzas y covarianzas son todas finitas. Por el contrario, un proceso estacionario en el sentido amplio no necesariamente es estacionario en el sentido estricto.

Si describimos un proceso de la vida real utilizando un proceso estocástico estacionario en el sentido estricto o estacionario en el sentido amplio, entonces estamos suponiendo que las correspondientes características distribucionales de este proceso no cambian en el transcurso del tiempo.

Ejemplo 3 (a) Si un proceso gaussiano X es determinado por sus esperanzas y covarianzas, entonces de la estacionariedad en el sentido estricto tenemos que

$$\mu(t+h) = \mu(t) \quad \text{y} \quad C(t, s) = C(t+h, s+h),$$

para toda $s, t \in \mathcal{T}$ y h tal que $s+h, t+h \in \mathcal{T}$. Esto significa que $\mu(t) = \mu(0)$ y $C(t, s) = c(|t-s|)$. Entonces un proceso gaussiano estacionario en el sentido estricto es también un proceso estacionario en el sentido amplio, y viceversa, debido a que los momentos de primer y segundo orden caracterizan completamente el proceso.

(b) El proceso Ornstein-Uhlenbeck (O-U) $X = \{X_t, t \in \mathbb{R}^+\}$ con parámetro $\gamma > 0$ y valor inicial $X_0 \sim N(0, 1)$ es un proceso gaussiano con medias $\mu(t) = 0$ y covarianzas $C(s, t) = \exp(-\gamma|t-s|)$, para toda $s, t \in \mathbb{R}^+$. Este es un proceso estacionario en el sentido amplio y también es estacionario en el sentido estricto.

Otra forma de clasificación importante de los procesos estocásticos son aquellos con *incrementos independientes*.

Definición 6 Sea $X = \{X_t, t \in \mathcal{T}\}$ un proceso estocástico y sea $\mathcal{T} \subset \mathbb{R}$ un intervalo. Se dice que X tiene **incrementos estacionarios** si

$$F_{X_t - X_s} = F_{X_{t+h} - X_{s+h}},$$

para toda $t, s \in \mathcal{T}$ y h tal que $t+h, s+h \in \mathcal{T}$. Se dice que X tiene **incrementos independientes** si para cada elección de $t_i \in \mathcal{T}$ con $0 = t_1 < \dots < t_n$ y $n \geq 1$,

$$X_{t_1}, X_{t_2} - X_{t_1}, \dots, X_{t_n} - X_{t_{n-1}}$$

son variables aleatorias independientes.

Ejemplo 4 *El movimiento browniano estándar $W = \{W_t, t \geq 0\}$ es un proceso gaussiano con incrementos independientes y*

$$W_0 = 0, \text{ casi seguramente,}$$

- para toda $0 \leq s \leq t$,

$$\mathbf{E}[W_t] = 0 \text{ y } \mathbf{Var}(W_t - W_s) = t - s, \quad (1.4)$$

- tiene trayectorias continuas casi seguramente

Un movimiento browniano estándar W tiene covarianzas iguales a $C(s, t) = \min(s, t)$, entonces no es estacionario en el sentido amplio.

1.3. Procesos de Markov

Las probabilidades condicionales son una buena herramienta en el análisis de las relaciones entre las variables aleatorias de un proceso estocástico en diferentes instantes de tiempo. Por ejemplo, consideremos una sucesión de variables aleatorias X_1, X_2, \dots que toman valores en un conjunto discreto $E = \{x_1, x_2, \dots\}$, $E \subset \mathbb{R}$, que puede ser finito o infinito. Supongamos que estamos en el instante de tiempo n y conocemos el valor actual de X_n que es x_{i_n} , y también conocemos los valores pasados $X_1 = x_{i_1}, X_2 = x_{i_2}, \dots, X_{n-1} = x_{i_{n-1}}$. Entonces deseamos predecir algo sobre los valores futuros al usar esta información, por ejemplo el valor de la variable aleatoria del futuro inmediato, es decir X_{n+1} . Para esto son útiles las probabilidades condicionales:

$$\begin{aligned} \mathbf{P}[X_{n+1} = x_j | X_1 = x_{i_1}, X_2 = x_{i_2}, \dots, X_n = x_{i_n}] \\ = \frac{\mathbf{P}[X_1 = x_{i_1}, X_2 = x_{i_2}, \dots, X_n = x_{i_n}, X_{n+1} = x_j]}{\mathbf{P}[X_1 = x_{i_1}, X_2 = x_{i_2}, \dots, X_n = x_{i_n}]} \end{aligned}$$

para cada $x_j \in E$.

Observación 2 *Por convención utilizamos la siguiente notación:*

$$\{X_1 = x_{i_1}\} = \{\omega \in \Omega : X_1(\omega) = x_{i_1}\}.$$

En principio podríamos determinar tales probabilidades condicionales para todos los estados $x_{i_1}, x_{i_2}, \dots, x_{i_n} \in E$ y para toda $n = 1, 2, \dots$. Sin embargo, en un sistema determinista gobernado por una ecuación diferencial de primer orden sólo el valor actual x_{i_n} de X_n es necesario para determinar el valor futuro X_{n+1} , mientras que los valores pasados de X_1, X_2, \dots, X_{n-1} son involucrados indirectamente para determinar el valor de X_n . Esta es la idea básica de la siguiente propiedad.

Definición 7 *Decimos que X cumple la propiedad de Markov, si*

$$\mathbf{P}[X_{n+1} = x | X_n = x_{i_n}, \dots, X_0 = x_{i_0}] = \mathbf{P}[X_{n+1} = x | X_n = x_{i_n}], \quad (1.5)$$

para todo $x, x_{i_1}, \dots, x_{i_n} \in E$ y $n = 1, 2, 3, \dots$

Una sucesión de variables aleatorias presentando esta propiedad es un ejemplo de un *proceso de Markov*. En particular llamamos a esta sucesión una *cadena de Markov en tiempo discreto*. En términos generales, los *procesos de Markov* se definen de la siguiente manera.

Recordemos que una clase \mathcal{F} de subconjuntos de Ω es una σ -álgebra si se cumple que $\Omega \in \mathcal{F}$; también que si $B \in \mathcal{F}$, entonces $B^c \in \mathcal{F}$; y que es cerrada bajo la formación de uniones numerables, es decir, si $B_1, B_2, \dots \in \mathcal{F}$, entonces $B_1 \cup B_2 \cup \dots \in \mathcal{F}$. Denotaremos por $\sigma(\Omega)$ a la σ -álgebra generada por Ω que además cumple que es una clase de subconjuntos de Ω que es la intersección de todas las σ -álgebras que contienen Ω . La σ -álgebra generada por \mathfrak{R} es la σ -álgebra de Borel, es decir $\sigma(\mathfrak{R}) = \mathcal{B}$.

Definición 8 Sea $\{X_t, t \in \mathcal{T}\}$ un proceso estocástico, donde $\mathcal{T} \in \mathfrak{R}$. X es un **proceso de Markov** si la siguiente condición es satisfecha:

$$\mathbf{P}(f(X_t) \in B | \sigma(X_u, u \leq s)) = \mathbf{P}(f(X_t) \in B | \sigma(X_s)), \quad (1.6)$$

para toda $B \in \mathcal{B}$, para toda f \mathcal{B} -medible y para toda $s, t > 0$, $s < t$.

Los *procesos de Markov* son modelos generales de probabilidad de sistemas dinámicos con perturbaciones que son independientes en tiempos diferentes. Esos sistemas dinámicos se determinan con una función específica para el estado del sistema al momento t bajo la condición de que su estado al momento s es $x \in \mathfrak{R}$.

Definición 9 Definimos las **probabilidades de transición** del proceso de Markov X por

$$P(s, x; t, B) = \mathbf{P}(X_t \in B | X_s = x),$$

para toda $x \in \mathfrak{R}$, $B \in \mathcal{B}$ y $s < t$.

Para s, t y x fijos, $P(s, x; t, \cdot)$ es una función de probabilidad o medida en la σ -álgebra \mathcal{B} de subconjuntos de Borel de \mathfrak{R} . Entonces tenemos una densidad $p(s, x; t, \cdot)$ llamada una *densidad de transición*, tal que

$$P(s, x; t, B) = \int_B p(s, x; t, y) dy$$

para todo $B \in \mathcal{B}$. También definimos $P(s, x; t, B) = I_B(x)$ para $t = s$, donde I_B es la función indicadora del conjunto B . La función $P(s, x; t, B)$ determina la probabilidad de que el proceso que inicia en x al tiempo s pertenezca al conjunto B al momento t .

Definición 10 Decimos que un proceso de Markov en tiempo continuo es **homogéneo** si todas sus densidades de transición $p(s, x; t, y)$ dependen sólo de la diferencia de tiempo $(t - s)$, $s < t$, y no del valor específico de t .

La *homogeneidad* es otro término para referirnos a la estacionariedad de los incrementos.

Ejemplo 5 *Los siguientes procesos son procesos de Markov homogéneos:*

- *Movimiento browniano estándar (1.4) tiene densidad de transición*

$$p(s, x; t, y) = \frac{1}{\sqrt{2\pi(t-s)}} \exp\left(-\frac{(y-x)^2}{2(t-s)}\right). \quad (1.7)$$

- *El proceso O-U con $\gamma = 1$ tiene densidad de transición*

$$p(s, x; t, y) = \frac{1}{\sqrt{2\pi(1-e^{-2(t-s)})}} \exp\left(-\frac{(y-xe^{-(t-s)})^2}{2(1-e^{-2(t-s)})}\right) \quad (1.8)$$

Cuando tenemos un proceso de Markov homogéneo tenemos que $p(s, x; t, y) = p(0, x; t-s, y)$, esta igualdad es equivalente a escribir la siguiente igualdad $p(x_s, y_t) = p(x_0, y_{t-s})$. La correspondiente función de probabilidad de transición de $p(0, x; t-s, y)$ es $P(0, x; t-s, B)$.

Por otro lado, se puede ver que todo proceso de Markov satisface la ecuación de Chapman-Kolmogorov que en términos de densidades se escribe como:

$$p(s, x; t, y) = \int_{-\infty}^{\infty} p(s, x; \tau, z) p(\tau, z; t, y) dz \quad (1.9)$$

para toda $s \leq \tau \leq t$ y $x, y \in \mathfrak{R}$.

Esta ecuación sugiere que para que un proceso que inicia desde un estado x en el tiempo s pertenezca a un conjunto B en el tiempo t es posible que pase después del estado inicial por un estado z en el tiempo τ y en seguida moverse al conjunto B en el tiempo t .

1.4. Movimiento browniano

El movimiento browniano juega un rol central en la teoría de probabilidad, en la teoría de procesos estocásticos, en física, en finanzas, etc.

Al considerar la definición dada en el ejemplo 4 del movimiento browniano se hace la siguiente afirmación.

Lema 1 *Las variables aleatorias $W_t - W_s$ y W_{t-s} , ambas, tienen una distribución $N(0, t-s)$ para $s < t$.*

Demostración Ver sección 1.3 de Mikosch [11]. □

Este lema es resultado de la estacionariedad de los incrementos del proceso, entonces $W_t - W_s$ tiene la misma distribución que $W_{t-s} - W_0 = W_{t-s}$ que es normal con media cero y varianza proporcional a la longitud del intervalo $[s, t]$.

Las descripciones del movimiento browniano y del proceso Poisson, dadas en los ejemplos 4 y 2, respectivamente, coinciden en que son procesos con incrementos estacionarios e independientes. La diferencia crucial entre ellos está en

la distribución de los incrementos. La suposición de que los incrementos se distribuyan Poisson es conveniente cuando las trayectorias son funciones con saltos, mientras que la suposición de que se distribuyan gaussianas es conveniente cuando las trayectorias son continuas.

De la descripción dada del movimiento browniano en el ejemplo 4 vemos que tiene como función esperanza

$$\mu(t) = \mathbf{E}[W_t] = 0, \quad t > 0,$$

y como los incrementos $W_s - W_0 = W_s$ y $W_t - W_s$ son independientes para $s < t$, recordemos que su función covarianza es

$$\begin{aligned} C(t, s) &= \mathbf{E}[(W_t - W_s) + W_s] W_s \\ &= \mathbf{E}[(W_t - W_s)W_s] + \mathbf{E}[W_s^2], \end{aligned}$$

entonces $\mathbf{E}[W_s^2] = \mathbf{Var}(W_s) = \mathbf{Var}(W_0 + (W_s - W_0))$. Por la independencia de los incrementos tenemos que $\mathbf{E}[W_s^2] = \mathbf{Var}(W_0) + \mathbf{Var}(W_s - W_0)$, y tomando la definición del movimiento browniano, que dice que $W_0 = 0$, y por el lema 1 entonces $\mathbf{E}[W_s^2] = s$. Por lo tanto,

$$\begin{aligned} C(t, s) &= \mathbf{E}[(W_t - W_s)] \mathbf{E}[W_s] + s \\ &= s. \end{aligned}$$

Lo anterior resultó de suponer que $s < t$; si suponemos que $s > t$, entonces tenemos que

$$\begin{aligned} C(t, s) &= \mathbf{E}[W_t [(W_s - W_t) + W_t]] \\ &= \mathbf{E}[W_t(W_s - W_t)] + \mathbf{E}[W_t^2] \\ &= \mathbf{E}[W_t] \mathbf{E}[(W_s - W_t)] + t \\ &= t. \end{aligned}$$

Entonces, como un proceso gaussiano es caracterizado por sus esperanzas y varianzas podemos afirmar que el movimiento browniano es un proceso gaussiano con $\mu(t) = 0$ y $C(s, t) = \min(s, t)$.

De la descripción del movimiento browniano dada en el ejemplo 4, sabemos que sus trayectorias son continuas. Sin embargo estas funciones de t son extremadamente irregulares, pues ellas oscilan violentamente. Esto se debe a que los incrementos de W son independientes. En particular, los incrementos del movimiento browniano en intervalos adyacentes son independientes cualquiera que sea la longitud de los intervalos. Entonces $W(\cdot, \omega)$ tiene en intervalos arbitrariamente pequeños pedazos de trayectoria con pendientes arbitrariamente grandes, lo que hace que $W(\cdot, \omega)$ no sea diferenciable en cero.

Teorema 2 *Las trayectorias del movimiento browniano $\{W(\cdot, \omega), \omega \in \Omega\}$ son con probabilidad 1 no-diferenciables.*

Demostración Ver teorema 37.3 de Billingsley [2]. \square

Incluso las irregularidades se pueden tornar más violentas bajo ciertas transformaciones del movimiento browniano.

Del movimiento browniano se pueden derivar varios procesos gaussianos y no gaussianos que son en la práctica relevantes. Introducimos ahora algunos de estos procesos:

(a) Puente browniano

Consideremos el proceso

$$X_t = W_t - tW_1, \quad 0 \leq t \leq 1,$$

donde $X_0 = W_0 - 0W_1 = 0$ y $X_1 = W_1 - 1W_1 = 0$. En estas dos últimas igualdades vemos la razón por la que este proceso X se le llama *Puente browniano*.

Teniendo presentes las propiedades del movimiento browniano sobre sus funciones esperanza y covarianza, tenemos para el puente browniano que $\mathbf{E}[X_t] = \mathbf{E}[W_t] - t\mathbf{E}[W_1] = 0$. Recordemos la definición de la covarianza,

$$\begin{aligned} C(t, s) &= \mathbf{E}[X_t X_s] - \mathbf{E}[X_t] \mathbf{E}[X_s] \\ &= \mathbf{E}[(W_t - tW_1)(W_s - sW_1)] \\ &= \mathbf{E}[W_t W_s - sW_t W_1 - tW_s W_1 + tsW_1^2] \\ &= \text{mín}(t, s) - s \text{mín}(t, 1) - t \text{mín}(1, s) + ts \\ &= \text{mín}(t, s) - ts. \end{aligned}$$

Las funciones esperanza y covarianza del puente browniano son:

$$\mu(t) = 0 \text{ y } C(t, s) = \text{mín}(t, s) - ts, \quad s, t \in [0, 1].$$

(b) Movimiento browniano con tendencia

Consideremos el proceso

$$X_t = \mu t + \sigma W_t, \quad t \geq 0,$$

para las constantes $\sigma > 0$ y $\mu \in \Re$. Teniendo en cuenta las propiedades del movimiento browniano tenemos que $\mathbf{E}[X_t] = \mathbf{E}[\mu t + \sigma W_t] = \mu t$, mientras que para el caso de la covarianza,

$$\begin{aligned} C(t, s) &= \mathbf{E}[X_t X_s] - \mathbf{E}[X_t] \mathbf{E}[X_s] \\ &= \mathbf{E}[(\mu t + \sigma W_t)(\mu s + \sigma W_s)] - \mu t \mu s \\ &= \mu t \mu s + \mu s \sigma \mathbf{E}[W_t] + \mu t \sigma \mathbf{E}[W_s] + \sigma^2 \mathbf{E}[W_t W_s] - \mu t \mu s \\ &= \sigma^2 \text{mín}(s, t). \end{aligned}$$

Este es un proceso gaussiano con esperanza y covarianzas:

$$\mu(t) = \mu t \text{ y } C(t, s) = \sigma^2 \min(t, s), \quad s, t \geq 0.$$

La función esperanza $\mu(t) = \mu t$, que es la tendencia del proceso, determina la forma característica de las trayectorias. Por lo tanto, X se le llama *movimiento browniano con tendencia (lineal)*.

(c) Movimiento browniano geométrico

El proceso sugerido por Black, Scholes y Merton está dado por

$$X_t = \exp \{ \mu t + \sigma W_t \}, \quad t \geq 0,$$

es decir, es la exponencial del movimiento browniano con tendencia, por esto se le llama *movimiento browniano geométrico*.

La transformada de Laplace de una variable aleatoria Z con distribución $N(0, 1)$ es $\mathbf{E} [e^{\xi Z}] = e^{\xi^2/2}$. Entonces

$$\begin{aligned} \mu(t) &= \mathbf{E} [X_t] = \mathbf{E} [\exp \{ \mu t + \sigma W_t \}] \\ &= \exp \{ \mu t \} \mathbf{E} \left[\exp \left\{ \sigma t^{1/2} W_1 \right\} \right] \\ &= \exp \left\{ \left(\mu + \frac{1}{2} \sigma^2 \right) t \right\}. \end{aligned}$$

Y la covarianza sería

$$\begin{aligned} C(t, s) &= \mathbf{E} [X_t X_s] - \mathbf{E} [X_t] \mathbf{E} [X_s] \\ &= \exp \mu (t + s) \mathbf{E} [\exp \{ \sigma (W_t + W_s) \}] - \exp \left\{ \left(\mu + \frac{1}{2} \sigma^2 \right) (t + s) \right\} \\ &= \exp \mu (t + s) \mathbf{E} [\exp \{ \sigma ((W_t - W_s) + 2W_s) \}] - \exp \left\{ \left(\mu + \frac{1}{2} \sigma^2 \right) (t + s) \right\} \\ &= \exp \mu (t + s) \mathbf{E} [\exp \{ \sigma (W_t - W_s) \}] \mathbf{E} [\exp \{ 2\sigma W_s \}] \\ &\quad - \exp \left\{ \left(\mu + \frac{1}{2} \sigma^2 \right) (t + s) \right\} \\ &= \exp \left\{ \left(\mu + \frac{1}{2} \sigma^2 \right) (t + s) \right\} (\exp \{ \sigma^2 s \} - 1). \end{aligned}$$

Las esperanzas y covarianzas del movimiento browniano geométrico son:

$$\mu(t) = \exp \left\{ \left(\mu + \frac{1}{2} \sigma^2 \right) t \right\}$$

y para $s \leq t$,

$$C(t, s) = \exp \left\{ \left(\mu + \frac{1}{2} \sigma^2 \right) (t + s) \right\} (\exp \{ \sigma^2 s \} - 1).$$

En particular, el movimiento browniano geométrico tiene varianza

$$\sigma^2(t) = \exp \{ (2\mu + \sigma^2) t \} (\exp \{ \sigma^2 t \} - 1).$$

1.5. Conclusiones

Hemos establecido las definiciones y propiedades básicas de los procesos estocásticos, que son necesarias para poder continuar con el siguiente capítulo, por ejemplo, definiciones sobre qué es un proceso estocástico, o sobre la función esperanza y covarianza de un proceso estocástico; y con respecto a sus propiedades tales como la estacionariedad, los incrementos independientes o la propiedad de Markov de un proceso estocástico. Todo esto era necesario explicarlo en este capítulo para así definir lo que son los procesos de difusión, tema central del siguiente capítulo.

Capítulo 2

Procesos de difusión

Cuando se tiene un número relativamente grande de observaciones de algún fenómeno de interés, se pueden presentar diferentes escenarios o estados transitorios como resultado del mismo comportamiento del fenómeno o de la relación de este con otros fenómenos. Ejemplos de estos fenómenos son el cambio de temperatura o la interacción entre las observaciones para el caso de la física; los nacimientos, muertes, mutaciones o infecciones en el área de la biología; o también, el consumo o el cambio de precios en economía; entre otros. Entonces, cuando el número de observaciones es grande, el tiempo de transición entre ellas puede ser relativamente pequeño si se elige una escala de tiempo conveniente para que sean relativamente frecuentes. Bajo una situación así es posible usar un modelo de aproximación de tipo difusión en el que la variable de interés y el tiempo sean continuos.

Las definiciones, resultados y parte de los cálculos de este capítulo se tomaron del libro Karlin y Taylor [7]. Una vez comprendido este capítulo es posible obtener los cálculos involucrados en los resultados del capítulo 3.

2.1. Definiciones

Sea I un intervalo de números reales que puede ser como (l, r) , $(l, r]$, $[l, r)$ o $[l, r]$, donde $l < r$ y l, r son números reales. Puede pasar que $l = -\infty$ y/o $r = +\infty$. Sea $\Omega_I = \{f; f: [0, \infty) \rightarrow I\}$, el conjunto de todas las funciones que van de $[0, \infty)$ a I . Sea $\Omega = \Omega_I$.

Sea \mathcal{B} la σ -álgebra de Borel de $[0, \infty)$ generada por el proceso estocástico $X = \{X(t, \omega), 0 \leq t < \infty, \omega \in \Omega\}$.

Sea τ una variable aleatoria en Ω con valores en $[0, \infty) \cup \infty$.

Definición 11 Decimos que τ es un **tiempo de Markov** asociado a X tal que para toda $t \geq 0$,

$$\{\tau \leq t\} \in \mathcal{F}_t^X,$$

donde $\mathcal{F}_t^X = \sigma\{X_s; 0 \leq s \leq t\}$.

Esta variable también es conocida como *tiempo de paro*. Un tiempo de Markov asociado a un proceso parametral continuo dado $\{X_t, 0 \leq t < \infty\}$, lo denotamos por $\tau(X_t)$.

Definición 12 Si para cualquier tiempo de Markov τ , la distribución de

$$X_{t_1+\tau}, X_{t_2+\tau}, \dots, X_{t_k+\tau},$$

con $t_1 < t_2 < \dots < t_k$, dado $X_s, s \leq \tau$ y $X_\tau = x$, es idéntica con la distribución de

$$X_{t_1}, X_{t_2}, \dots, X_{t_k},$$

con $t_1 < t_2 < \dots < t_k$, dado que $X_0 = x$, entonces el proceso de Markov se dice que posee la **propiedad fuerte de Markov** o es un **proceso fuerte de Markov**.

Con esto podemos dar una descripción general de un proceso de difusión.

Definición 13 Un proceso estocástico X con parámetros μ y σ^2 en tiempo continuo que posee la propiedad fuerte de Markov y que sus curvas aleatorias X_t son continuas con respecto a t , con probabilidad uno es llamado un **proceso de difusión**.

Consideremos un proceso de difusión $X = \{X_t, t \geq 0\}$ con I , un intervalo con puntos finales $l < r$, como su espacio de estados.

Definición 14 A los procesos estocásticos que inician desde cualquier punto en el interior de I y llegan a cualquier otro punto en el interior de I , que puede ser alcanzado con probabilidad positiva, son llamados **regulares**, es decir, sean $z \in I$, T_z el tiempo en que el proceso alcanza el valor z por primera vez y x el estado inicial del proceso, entonces el proceso es **regular** si

$$\mathbf{P}(T_z < \infty | X_0 = x) > 0,$$

para toda $z, x \in I$, donde $l < x, z < r$.

En el caso en el que z nunca sea alcanzado decimos por convención que $T_z = \infty$. De ahora en adelante sólo consideramos procesos de difusión regulares.

Cada proceso de difusión satisface la siguiente condición, que se le conoce por la *condición de Dynkin*.

Condición 1 Para cada $\epsilon > 0$,

$$\lim_{h \downarrow 0} \frac{1}{h} \mathbf{P}(|X_{t+h} - X_t| > \epsilon | X_t = x) = 0, \quad (2.1)$$

para toda $x \in I$.

Esta relación afirma que los desplazamientos grandes de orden mayor a un ϵ fijo son muy poco probables sobre intervalos de tiempo lo suficientemente pequeños. Con esto podemos decir que las trayectorias del proceso son continuas.

Comúnmente en la literatura los procesos de difusión que han aparecido como modelos de fenómenos físicos o biológicos son caracterizados por dos condiciones básicas además de (2.1), que además describen la media y la varianza de los desplazamientos infinitesimales.

Sea $\Delta_h X_t = X_{t+h} - X_t$. Con lo siguiente afirmamos la existencia de los límites:

$$\lim_{h \downarrow 0} \frac{1}{h} \mathbf{E} [\Delta_h X_t | X_t = x] = \mu(x, t) \quad (2.2)$$

y

$$\lim_{h \downarrow 0} \frac{1}{h} \mathbf{E} [(\Delta_h X_t)^2 | X_t = x] = \sigma^2(x, t), \quad (2.3)$$

cuando $l < x < r$.

Definición 15 Las funciones $\mu(x, t)$ y $\sigma^2(x, t)$ son llamadas los **parámetros infinitesimales** del proceso. En particular, $\mu(x, t)$ se le llama el **parámetro de tendencia, media infinitesimal o desplazamiento infinitesimal esperado** y a $\sigma^2(x, t)$ se le llama el **parámetro de difusión o varianza infinitesimal**.

Generalmente $\mu(x, t)$ y $\sigma^2(x, t)$ son funciones continuas de x y t , y un proceso regular cumple que $\sigma^2(x, t)$ es positiva para toda $l < x < r$ y $t > 0$. Otra forma de definir a $X = \{X_t, t \geq 0\}$ como un proceso de difusión es si se cumple que X es un proceso de Markov que satisface (2.1) y además que (2.2) y (2.3) sean funciones continuas de x y t .

Los *tiempos de alcance* de puntos y conjuntos son útiles en el estudio de procesos de difusión uni-dimensionales.

Definición 16 Definimos los **tiempos de alcance** del proceso estocástico $X = \{X_t, 0 \leq t < \infty\}$ al nivel z o del estado z por

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

Usamos la notación

$$T_{a,b} =: \min \{T_a, T_b\} = T_a \wedge T_b,$$

para referirnos al tiempo de alcance en a o en b , al que se llegue primero, ya sea $X_t = a$ o $X_t = b$.

Para procesos que inician en $X_0 = x$, $x \in (a, b)$, el tiempo de alcance es similar al **tiempo de salida** del intervalo (a, b) , porque lo que se quiere alcanzar es algún elemento de $(a, b)^c$,

$$T(a, b) =: \inf \{t \geq 0; X_t \notin (a, b)\},$$

con $X_0 = x$ en (a, b) .

Es conveniente tener condiciones suficientes bajo las cuales un proceso de Markov es un proceso de difusión. Introduzcamos el concepto de un proceso estocástico *estándar*.

Definición 17 Un proceso estocástico con la propiedad fuerte de Markov $X = \{X_t, t \geq 0\}$ es llamado un proceso **estándar** si las trayectorias del proceso poseen las siguientes **propiedades o condiciones de regularidad**:

(i) X_t es continua por la derecha, es decir, para toda $s \geq 0$,

$$\lim_{t \rightarrow s^+} X_t = X_s;$$

(ii) los límites izquierdos de X_t existen, es decir

$$\lim_{t \rightarrow s^-} X_t \text{ existe para toda } s > 0;$$

y

(iii) si $t_1 \leq t_2 \leq \dots$ son tiempos de Markov convergiendo a $t \leq \infty$, entonces $\lim_{n \rightarrow \infty} X_{t_n} = X_t$, cuando $t < \infty$.

Un resultado relevante es que cada proceso fuerte de Markov $X = \{X_t, t \geq 0\}$ continuo en probabilidad y bajo condiciones de regularidad, posee una versión equivalente, $\tilde{X} = \{\tilde{X}_t, t \geq 0\}$, que es un proceso estándar. Es decir,

Definición 18 Decimos que los procesos \tilde{X}_t y X_t son **equivalentes** si \tilde{X}_t y X_t tienen las mismas distribuciones finito-dimensionales.

Supondremos de ahora en adelante que los procesos de Markov son procesos estándares.

Para un proceso de difusión estacionario en \mathfrak{R} puede existir una *densidad de probabilidad estacionaria* $\bar{p}(y)$ tal que

$$\bar{p}(y) = \int_{-\infty}^{\infty} p(s, x; t, y) \bar{p}(x) dx = \int_{-\infty}^{\infty} p(x_s, y_t) \bar{p}(x) dx, \quad (2.4)$$

para toda $0 \leq s \leq t$ y $y \in \mathfrak{R}$, aunque no siempre se conoce de forma explícita. Cabe mencionar que $\bar{p}(y) \geq 0$ para toda $y \in \mathfrak{R}$ y $\int_{-\infty}^{\infty} \bar{p}(y) dy = 1$.

Una propiedad muy importante de los procesos de Markov, que se le llama *ergodicidad*, relaciona los promedios de tiempo a largo plazo de sus realizaciones con el promedio con respecto a la distribución estacionaria.

Definición 19 Un proceso de difusión estacionario $X = \{X_t, t \geq 0\}$ se dice que es **ergódico** si el siguiente límite del promedio en el tiempo existe y es igual al promedio con respecto a \bar{p} , con probabilidad 1, es decir si

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T f(X_t) dt = \int_{-\infty}^{\infty} f(x) \bar{p}(x) dx, \quad (2.5)$$

para todas las funciones medibles acotadas $f : \mathfrak{R} \rightarrow \mathfrak{R}$.

Las trayectorias de un proceso de difusión son casi seguramente funciones continuas en el tiempo, aunque no son necesariamente diferenciables. Podemos definir la continuidad de un proceso estocástico $X(t)$ en t de diferentes formas de modo correspondiente a las diferentes convergencias de sucesiones de variables aleatorias. En particular, tenemos que para un instante de tiempo t fijo:

I. *Continuidad con probabilidad uno:*

$$\mathbf{P} \left(\left\{ \omega \in \Omega : \lim_{s \rightarrow t} |X(s, \omega) - X(t, \omega)| = 0 \right\} \right) = 1. \quad (2.6)$$

II. *Continuidad en media cuadrada:* $\mathbf{E} [X_t^2] < \infty$ y

$$\lim_{s \rightarrow t} \mathbf{E} [|X_s - X_t|^2] = 0. \quad (2.7)$$

III. *Continuidad en probabilidad:* para toda $\epsilon > 0$

$$\lim_{s \rightarrow t} \mathbf{P} (\{ \omega \in \Omega : |X(s, \omega) - X(t, \omega)| \geq \epsilon \}) = 0. \quad (2.8)$$

IV. *Continuidad en distribución:* para todos los puntos de continuidad de F_t

$$\lim_{s \rightarrow t} F_s(x) = F_t(x). \quad (2.9)$$

Estas formas de convergencia están relacionadas unas con otras del mismo modo que las convergencias de sucesiones de variables aleatorias. En particular: **(I)** implica **(III)**, **(II)** implica **(III)** y **(III)** implica **(IV)**. Sin embargo **(III)** se puede escribir en términos de la métrica $\lim_{n \rightarrow \infty} \mathbf{E} [|X_n - X| / (1 + |X_n - X|)] = 0$ como

$$\lim_{s \rightarrow t} \mathbf{E} \left[\frac{|X_s - X_t|}{1 + |X_s - X_t|} \right] = 0. \quad (2.10)$$

Para un proceso de difusión, el concepto de continuidad apropiado es la continuidad de trayectorias. Hay un criterio de Kolmogorov que trata de la continuidad de trayectorias de un proceso estocástico en tiempo continuo $X = \{X_t, t \in \mathcal{T}\}$ conocido como el *teorema de continuidad de Kolmogorov* que establece que si $\mathbf{E} [|X_t - X_s|^\gamma] \rightarrow 0$ rápidamente, donde γ es una constante positiva, entonces X tiene una versión equivalente continua.

Una condición suficiente para que un proceso estándar sea una difusión es que cumplan la condición de Dynkin dada en (2.1). El siguiente teorema afirma la suficiencia de esta condición.

Teorema 3 *Sea $\{X_t, t \geq 0\}$ un proceso estándar y supongamos que se cumple la condición de Dynkin (2.1). Entonces $\{X_t, t \geq 0\}$ es un proceso de difusión.*

Demostración Ver teorema 1.1 en la página 163 de Karlin y Taylor [7]. \square

Entonces podemos enfocarnos en comprobar si un proceso cumple la condición de Dynkin y es aquí donde el siguiente lema cubre esa función.

Lema 2 Si un proceso estándar satisface la siguiente condición, a veces llamada condición de Dynkin, de momentos infinitesimales,

$$\lim_{h \downarrow 0} \frac{1}{h} \mathbf{E} [|\Delta_h X_t|^p | X_t = x] = 0, \quad (2.11)$$

para alguna $p > 2$ uniformemente en x en cualquier subintervalo compacto de (l, r) y t en cualquier intervalo acotado $[0, N]$, entonces la condición de Dynkin (2.1) es satisfecha.

Demostración Ver lema 1.1 en la página 165 de Karlin y Taylor [7]. \square

Otro criterio muy utilizado para verificar que un proceso estocástico unidimensional $\{X_t, t \geq 0\}$ tiene una trayectoria continua, aún cuando no posea la propiedad de Markov, es la condición de Kolmogorov:

Condición 2 Sea $\{X_t, t \geq 0\}$ un proceso estocástico que cumple

$$\mathbf{E} [|X_t - X_s|^\gamma] \leq C |\varphi_t - \varphi_s|^{1+\alpha}, \quad (2.12)$$

para toda $s, t \geq 0$, donde α, γ y C son constantes positivas independientes de s y t , y φ es una función no decreciente continua. Entonces existe una versión equivalente \tilde{X} que posee trayectorias continuas.

El movimiento browniano estándar $W = \{W_t, t \geq 0\}$ satisface la condición 2 y por lo tanto tenemos que

$$\mathbf{E} [|W_t - W_s|^4] = 3 |t - s|^2$$

para toda $s, t \geq 0$. Entonces el movimiento browniano tiene casi seguramente trayectorias continuas.

Para procesos d -dimensionales la condición de Kolmogorov tiene que ser modificada a

$$\mathbf{E} [|X_t - X_s|^\gamma] \leq C |\varphi_t - \varphi_s|^{d+\alpha},$$

con α, γ y C constantes positivas y el exponente de la derecha ahora excede d .

Transformaciones elementales de procesos

A partir de ciertas transformaciones aplicadas a algunos procesos dados podemos determinar la forma y los parámetros infinitesimales de esos procesos.

Una función continua estrictamente creciente g puede ser usada para transformar un proceso estocástico arbitrario $\{X_t, t \geq 0\}$ en otro proceso $\{Y_t, t \geq 0\}$ definido por $Y_t = g(X_t)$. Si $\{X_t, t \geq 0\}$ es un proceso de Markov con trayectorias continuas, es decir una difusión, entonces también lo es $\{Y_t, t \geq 0\}$, ya que g se ha supuesto continua y monótona, por lo que $\{Y_t, t \geq 0\}$ es un proceso de Markov con trayectorias continuas. Además, si $\{X_t, t \geq 0\}$ tiene parámetros

infinitesimales $\mu(x)$ y $\sigma^2(x)$ dados por (2.2) y (2.3), y g tiene dos derivadas continuas uniformemente g' y g'' , entonces $\{Y_t, t \geq 0\}$ también tiene parámetros infinitesimales.

Con el siguiente teorema podemos determinar los parámetros infinitesimales de $\{Y_t = g(X_t), t \geq 0\}$.

Teorema 4 *Sea $\{X_t, t \geq 0\}$ una difusión regular con el intervalo I como su espacio de estados teniendo como puntos finales l y r . Supongamos que $\{X_t, t \geq 0\}$ tiene parámetros infinitesimales $\mu(x)$ y $\sigma^2(x)$. Sea g una función estrictamente monótona en I con segunda derivada continua $g''(x)$, para $l < x < r$. Entonces $\{Y_t = g(X_t), t \geq 0\}$ define un proceso de difusión en el intervalo con puntos finales $g(l)$ y $g(r)$, y $\{Y_t, t \geq 0\}$ tiene parámetros infinitesimales*

$$\begin{aligned}\mu_Y(y) &= \frac{1}{2}\sigma^2(x)g''(x) + \mu(x)g'(x) \\ \sigma_Y^2(y) &= \sigma^2(x)[g'(x)]^2\end{aligned}\tag{2.13}$$

donde $y = g(x)$.

Demostración Ver teorema 2.1 en la página 173 de Karlin y Taylor [7]. \square

2.2. Ejemplos de procesos de difusión

A. Movimiento browniano

Un movimiento browniano es un proceso de difusión regular en el intervalo $(-\infty, +\infty)$ con $\mu(x) = 0$ y $\sigma^2(x) = \sigma^2$, para toda x . Podemos calcular estos parámetros infinitesimales tomando en cuenta que $\Delta_h X = X_h - X_0$ se distribuye normal con media cero y varianza $\sigma^2 h$, es decir,

$$\mathbf{E}[\Delta_h X | X_0 = x] = 0$$

y

$$\mathbf{E}[(\Delta_h X)^2 | X_0 = x] = \sigma^2 h.$$

Para este caso, se comprueba solamente la ecuación (2.11) cuando $p = 4$, es decir la condición de Dynkin. Entonces tenemos

$$\mathbf{E}[(\Delta_h X)^4 | X_0 = x] = \frac{1}{\sqrt{2\pi h}\sigma} \int_{-\infty}^{\infty} y^4 \exp\left\{-\frac{1}{2} \frac{y^2}{\sigma^2 h}\right\} dy,$$

sea $\frac{y^2}{\sigma^2 h} = 2\xi$, entonces

$$\mathbf{E}[(\Delta_h X)^4 | X_0 = x] = \frac{4\sigma^4 h^2}{\sqrt{\pi}} \int_0^{\infty} \xi^{3/2} e^{-\xi} d\xi = 3h^2 \sigma^4$$

y (2.11) se cumple.

Ahora si sumamos la tendencia μt a un movimiento browniano $\{X_t, t \geq 0\}$ obtenemos un movimiento browniano con tendencia $\{X_t + \mu t, t \geq 0\}$. En este caso el parámetro de tendencia es μ , mientras que el parámetro varianza es σ^2 .

B. Movimiento browniano absorbido y el reflejado

El movimiento browniano absorbido y el reflejado son procesos de difusión regulares definidos en el espacio de estados $I = [0, \infty)$. Estos dos procesos inician desde un punto $X_0 = x$ en el interior del intervalo, es decir, $x > 0$ y tanto el proceso absorbido como el reflejado actúan como un movimiento browniano hasta el nivel cero, es decir, hasta que el estado cero es alcanzado por primera vez. Por lo tanto, los parámetros infinitesimales son $\mu(x) = 0$ y $\sigma^2(x) = \sigma^2$ para $0 < x < \infty$.

Entonces tenemos dos procesos de difusión con el mismo espacio de estados y los mismos parámetros infinitesimales, esto significa que estos parámetros, $\mu(x)$ y $\sigma^2(x)$, no definen únicamente un proceso de difusión. Los parámetros infinitesimales dirigen la evolución del proceso sólo mientras el proceso este en el interior del espacio de estados I . Para definir completamente un proceso de difusión, tenemos que anexar condiciones de frontera para especificar el comportamiento en cualquiera de los puntos finales de I de modo que el proceso pueda alcanzar estos puntos finales. Y es así como podemos diferenciar estos dos procesos.

C. Proceso Ornstein-Uhlenbeck

Este proceso de difusión tiene los números reales, $I = (-\infty, \infty)$, como su espacio de estados y $\mu(t) = -\alpha t$, $\sigma^2(t) = \sigma^2$, como sus parámetros infinitesimales, donde α y σ son constantes positivas arbitrarias. El parámetro de tendencia infinitesimal refleja una fuerza que regresa hacia el origen y de una magnitud proporcional a la distancia.

Si el movimiento browniano representa la posición de una partícula, la derivada del movimiento browniano representa la velocidad de la partícula. Pero esta derivada no existe en cualquier punto del tiempo. El proceso Ornstein-Uhlenbeck es un modelo alternativo que supera este defecto al modelar directamente la velocidad de la partícula como una función del tiempo. Dos factores son considerados que afectan esta velocidad durante una duración pequeña de tiempo. Primero, la resistencia por la fricción de medio circunvecino es tomado en cuenta para reducir la magnitud de la velocidad por una cantidad proporcional. Segundo, hay un cambio en la velocidad causado por las coaliciones aleatorias con las partículas vecinas. Estos dos factores nos llevan a las especificaciones $\mu(t) = -\alpha t$ y $\sigma^2(t) = \sigma^2$, respectivamente.

D. Movimiento browniano geométrico

Aplicaremos una transformación como la planteada en el teorema 4 para definir el movimiento browniano geométrico.

Sea $\{X_t, t \geq 0\}$ el proceso proveniente de un movimiento browniano con parámetros de tendencia μ y de difusión σ^2 . El proceso $\{Y_t, t \geq 0\}$ definido por $Y_t = \exp\{X_t\}$ se llama movimiento browniano geométrico. El espacio de estados es el intervalo $(0, \infty)$. Si $t_0 < t_1 < \dots < t_n$ son puntos en el

tiempo, los radios sucesivos

$$\frac{Y_{t_1}}{Y_{t_0}}, \dots, \frac{Y_{t_n}}{Y_{t_{n-1}}},$$

son variables aleatorias independientes para el movimiento browniano geométrico, de manera que los cambios porcentuales sobre intervalos de tiempo ajenos son independientes. Con $y = g(x) = e^x$ tenemos que $g'(x) = g''(x) = y$, por lo tanto los parámetros infinitesimales para el movimiento browniano geométrico Y son

$$\mu(x) = \left(\mu + \frac{1}{2}\sigma^2 \right) x \quad \text{y} \quad \sigma^2(x) = \sigma^2 x^2.$$

El movimiento browniano geométrico es usado a menudo para modelar precios de activos, como lo son las acciones, que son comerciadas en un mercado perfecto. Los precios no son negativos y exhiben un crecimiento exponencial a lo largo de su trayecto, dos propiedades muy marcadas para el movimiento browniano geométrico.

E. El proceso de Bessel

El proceso de Bessel es definido como la distancia Euclideana desde el origen de un movimiento browniano n -dimensional hasta el tiempo t y es un proceso de Markov. Usemos el teorema 4 para determinar los parámetros infinitesimales apropiados. Primero establezcamos a

$$Z(t) = X_1(t)^2 + \dots + X_n(t)^2,$$

donde $\{X_i(t), t \geq 0\}$ son procesos de movimientos brownianos estándares independientes. Condicionemos que $X_i(t) = x_i$, $i = 1, \dots, n$, y denotemos por

$$X_i(t + \Delta t) = x_i + \Delta_h X_i \quad \text{y} \quad Z(t + \Delta t) = z + \Delta Z,$$

donde $z = x_1^2 + \dots + x_n^2$. Entonces desarrollando ΔZ de la última igualdad

$$\begin{aligned} \Delta Z &= (X_1(t + \Delta t))^2 - x_1^2 + \dots + (X_n(t + \Delta t))^2 - x_n^2 \\ &= 2(x_1 \Delta_h X_1 + \dots + x_n \Delta_h X_n) + [(\Delta_h X_1)^2 + \dots + (\Delta_h X_n)^2], \end{aligned}$$

como $\Delta_h X_1, \dots, \Delta_h X_n$ son independientes y tienen distribución normal con media cero y varianza Δt , se sigue que

$$\mathbf{E}[\Delta Z | Z(t) = z] = n\Delta t$$

y

$$\begin{aligned} \mathbf{E}[(\Delta Z)^2 | Z(t) = z] &= 4(x_1^2 + \dots + x_n^2) \Delta t + o(\Delta t) \\ &= 4z\Delta t + o(\Delta t), \end{aligned}$$

donde $o(\Delta t)$ representa las esperanzas de términos de la forma $(\Delta_h X_i)^4$ y de ordenes mayores. Similarmente encontramos que

$$\mathbf{E} [(\Delta Z)^4 | Z(t) = z] = o((\Delta t)^2)$$

y así la condición (2.11) se cumple cuando $p = 4$. Concluimos por lo tanto que $Z(t)$ es un proceso de difusión con los parámetros infinitesimales

$$\mu(x) = n \quad \text{y} \quad \sigma^2(x) = 4x.$$

El proceso de Bessel $\{Y_t, t \geq 0\}$ se define por $Y_t = g[Z(t)]$ para $g(z) = \sqrt{z}$. Si $y = g(z)$, entonces

$$\mu(z) = n, \quad \sigma^2(z) = 4y^2,$$

y

$$g'(z) = \frac{1}{2\sqrt{z}} = \frac{1}{2y}, \quad g''(z) = -\frac{1}{4z^{3/2}} = -\frac{1}{4y^3}.$$

Aplicando el teorema 4 para obtener los parámetros infinitesimales del proceso de Bessel Y , así

$$\mu(y) = \frac{n-1}{2y} \quad \text{y} \quad \sigma^2(y) = 1.$$

Para $n = 1$, el proceso de Bessel se puede identificar con el movimiento browniano reflejado que tiene a $\mu(y) = 0$ y $\sigma^2(y) = 1$.

2.3. Propiedades generales

Supongamos que $\{X_t, t \geq 0\}$ sea un proceso de difusión homogéneo en el tiempo que satisface las siguientes condiciones:

1. el espacio de estados es un intervalo I de la forma $[l, r]$, $(l, r]$, $[l, r)$ o (l, r) , donde $-\infty \leq l < r \leq \infty$,
2. el proceso es regular en el interior de I , es decir,

$$\mathbf{P}(T_z < \infty | X_0 = x) > 0,$$

$l < x, z < r$, donde T_z es el primer tiempo, si lo hay, que el proceso alcanza el valor z ,

3. el proceso tiene parámetros infinitesimales $\mu(x)$ y $\sigma^2(x)$, para $l < x < r$, donde $\Delta_h X = X_h - X_0$ y

$$\mu(x) = \lim_{h \downarrow 0} \frac{1}{h} \mathbf{E} [\Delta_h X | X_0 = x]$$

y

$$\sigma^2(x) = \lim_{h \downarrow 0} \frac{1}{h} \mathbf{E} [(\Delta_h X)^2 | X_0 = x],$$

4. los parámetros infinitesimales $\mu(x)$ y $\sigma^2(x)$ son funciones continuas de x , y $\sigma^2 > 0$, para $l < x < r$.

Sean a y b fijos tales que $l < a < b < r$, sean T_z el primer tiempo de alcance de z y $T_{a,b}$ el primer tiempo en que el proceso alcanza a a o a b .

Por ahora concentrémonos en tres problemas:

Problema A. Encontrar

$$u(x) = \mathbf{P}(T_b < T_a | X_0 = x),$$

$a < x < b$, la probabilidad de que el proceso alcance b antes que a a .

Problema B. Encontrar

$$v(x) = \mathbf{E}[T_{a,b} | X_0 = x]$$

$a < x < b$, el tiempo esperado para alcanzar a a o alcanzar a b .

Problema C. Para una función acotada y continua g , encontrar

$$w(x) = \mathbf{E} \left[\int_0^{T_{a,b}} g(X_s) ds | X_0 = x \right],$$

$a < x < b$.

Como las trayectorias del proceso de difusión son continuas, la integral $A = \int_0^{T_{a,b}} g(X_s) ds$ está bien definida. Si $g(x)$ representa un valor de costo cuando el proceso este en el estado x , entonces A es el costo total hasta el tiempo en que a o b sean alcanzados por primera vez. Si $g(x) = 1$ para toda x , entonces $A = T_{a,b}$, que es el tiempo para alcanzar a a o alcanzar a b , de manera que el Problema B es un caso especial del Problema C.

Bajo las condiciones (1)-(4), mencionadas al inicio de esta sección, mostramos que $v(x)$ y $w(x)$ son finitas, que $u(x)$, $v(x)$ y $w(x)$ poseen dos derivadas acotadas para $a < x < b$, y que estas funciones satisfacen las siguientes ecuaciones diferenciales:

■ **Ecuación A**

$$0 = \mu(x) \frac{du}{dx} + \frac{1}{2} \sigma^2(x) \frac{d^2u}{dx^2}; \quad a < x < b, \quad u(a) = 0, \quad u(b) = 1, \quad (2.14)$$

■ **Ecuación B**

$$-1 = \mu(x) \frac{dv}{dx} + \frac{1}{2} \sigma^2(x) \frac{d^2v}{dx^2}; \quad a < x < b, \quad v(a) = v(b) = 0, \quad (2.15)$$

■ **Ecuación C**

$$-g(x) = \mu(x) \frac{dw}{dx} + \frac{1}{2} \sigma^2(x) \frac{d^2w}{dx^2}; \quad a < x < b, \quad w(a) = w(b) = 0, \quad (2.16)$$

Las soluciones explícitas de estas ecuaciones diferenciales están en (2.25) - (2.27).

Consideremos el Problema A. Las condiciones de frontera son $u(a) = 0$ y $u(b) = 1$, ya que $u(x)$ es la probabilidad de que el proceso alcance b antes que a con valor inicial en x . Ahora consideremos $a < x < b$ y elijamos un tiempo h de duración lo suficientemente pequeño, de manera que la probabilidad de alcanzar a o b antes del tiempo h es insignificante. Condicionemos la posición de X_h , entonces al tiempo h la probabilidad de alcanzar b antes que a es $u(X_h)$. Al usar la ley de probabilidad total tenemos

$$u(x) = \mathbf{E}[u(X_h) | X_0 = x] + o(h),$$

donde el término de error $o(h)$ es de orden más pequeño que h . Ahora escribimos $\Delta_h X = X_h - x$ y usemos la expansión de Taylor

$$\begin{aligned} u(X_h) &= u(x + \Delta_h X) \\ &= u(x) + (\Delta_h X)u'(x) + \frac{1}{2}(\Delta_h X)^2 u''(x) + \sum_{i=3}^{\infty} \frac{1}{i} (\Delta_h X)^i \frac{d^i(u(x))}{dx^i}, \end{aligned}$$

donde el cuarto y demás términos son de orden más pequeño que $(\Delta_h X)^2$ y pueden ser omitidos. Como

$$\mathbf{E}[\Delta_h X | X_0 = x] = \mu(x)h + o(h)$$

y

$$\mathbf{E}[(\Delta_h X)^2 | X_0 = x] = \sigma^2(x)h + o(h)$$

tenemos

$$\begin{aligned} u(x) &= \mathbf{E}[u(X_h) | X_0 = x] + o(h) \\ &= \mathbf{E}[u(x + \Delta_h X) | X_0 = x] + o(h) \\ &= u(x) + \mathbf{E}[\Delta_h X | X_0 = x] u'(x) + \frac{1}{2} \mathbf{E}[(\Delta_h X)^2 | X_0 = x] u''(x) + o(h) \\ &= u(x) + \mu(x)u'(x)h + \frac{1}{2}\sigma^2(x)u''(x)h + o(h). \end{aligned}$$

Sustraemos $u(x)$ de ambos lados, dividimos por h y hacemos decrecer h hasta cero para concluir que

$$0 = \mu(x)u'(x) + \frac{1}{2}\sigma^2(x)u''(x),$$

$a < x < b$, que es la ecuación diferencial para el Problema A.

La derivación de (2.16) es similar a la derivación de (2.14). Las condiciones de frontera son $w(a) = w(b) = 0$. Elegimos un tiempo h de duración corta como antes. Condicionamos la posición de X_h , entonces en el tiempo h la esperanza de la integral total es la esperanza de la función g hasta el tiempo h , $\int_0^h g(X_s) ds$,

más la esperanza de la función g en el tiempo restante, $\int_h^{T_{a,b}} g(X_s) ds$. Condicionando en $X_h = z$, la media condicional de $\int_h^{T_{a,b}} g(X_s) ds$ es

$$\mathbf{E} \left[\int_h^{T_{a,b}} g(X_s) ds \mid X_h = z \right] = \mathbf{E} \left[\int_0^{T_{a,b}} g(X_s) ds \mid X_0 = z \right]$$

la última igualdad es debido a la propiedad de Markov y la estacionariedad del proceso, y si tomamos la definición de w , obtenemos la siguiente igualdad

$$\mathbf{E} \left[\int_h^{T_{a,b}} g(X_s) ds \mid X_h = z \right] = w(z). \quad (2.17)$$

Por su parte, enfoquémonos en la primera integral, $\int_0^h g(X_s) ds$. Como las trayectorias del proceso y la función g son continuas, tenemos la aproximación

$$\mathbf{E} \left[\int_0^h g(X_s) ds \mid X_0 = x \right] = g(x)h + o(h). \quad (2.18)$$

Ahora, tomemos en cuenta la definición de w y las integrales en (2.17) y (2.18). Para $a < x < b$ tenemos

$$w(x) = \mathbf{E} \left[\int_0^h g(X_s) ds \mid X_0 = x \right] + \mathbf{E} [w(X_h) \mid X_0 = x], \quad (2.19)$$

y como ya habíamos definido antes $\Delta_h X = X_h - x$, entonces

$$\begin{aligned} \mathbf{E} [w(X_h) \mid X_0 = x] &= \mathbf{E} [w(x + \Delta_h X) \mid X_0 = x] \\ &= w(x) + \mu(x)w'(x)h + \frac{1}{2}\sigma^2(x)w''(x)h + o(h), \end{aligned}$$

de manera que (2.19) se vuelve

$$w(x) = g(x)h + w(x) + \mu(x)w'(x)h + \frac{1}{2}\sigma^2(x)w''(x)h + o(h).$$

Así también hemos obtenido la ecuación diferencial del Problema B, pues como ya habíamos mencionado la ecuación B es un caso particular de la ecuación C. Sustraemos $w(x)$ de ambos lados, dividimos por h y hacemos tender h a cero obteniendo la ecuación diferencial para el Problema C.

Volvamos a las soluciones de los tres problemas. Pero antes definimos tres funciones que se usan a lo largo de este trabajo.

Definición 20 Recordemos que $\sigma^2(x) > 0$ para $l < x < r$. Definimos a la *densidad de escala* por

$$s(x) = \exp \left\{ - \int^x \frac{2\mu(\xi)}{\sigma^2(\xi)} d\xi \right\}, \quad (2.20)$$

para $l < x < r$. A la **función de escala** del proceso la definimos por

$$\begin{aligned} S(x) &= \int^x s(\eta) d\eta \\ &= \int^x \exp \left\{ - \int^\eta \frac{2\mu(\xi)}{\sigma^2(\xi)} d\xi \right\} d\eta, \end{aligned} \quad (2.21)$$

para $l < x < r$. Y la **densidad de velocidad** la definimos por

$$m(x) = \frac{1}{\sigma^2(x)s(x)} \quad (2.22)$$

para $l < x < r$.

Usamos integrales indefinidas aquí por razones que más adelante se aclararán.

Cada una de las ecuaciones (2.14)-(2.16) incluye el operador diferencial L definido por

$$Lf(x) = \mu(x)f'(x) + \frac{1}{2}\sigma^2(x)f''(x),$$

donde f es una función dos veces diferenciable continuamente en (a, b) . Una aproximación a la solución de las ecuaciones (2.14)-(2.16) es expresar este operador como diferenciaciones sucesivas con respecto a la medida de escala y de velocidad. Con este fin, primero verificamos que $s'(x)/s(x) = -2\mu(x)/\sigma^2(x)$. Entonces introducimos $1/s(x)$ como un factor de integración, por lo que reagrupamos variables y obtenemos

$$Lf(x) = \frac{1}{2} \left(\frac{1}{1/(\sigma^2(x)s(x))} \right) \frac{d}{dx} \left[\frac{1}{s(x)} \frac{df(x)}{dx} \right]. \quad (2.23)$$

Para obtener una expresión más breve y significativa para L , escribimos $s(x) = dS(x)/dx$ en la forma diferencial $dS = s(x)dx$ y similarmente escribimos la densidad de velocidad como una diferencial de una *medida de velocidad* M en la forma $dM = m(x)dx$. En términos de estas diferenciales, el operador L en (2.23) es

$$Lf(x) = \frac{1}{2} \frac{d}{dM} \left[\frac{df(x)}{dS} \right]. \quad (2.24)$$

A la expresión anterior se le llama la *representación canónica del operador diferencial infinitesimal* asociado con el proceso de difusión.

En términos de esta representación canónica, la ecuación diferencial para el Problema A es

$$\frac{1}{2} \frac{d}{dM} \left[\frac{du(x)}{dS} \right] = 0,$$

para $a < x < b$, $u(a) = 0$ y $u(b) = 1$. La solución se sigue directamente de dos integraciones sucesivas.

Denotemos las constantes de integración por α y β , integremos una vez para obtener $du(x)/dS(x) = \beta$, y de nuevo para obtener $u(x) = \alpha + \beta S(x)$, $a \leq x \leq b$. Con la condición de frontera $u(a) = 0$, podemos determinar que $\alpha = -\beta S(a)$, y con $u(b) = 1$ que $\beta = 1/[S(b) - S(a)]$. Haciendo un resumen de este análisis tenemos:

Solución A.

$$u(x) = \frac{S(x) - S(a)}{S(b) - S(a)}, \text{ para } a \leq x \leq b. \quad (2.25)$$

Note que $u(x)$ no se altera si $S(x)$ es reemplazado por $S^*(x) = \alpha + \beta S(x)$ para cualesquiera α y $\beta \neq 0$. Así que, si $S(x)$ es una función de escala, también lo es $S^*(x)$. Entonces se sigue que podemos especificar la función de escala $S(x)$ si usamos integrales indefinidas pues el resultado $u(x)$ no depende de los límites inferiores de integración.

Observación 3 *La función de escala puede ser usada para reescalar el espacio de estados (l, r) en términos de las probabilidades de alcance de varios niveles, y este uso motiva su nombre. Fijemos un punto x_0 como el origen y determinemos la función de escala al realizar una traslación, si es necesario, lo que causa que $S(x_0) = 0$. Entonces formamos el proceso $Y(t) = S(X_t)$ en el intervalo $(S(l), S(r))$. Como S es estrictamente monótona y dos veces continuamente diferenciable, recurrimos al teorema 4 que establece que los parámetros infinitesimales del proceso $\{Y_t, t \geq 0\}$ son*

$$\mu_Y(y) = \frac{1}{2}\sigma^2(x)S''(x) + \mu(x)S'(x) = 0$$

y

$$\sigma_Y^2(y) = \sigma^2(x) [S'(x)]^2 = \sigma^2(x)S^2(x),$$

donde $y = S(x)$. La medida de escala para el proceso $\{Y_t, t \geq 0\}$ es por lo tanto $S_Y(y) = y$, o lo que es equivalente, $S_Y(y) = y = \alpha + \beta y$, con $\alpha, \beta \neq 0$, constantes. Un proceso $\{Y_t, t \geq 0\}$ cuya función de escala es lineal se dice que esta en escala natural o canónica ya que las probabilidades de alcance

$$\mathbf{P}(T_a(Y) < T_b(Y) | Y(0) = y) = \frac{b-y}{b-a}, \text{ para } a < y < b,$$

son proporcionales a las distancias actuales.

Continuemos con el Problema C pues el Problema B es un caso especial del Problema C. En la representación canónica, la ecuación diferencial es escrita como

$$\frac{1}{2} \frac{d}{dM} \left[\frac{dw(x)}{dS(x)} \right] = -g(x),$$

para $a < x < b$, sujeto a las condiciones de frontera $w(a) = w(b) = 0$. En la primera integración, obtenemos

$$\frac{dw(\eta)}{dS} = -2 \int_a^\eta g(\xi) dM(\xi) + \beta$$

$$= -2 \int_a^\eta g(\xi) m(\xi) d\xi + \beta,$$

y después de la segunda integración tenemos

$$w(x) = -2 \int_a^x \left[\int_a^\eta g(\xi) m(\xi) d\xi \right] dS(\eta) + \beta [S(x) - S(a)] + \alpha.$$

Como $w(a) = 0$, entonces $\alpha = 0$ y $w(b) = 0$ requiere que

$$\beta = \frac{2}{S(b) - S(a)} \int_a^b \left[\int_a^\eta g(\xi) m(\xi) d\xi \right] dS(\eta).$$

Al usar $u(x) = [S(x) - S(a)] / [S(b) - S(a)]$, obtenemos

$$w(x) = 2 \left\{ u(x) \int_a^b \left[\int_a^\eta g(\xi) m(\xi) d\xi \right] dS(\eta) - \int_a^x \left[\int_a^\eta g(\xi) m(\xi) d\xi \right] dS(\eta) \right\}.$$

Cambiamos el orden de integración y hagamos manipulaciones algebraicas y se reduce a lo siguiente:

Solución C.

$$w(x) = 2 \left\{ u(x) \int_x^b [S(b) - S(\xi)] m(\xi) g(\xi) d\xi \right. \\ \left. + [1 - u(x)] \int_a^x [S(\xi) - S(a)] m(\xi) g(\xi) d\xi \right\} \quad (2.26)$$

Como mencionamos antes, el caso especial $g(\xi) \equiv 1$ da la solución al Problema B.

Solución B.

$$v(x) = 2 \left\{ u(x) \int_x^b [S(b) - S(\xi)] m(\xi) d\xi \right. \\ \left. + [1 - u(x)] \int_a^x [S(\xi) - S(a)] m(\xi) d\xi \right\} \quad (2.27)$$

Veamos lo que obtenemos si consideramos por un momento un proceso en escala natural en donde $S(x) = x$ y $s(x) = S'(x) = 1$. Calculemos el tiempo esperado para salir del intervalo $(x - \epsilon, x + \epsilon)$ iniciando en x . Insertemos $a = x - \epsilon$, $b = x + \epsilon$, $u(x) = \frac{1}{2}$ y $S(x) = x$ en (2.27), así obtenemos

$$v(x) = \mathbf{E} [T_{x-\epsilon, x+\epsilon} | X_0 = x] = \int_x^{x+\epsilon} (x + \epsilon - \xi) m(\xi) d\xi + \int_{x-\epsilon}^x (\xi - x + \epsilon) m(\xi) d\xi,$$

de donde

$$\begin{aligned} & \lim_{\epsilon \downarrow 0} \frac{1}{\epsilon^2} \mathbf{E} [T_{x-\epsilon, x+\epsilon} | X_0 = x] \\ &= \lim_{\epsilon \downarrow 0} \frac{1}{\epsilon^2} \int_x^{x+\epsilon} (x + \epsilon - \xi) m(\xi) d\xi + \lim_{\epsilon \downarrow 0} \frac{1}{\epsilon^2} \int_{x-\epsilon}^x (\xi - x + \epsilon) m(\xi) d\xi. \end{aligned}$$

Entonces la densidad de velocidad $m(x)$ se puede construir como la velocidad en la que el proceso tarda en encontrarse en el estado x .

Observación 4 $v(x)$ es útil para calcular la cantidad de tiempo esperado que el proceso pasa en un intervalo $[\xi, \xi + \Delta)$ antes de alcanzar a $T_{a,b}$. Si hacemos que Δ sea lo más pequeño posible hasta que sea $d\xi$ obtenemos el tiempo esperado local en ξ . Para hacer estas ideas más precisas, escribimos la solución del Problema C,

$$w(x) = \mathbf{E} \left[\int_0^{T_{a,b}} g(X_s) ds | X_0 = x \right], \quad (2.28)$$

para $a \leq x \leq b$, en la forma

$$w(x) = \int_a^b G(x, \xi) g(\xi) d\xi, \quad (2.29)$$

donde

$$G(x, \xi) = \begin{cases} 2 \frac{[S(x)-S(a)][S(b)-S(\xi)]}{S(b)-S(a)} \frac{1}{\sigma^2(\xi)s(\xi)}, & a \leq x \leq \xi \leq b, \\ 2 \frac{[S(b)-S(x)][S(\xi)-S(a)]}{S(b)-S(a)} \frac{1}{\sigma^2(\xi)s(\xi)}, & a \leq \xi \leq x \leq b. \end{cases} \quad (2.30)$$

A la función $G(x, \xi)$ se le llama la función de Green del proceso en el intervalo $[a, b]$. Al determinar el tiempo esperado que el proceso pasa en el intervalo $[\xi, \xi + \Delta)$ antes de alcanzar a $T_{a,b}$ es equivalente a evaluar

$$w(x) = \mathbf{E} \left[\int_0^{T_{a,b}} g(X_s) ds | X_0 = x \right],$$

para

$$g(x) = \begin{cases} 1, & \xi \leq x < \xi + \Delta, \\ 0, & \text{en otro caso.} \end{cases} \quad (2.31)$$

Sigamos el formato de (2.29), es decir

$$w(x) = \int_{\xi}^{\xi+\Delta} G(x, \eta) d\eta. \quad (2.32)$$

Mientras la función $g(x)$, como se define en (2.31), no satisfaga la suposición de continuidad del Problema C, (2.32) se puede establecer definiendo funciones continuas aproximadas convenientes.

Al hacer que Δ sea lo más pequeño posible como $d\xi$, vemos de (2.32) que $G(x, \xi) d\xi$ mide el tiempo esperado que el proceso pasa en el intervalo infinitesimal $[\xi, \xi + d\xi)$ antes de alcanzar a $T_{a,b}$ dado que $X_0 = x$, $x \in [\xi, \xi + d\xi)$.

Algunos casos concretos

Tomando en cuenta las soluciones (2.25)-(2.27) expondremos dos ejemplos:

A. Movimiento browniano estándar

Sea $\{X_t, t \geq 0\}$ un movimiento browniano estándar con parámetros infinitesimales $\mu(x) = 0$, $\sigma^2(x) = 1$. Entonces tomemos

$$s(x) = \exp \left\{ -2 \int^x \frac{\mu(\xi)}{\sigma^2(\xi)} d\xi \right\} = 1$$

y para la medida de escala $S(x) = x$, de manera que $u(x)$, la probabilidad de alcanzar a b antes que a a con estado inicial x , es

$$u(x) = \frac{x - a}{b - a}, \quad (2.33)$$

para $a \leq x \leq b$. La densidad de velocidad es

$$m(\xi) = \frac{1}{\sigma^2(\xi)s(\xi)} = 1,$$

y la función de Green, definida en (2.30) para el intervalo $[a, b]$ es

$$G(x, \xi) = \begin{cases} 2 \frac{(x-a)(b-\xi)}{b-a}, & a \leq x \leq \xi \leq b, \\ 2 \frac{(\xi-a)(b-x)}{b-a}, & a \leq \xi \leq x \leq b. \end{cases} \quad (2.34)$$

Entonces un cálculo directo de (2.27) da

$$v(x) = \mathbf{E}[T_{a,b} | X_0 = x] = (x - a)(b - x), \quad (2.35)$$

para $a \leq x \leq b$.

B. Movimiento browniano con tendencia

Si $\{X_t, t \geq 0\}$ es un movimiento browniano con tendencia no cero $\mu(x) = \mu$ y varianza σ^2 , entonces

$$s(x) = \exp \left\{ -2\mu x / \sigma^2 \right\},$$

$$S(x) = A \exp \left\{ -2\mu x / \sigma^2 \right\} + B,$$

donde A y B son constantes y

$$u(x) = \frac{\exp \left\{ -2\mu x / \sigma^2 \right\} - \exp \left\{ -2\mu a / \sigma^2 \right\}}{\exp \left\{ -2\mu b / \sigma^2 \right\} - \exp \left\{ -2\mu a / \sigma^2 \right\}},$$

$$m(\xi) = \frac{1}{\sigma^2(\xi)s(\xi)} = \frac{1}{\sigma^2} \exp \left\{ -2\mu x / \sigma^2 \right\}.$$

2.4. Conclusiones

El objetivo de este capítulo es la definición de los procesos de difusión, al igual que enunciar algunas de sus propiedades con respecto a sus parámetros infinitesimales. Sin embargo, las definiciones obtenidas a través de estos parámetros infinitesimales, como lo son la densidad de escala y la densidad de velocidad, son el móvil principal de este capítulo, porque estas definiciones y algunas propiedades desarrolladas de ellas en este capítulo son importantes en la demostración del teorema 5 del siguiente capítulo.

Capítulo 3

Procesos de difusión con densidad marginal y función de autocorrelación dadas

Los modelos que se basan en procesos estacionarios tipo difusión son modelos que se pueden ajustar por medio de la distribución marginal y la estructura de correlación encontradas en muchas series de tiempo, por ejemplo en finanzas y en turbulencias. En este trabajo estudiamos los modelos de difusión con tendencia lineal, con una distribución marginal conocida y preestablecida, y con coeficientes de difusión que podemos encontrar explícitamente.

El trabajo en este capítulo está basado en el artículo de Bibby, Skovgaard y Sorensen [1]. Básicamente es un desglose de los cálculos "tras bambalinas" de algunos resultados del mencionado artículo.

3.1. Introducción

Consideremos el problema de elegir un modelo en tiempo continuo basado en observaciones en tiempo discreto X_{t_1}, \dots, X_{t_n} . En la elección de un modelo para estas observaciones es conveniente la comprensión de los procesos que son gobernados por el sistema del cual los datos fueron obtenidos. Una forma de hacer la descripción de un sistema es por medio de un número de ecuaciones diferenciales ordinarias, es decir,

$$\frac{dX_t}{dt} = b(X_t), \quad t \geq 0,$$

en el caso de una sola ecuación diferencial ordinaria. Con este modelo se puede hacer una extensión al sumar un término de ruido blanco

$$dX_t = b(X_t)dt + \sqrt{v(X_t)}dW_t, \quad t \geq 0,$$

donde W es un movimiento browniano estándar. Esto introduce incertidumbre al tratar de describir con este modelo al sistema de los datos, lo cual también es resultado de la dependencia entre las observaciones. Así que, al tener una función de tendencia dada por b , mostramos cómo cualquier densidad de probabilidad que satisfaga débilmente las condiciones de regularidad, puede ser tratada como una densidad marginal cuando elegimos convenientemente al coeficiente de difusión v . Este resultado es útil cuando es posible elegir un coeficiente de difusión que se base en los datos.

En muchos casos los métodos en el manejo de los procesos no son entendidos lo suficiente o son muy complicados para describirlos por medio de una función simple de tendencia b , por lo que se tiene que tomar una aproximación obtenida con más datos, si es que es posible. El principal objetivo de este trabajo es proponer un método para elegir un modelo que se base en los datos. Es decir, mostramos cómo construir un modelo para $X = \{X_t, t \geq 0\}$ con una densidad marginal dada por f , es decir que $X_t \sim f$, también donde f es infinitamente divisible y satisface algunas condiciones de regularidad, y con una función de autocorrelación dada por ρ , $\rho(t) = \mathbf{Corr}(X_t, X_{s+t})$, $s, t \geq 0$, y donde $\rho(t)$ pertenece a una clase de funciones de autocorrelación. El modelo usualmente no es markoviano. Las expresiones para f y ρ son elegidas de manera que ellas se ajustan a un histograma de los datos y a la función de autocorrelación empírica, respectivamente.

Por lo tanto, los modelos introducidos en este trabajo se pueden usar para construir modelos estocásticos con volatilidad.

3.2. Construcción de procesos de difusión

En esta sección describimos la construcción de procesos de difusión con una función de autocorrelación exponencial y una distribución marginal específica. La difusión se construye de tal manera que la distribución marginal está concentrada en el conjunto (l, u) , $-\infty \leq l < u \leq \infty$, y tiene una función de densidad preestablecida f con respecto a la medida de Lebesgue en el espacio de estados (l, u) . En el resto de esta sección una densidad de probabilidad f satisface la siguiente condición.

Condición 3 *La densidad de probabilidad f es continua, acotada, estrictamente positiva en (l, u) , cero fuera de (l, u) y tiene varianza finita.*

Consideremos la ecuación diferencial estocástica

$$dX_t = -\theta(X_t - \mu) dt + \sqrt{v(X_t)} dW_t, \quad t \geq 0, \quad (3.1)$$

donde $\theta > 0$, $\mu \in (l, u)$ y v es una función no negativa definida en el conjunto (l, u) .

Lo que queremos hacer es elegir una v de manera que X sea ergódico con densidad invariante igual a la función de densidad dada f . Entonces supongamos

que esto se ha conseguido y que

$$\int_l^u v(x)f(x)dx < \infty. \quad (3.2)$$

Entonces la solución X es un proceso de reversión a la media, y si este proceso es estacionario la función de autocorrelación es $e^{-\theta t}$. El teorema 5, que se demuestra más adelante, dice que si

$$v(x) = \frac{2\theta \int_l^x (\mu - y)f(y)dy}{f(x)} = \frac{2\theta \mu F(x) - 2\theta \int_l^x yf(y)dy}{f(x)}, \quad l < x < u, \quad (3.3)$$

donde F es la función de distribución asociada con la función de densidad f , entonces X es ergódico con función de densidad invariante f , y (3.2) se satisface.

Lema 3 *Supongamos que la esperanza de f es menor o igual a μ , y que v está dada por (3.3). Entonces la función*

$$g(x) = f(x)v(x) \quad (3.4)$$

es estrictamente positiva para toda $l < x < u$, y $\lim_{x \rightarrow l} g(x) = 0$. Si f tiene esperanza igual a μ , entonces $\lim_{x \rightarrow u} g(x) = 0$.

Demostración De (3.3) y (3.4) tenemos que

$$g(x) = 2\theta \int_l^x (\mu - y)f(y)dy$$

y podemos ver que g es estrictamente creciente en (l, μ) y es decreciente en (μ, u) . Entonces por la definición de g tenemos que $\lim_{x \rightarrow l} g(x) = 0$, cuando la esperanza de f es menor a μ , es decir $\mu_f < \mu$, entonces $\lim_{x \rightarrow u} g(x) > 0$ y si $\mu_f = \mu$ tenemos que $\lim_{x \rightarrow u} g(x) = 0$. Por lo tanto $g(x) > 0$ para toda $l < x < u$.

□

Teorema 5 *Supongamos que la densidad de probabilidad f tiene esperanza μ y satisface la condición 3. Entonces se cumple lo siguiente.*

- (i) *La ecuación diferencial estocástica dada por (3.1) y (3.3) tiene una única solución débil markoviana. El coeficiente de difusión es estrictamente positivo para todo $l < x < u$.*
- (ii) *El proceso de difusión X que es solución de (3.1) y (3.3) es ergódico con densidad invariante f .*
- (iii) *La ecuación (3.2) es satisfecha. Además, si $X_0 \sim f$, entonces X es estacionario, $\mathbf{E}(X_{s+t} | X_s = x) = xe^{-\theta t}$, y la función de autocorrelación para X está dada por*

$$\mathbf{Corr}(X_{s+t}, X_s) = e^{-\theta t} \quad (3.5)$$

(iv) Si $-\infty < l$ ó $u < \infty$, entonces la difusión dada por (3.1) y (3.3) es la única difusión ergódica con tendencia $-\theta(x - \mu)$ y con densidad invariante f . Si el espacio de estados es \mathfrak{R} entonces la difusión dada por (3.1) y (3.3) es la única difusión con tendencia $-\theta(x - \mu)$ y con densidad invariante f que satisfaga (3.2).

Demostración (i) Tenemos que f tiene esperanza μ y al usar el lema 3 tenemos que $g(x) > 0$, $x \in (l, u)$. Como f es continua y estrictamente positiva en (l, u) , entonces $v(x) > 0$, $x \in (l, u)$. Por lo tanto $\sqrt{v(x)} > 0$, $x \in (l, u)$.

Para $x \in (l, u)$, la densidad de escala es

$$s(x) = \exp \left\{ 2\theta \int_{x^*}^x \frac{y - \mu}{v(y)} dy \right\}$$

para algún punto interior $x^* \in (l, u)$. Sea $h'(x) = 2\theta(x - \mu)f(x)/g(x)$, si integramos $h'(x)$ tenemos que $h(x) = 2\theta \int_l^x [(y - \mu)f(y)/g(y)] dy$.

Por otro lado, como $g(x) = 2\theta \int_l^x (\mu - y)f(y)dy$, entonces

$$h'(x) = \frac{-2\theta(\mu - x)f(x)}{g(x)} = -\frac{g'(x)}{g(x)}.$$

Además sabemos que $\int_l^x [g'(y)/g(y)] dy = \ln g(y)|_l^x$, si $g(y) > 0$, $y \in (l, u)$. Entonces $h(x) = -\ln g(x)$, $x \in (l, u)$. Si integramos sobre (x, x^*) , $x \in (l, u)$, tenemos por un lado que

$$h(y)|_{x^*}^x = 2\theta \int_{x^*}^x \frac{(y - \mu)f(y)}{g(y)} dy = 2\theta \int \frac{(y - \mu)}{v(y)} dy$$

y por otro,

$$-\ln g(y)|_{x^*}^x = -\ln g(x) + \ln g(x^*) = \ln \frac{g(x^*)}{g(x)}.$$

Entonces $2\theta \int_{x^*}^x \frac{(y - \mu)}{v(y)} dy = \ln (g(x^*)/g(x))$, si usamos la función exponencial en ambos lados tenemos que

$$\exp \left\{ 2\theta \int_{x^*}^x [(y - \mu)/v(y)] dy \right\} = \frac{g(x^*)}{g(x)}. \quad (3.6)$$

Entonces la función de escala está dada por,

$$S(x) = \int_{x^*}^x s(y)dy = g(x^*) \int_{x^*}^x \frac{1}{g(y)} dy.$$

La función S es estrictamente creciente, dos veces diferenciable continuamente y mapea (l, u) a \mathfrak{R} .

Si $u < \infty$, se sigue de la condición 3 que existe $k > 0$ tal que $g(x) = 2\theta \int_x^u (y - \mu)f(y)dy \leq k(u - x)$. Entonces

$$S(x) \geq g(x^*) \frac{1}{k} \int_{x^*}^x \frac{1}{u - y} dy = g(x^*) \frac{1}{k} \ln \left(\frac{x^* - u}{x - u} \right).$$

Si x tiende a u entonces $\ln[(x^* - u)/(x - u)]$ tiende a ∞ . Por lo tanto $S(x)$ tiende a ∞ cuando x tiende a u .

Si $l < \infty$, tomamos a $g(x) = 2\theta \int_l^x (\mu - y)f(y)dy$ y como f esta acotada, entonces $g(x) \geq k(x - l)$. Así

$$S(x) \leq g(x^*) \frac{1}{k} \int_{x^*}^x \frac{1}{y - l} dy = g(x^*) \frac{1}{k} \ln \left(\frac{x - l}{x^* - l} \right).$$

Si x tiende a l entonces $\ln[(x - l)/(x^* - l)]$ tiende a $-\infty$. Por lo que $S(x)$ tiende a $-\infty$ si x tiende a l .

Ahora, la ecuación diferencial estocástica

$$dY_t = s(S^{-1}(Y_t)) \sqrt{v(S^{-1}(Y_t))} dW_t \quad (3.7)$$

satisface las condiciones del teorema 2.2 en Engelbert y Schmidt [3] porque la función $s(S^{-1}(x)) \sqrt{v(S^{-1}(x))}$ es continua en \mathfrak{R} .

Por lo tanto tiene una única solución débil markoviana con espacio de estados \mathfrak{R} . Por la fórmula de Itô, el proceso $S(Y_t)$ es solución de (3.1). Sabemos que X_t es una solución de (3.1) y como $S(x)$ es dos veces diferenciable continuamente podemos usar la fórmula de Itô de 1-dimensión. Entonces $S(X_t)$ es solución de (3.7), por la fórmula de Itô. Por lo tanto, hemos demostrado (i).

(ii) Se define a la densidad de velocidad por

$$m(x) = \frac{1}{\sigma^2(x)s(x)} = \frac{f(x)}{g(x^*)},$$

esta última igualdad se obtiene de las expresiones $\sigma^2(x) = v(x) = g(x)/f(x)$ y $s(x) = g(x^*)/g(x)$. Entonces la densidad es proporcional a la densidad de la medida de velocidad. Mientras que por el teorema 18 en Skorokhod [16] podemos concluir que la densidad es única, finita e invariante. Hemos demostrado (ii).

(iii) Tenemos que

$$\begin{aligned} \int_{\mu}^{\infty} g(x) dx &= \int_{\mu}^{\infty} v(x) f(x) dx \\ &= 2\theta \int_{\mu}^{\infty} \int_{\mu}^y (y - \mu) f(y) dy dx \\ &= 2\theta \int_{\mu}^{\infty} (y - \mu) f(y) \int_{\mu}^y dx dy \\ &= 2\theta \int_{\mu}^{\infty} (y - \mu)^2 f(y) dy \\ &< \infty \end{aligned}$$

pues f tiene varianza finita. Ahora calculemos la siguiente esperanza condicional $\mathbf{E}[X(t) | X(0) = x]$. Recordemos que:

$$dX_t = -\theta(X_t - \mu)dt + \sqrt{v(X_t)} dW_t, \quad (3.8)$$

y tomemos en cuenta la siguiente función $h(t, x) = e^{\theta t} x$. Entonces multipliquemos por $e^{\theta t}$ a X_t que se define en (3.8) y tenemos a $Y_t = e^{\theta t} X_t$. Usemos la fórmula de Itô-Doebelin para resolver la ecuación (3.8) y obtenemos que

$$Y_t - Y_0 = \int_0^t \left[h_t(s, X_s) + \theta(\mu - X_s) h_x(s, X_s) + \frac{1}{2} \sigma^2 h_{xx}(s, X_s) \right] ds \\ + \int_0^t \sqrt{v(X_s)} h_x(s, X_s) dW_s,$$

donde h_t y h_x son las primeras derivadas de $h(t, x)$ con respecto a t y a x , respectivamente, y h_{xx} es la segunda derivada con respecto a x . Al desarrollar obtenemos que

$$Y_t - Y_0 = \int_0^t \left[\theta e^{\theta s} X_s + \theta(\mu - X_s) e^{\theta s} + \frac{1}{2} v(X_s) * 0 \right] ds + \int_0^t \sqrt{v(X_s)} e^{\theta s} dW_s \\ = \int_0^t [\theta e^{\theta s} X_s + \theta \mu e^{\theta s} - \theta e^{\theta s} X_s] ds + \int_0^t \sqrt{v(X_s)} e^{\theta s} dW_s \\ = \int_0^t \theta \mu e^{\theta s} ds + \int_0^t \sqrt{v(X_s)} e^{\theta s} dW_s.$$

Por lo tanto, al sustituir Y_t obtenemos,

$$e^{\theta t} X_t = X_0 + \theta \mu \int_0^t e^{\theta s} ds + \int_0^t e^{\theta s} \sqrt{v(X_s)} dW_s.$$

Calculemos la esperanza de la ecuación anterior y como la esperanza de la integral de Itô es cero, debido a que la esperanza de las diferencias dW_t es cero, entonces

$$\mathbf{E} [e^{\theta t} X_t] = \mathbf{E} [X_0] + \frac{\theta \mu}{\theta} (e^{\theta t} - 1),$$

multipliquemos por $e^{-\theta t}$,

$$\mathbf{E} [X_t] = e^{-\theta t} \mathbf{E} [X_0] + \mu(1 - e^{-\theta t}).$$

Como $X_0 = x$, entonces tenemos que

$$\mathbf{E} [X_t | X_0 = x] = x e^{-\theta t} + \mu(1 - e^{-\theta t}).$$

Por la estacionariedad de X podemos afirmar que:

$$\mathbf{E} [X_{s+t} | X_s = x] = x e^{-\theta t} + \mu(1 - e^{-\theta t}). \quad (3.9)$$

Ahora calculemos la función de autocorrelación para X . Sabemos por definición que

$$\mathbf{Corr} (X_{s+t}, X_s) = \frac{\mathbf{E} [X_{s+t} X_s] - \mathbf{E} [X_{s+t}] \mathbf{E} [X_s]}{\sqrt{\mathbf{Var} (X_{s+t})} \sqrt{\mathbf{Var} (X_s)}}$$

Enfoquémonos en el cálculo de $\mathbf{E}[X_{s+t}X_s]$, y usemos la propiedad de la esperanza que dice que la esperanza de X y $\mathbf{E}[\mathbf{E}[X|Y]]$ son las mismas, entonces

$$\begin{aligned}\mathbf{E}[X_{s+t}X_s] &= \mathbf{E}[\mathbf{E}[X_{s+t}X_s|X_s = x]] \\ &= \mathbf{E}[X_s\mathbf{E}[X_{s+t}|X_s = x]].\end{aligned}$$

Usemos la ecuación (3.9)

$$\begin{aligned}\mathbf{E}[X_{s+t}X_s] &= \mathbf{E}[X_s(X_s e^{-\theta t} + \mu(1 - e^{-\theta t}))] \\ &= \mathbf{E}[X_s^2] e^{-\theta t} + \mathbf{E}[X_s]\mu(1 - e^{-\theta t}) \\ &= \mathbf{E}[X_s^2] e^{-\theta t} + (\mathbf{E}[X_s])^2(1 - e^{-\theta t}) \\ &= (\mathbf{E}[X_s^2] - (\mathbf{E}[X_s])^2) e^{-\theta t} + (\mathbf{E}[X_s])^2.\end{aligned}$$

Tomando en cuenta este resultado y que el proceso es estacionario obtenemos la función de autocorrelación

$$\begin{aligned}\mathbf{Corr}(X_{s+t}, X_s) &= \frac{(\mathbf{E}[X_s^2] - (\mathbf{E}[X_s])^2) e^{-\theta t} + (\mathbf{E}[X_s])^2 - \mathbf{E}[X_{s+t}]\mathbf{E}[X_s]}{\sigma^2} \\ &= \frac{(\mathbf{E}[X_s^2] - (\mathbf{E}[X_s])^2)}{\sigma^2} e^{-\theta t} \\ &= \frac{\mathbf{Var}(X_s)}{\sigma^2} e^{-\theta t} \\ &= e^{-\theta t}.\end{aligned}$$

Hemos demostrado (iii).

(iv) Por el inciso (ii) notemos que para una difusión ergódica de la forma (3.1) con densidad invariante f cuando $g(x^*) = K$, tenemos que

$$\begin{aligned}f(x) &= \frac{K}{v(x)} \exp\left(-2\theta \int_{x^*}^x \frac{y - \mu}{v(y)} dy\right), \\ &= Km(x),\end{aligned}$$

para alguna constante positiva K , donde m es la densidad de la medida de velocidad. Entonces tomando en cuenta la ecuación anterior, tenemos que

$$f(x)v(x) = K \exp\left(-2\theta \int_{x^*}^x \frac{y - \mu}{v(y)} dy\right),$$

si tomamos el logaritmo obtenemos que

$$\log g(x) = \log K - 2\theta \int_{x^*}^x \frac{y - \mu}{v(y)} dy,$$

también tenemos que la función $g = fv$ es diferenciable, entonces

$$g'(x) = -2\theta(x - \mu) \frac{g(x)}{v(x)} = -2\theta(x - \mu)f(x),$$

al integrar llegamos a

$$g(x) = \int_l^x 2\theta(\mu - y)f(y)dy + C$$

donde C es alguna constante. Como $g(x) = f(x)v(x)$, al sustituir y desarrollar conseguimos que

$$v(x) = \frac{\int_l^x 2\theta(\mu - y)f(y)dy + C}{f(x)}. \quad (3.10)$$

Ahora, para poder afirmar que $v(x) > 0$, para toda $l < x < u$, es necesario que $C \geq 0$, pues en el lema 3 hemos visto que la integral tiende a cero en la frontera. Si una de las fronteras es finita, necesitamos que $C = 0$ para que la medida de escala $1/fv$ diverja en esa frontera porque la integral en (3.10) tiende a cero en la frontera. Por otro lado, si ambas fronteras son infinitas, (3.1) y (3.10) definen una difusión ergódica con densidad invariante f , para toda $C \geq 0$. Pero observemos que la condición (3.2) sólo se cumple cuando $C = 0$.

□

3.3. Expresiones generales de procesos de difusión

Cuando la densidad invariante pertenece a una *familia exponencial* con una componente lineal en el estadístico canónico del coeficiente de difusión al cuadrado podemos enunciar el siguiente teorema.

Teorema 6 *Consideremos una densidad invariante para un proceso de difusión que pertenece a una familia exponencial de la forma siguiente,*

$$f(x; \xi) = a(\xi)b(x) \exp \{ \xi_1 x + \alpha(\xi') \cdot t(x) \}, \quad (3.11)$$

donde $\xi = (\xi_1, \dots, \xi_p)$ y $\xi' = (0, \xi_2, \dots, \xi_p)$, α y t pueden ser vectores, y " \cdot " puede ser el producto punto. Entonces el coeficiente de difusión al cuadrado está dado por

$$v(x; \xi) = -2\theta \frac{\frac{\partial}{\partial \xi_1} F(x; \xi)}{f(x; \xi)}, \quad l < x < u. \quad (3.12)$$

Demostración Desarrollemos algunos cálculos en la densidad de la familia exponencial,

$$\begin{aligned} f(x; \xi) &= a(\xi)b(x) \exp \{ \xi_1 x + \alpha(\xi) \cdot t(x) \} \\ &= \exp \{ -(-\log a(\xi)) \} b(x) \exp \{ \xi_1 x + \alpha(\xi) \cdot t(x) \} \\ &= \exp \{ \xi_1 x + \alpha(\xi) \cdot t(x) - (-\log a(\xi)) \} b(x) \\ &= \exp \{ \xi_1 x + \alpha(\xi) \cdot t(x) - \psi(\xi) \} b(x) \end{aligned}$$

donde definimos a $\psi(\xi) = -\log a(\xi)$. De acuerdo con la sección 2 en Morris [13], entonces

$$\frac{\partial \psi(\xi)}{\partial \xi_1} = \mu, \quad (3.13)$$

donde μ es la esperanza univariada calculada por $\int_l^x x \frac{\partial F(x; \xi)}{\partial \xi_1}$, y $F(x; \xi)$ es la función de distribución de $f(x; \xi)$.

Entonces realicemos algunos cálculos en el lado izquierdo de la ecuación (3.13),

$$\begin{aligned} \frac{\partial \psi(\xi)}{\partial \xi_1} &= \frac{\partial(-\log a(\xi))}{\partial \xi_1} \\ &= -\frac{1}{a(\xi)} \frac{\partial a(\xi)}{\partial \xi_1}. \end{aligned}$$

Por lo tanto, tenemos que

$$-\frac{1}{a(\xi)} \frac{\partial a(\xi)}{\partial \xi_1} = \mu. \quad (3.14)$$

Por otro lado, calculemos $\frac{\partial F(x; \xi)}{\partial \xi_1}$. Entonces,

$$\begin{aligned} \frac{\partial F(x; \xi)}{\partial \xi_1} &= \frac{\partial}{\partial \xi_1} \int_l^x f(y; \xi) dy \\ &= \frac{\partial}{\partial \xi_1} \int_l^x [a(\xi)b(y) \exp \{ \xi_1 y + \alpha(\xi) \cdot t(y) \}] dy \\ &= \int_l^x \frac{\partial}{\partial \xi_1} [a(\xi)b(y) \exp \{ \xi_1 y + \alpha(\xi) \cdot t(y) \}] dy \\ &= \int_l^x \frac{\partial a(\xi)}{\partial \xi_1} (b(y) \exp \{ \xi_1 y + \alpha(\xi) \cdot t(y) \}) dy \\ &\quad + \int_l^x a(\xi) \frac{\partial}{\partial \xi_1} \exp \{ \xi_1 y + \alpha(\xi) \cdot t(y) \} dy \\ &= \frac{\partial a(\xi)}{\partial \xi_1} \int_l^x a(\xi)b(y) \exp \{ \xi_1 y + \alpha(\xi) \cdot t(y) \} dy \\ &\quad + \int_l^x ya(\xi) \exp \{ \xi_1 y + \alpha(\xi) \cdot t(y) \} dy \end{aligned}$$

$$= \frac{\partial a(\xi)}{\partial \xi_1} F(x; \xi) + \int_l^x y f(y; \xi) dy.$$

Multiplicando por $\frac{-2\theta}{f(x; \xi)}$, obtenemos que

$$-\frac{2\theta \frac{\partial}{\partial \xi_1} F(x; \xi)}{f(x; \xi)} = \frac{-2\theta \frac{1}{a(\xi)} \frac{\partial a(\xi)}{\partial \xi_1} F(x; \xi) - 2\theta \int_l^x y f(y; \xi) dy}{f(x; \xi)}.$$

Sutituyamos por lo obtenido en la ecuación (3.14),

$$-\frac{2\theta \frac{\partial}{\partial \xi_1} F(x; \xi)}{f(x; \xi)} = \frac{2\theta \mu F(x; \xi) - 2\theta \int_l^x y f(y; \xi) dy}{f(x; \xi)},$$

y tomemos en cuenta (3.3), entonces tenemos que

$$-\frac{2\theta \frac{\partial}{\partial \xi_1} F(x; \xi)}{f(x; \xi)} = v(x; \xi),$$

$l < x < u$.

□

El resultado de las transformaciones lineales simples está dado en el siguiente lema del que se sigue que sólo necesitamos considerar distribuciones centradas y estandarizadas.

Lema 4 *Sea $\{X_t, t \geq 0\}$ un proceso de difusión estacionario con tendencia lineal y densidad invariante f . Consideremos la transformación lineal dada por*

$$Y_t = \alpha + \sigma X_t, \quad \sigma > 0, \quad \alpha \in \mathfrak{R}.$$

Entonces

$$v_g(y) = \sigma^2 \cdot v_f\left(\frac{y - \alpha}{\sigma}\right),$$

donde g denota la densidad invariante de $\{Y_t, t \geq 0\}$, y donde v_f y v_g denotan los coeficientes de difusión al cuadrado obtenidos por (3.3) de f y g , respectivamente.

Demostración Sean f y g las densidades de X_t y Y_t , respectivamente. Como

$$Y_t = \alpha + \sigma X_t,$$

$\sigma > 0$ y $\alpha \in \mathfrak{R}$, entonces $X_t = \frac{Y_t - \alpha}{\sigma}$. Así, tenemos que

$$\begin{aligned} G(y) &= \mathbf{P}(Y_t \leq y) \\ &= \mathbf{P}(\alpha + \sigma X_t \leq y) \end{aligned}$$

$$\begin{aligned}
&= \mathbf{P} \left(X_t \leq \frac{y - \alpha}{\sigma} \right) \\
&= F \left(\frac{y - \alpha}{\sigma} \right),
\end{aligned}$$

entonces $g(y) = f((y - \alpha) / \sigma) \frac{1}{\sigma}$. Por lo tanto

$$\begin{aligned}
v_g(y) &= \frac{2\theta \int_l^y (\mu - z)g(z)dz}{g(y)} \\
&= \frac{2\theta \int_l^y (\mu - z) \frac{1}{\sigma} f\left(\frac{z-\alpha}{\sigma}\right) dz}{\frac{1}{\sigma} f\left(\frac{y-\alpha}{\sigma}\right)} \\
&= \frac{2\theta \int_{\frac{l-\alpha}{\sigma}}^{\frac{y-\alpha}{\sigma}} (\mu - \alpha - \sigma u) f(u) \sigma du}{f\left(\frac{y-\alpha}{\sigma}\right)} \\
&= \frac{2\theta \int_{\frac{l-\alpha}{\sigma}}^{\frac{y-\alpha}{\sigma}} \sigma^2 \left(\frac{\mu-\alpha}{\sigma} - u\right) f(u) du}{f\left(\frac{y-\alpha}{\sigma}\right)} \\
&= \sigma^2 \frac{2\theta \int_l^{\frac{y-\alpha}{\sigma}} (\mu' - u) f(u) du}{f\left(\frac{y-\alpha}{\sigma}\right)} \\
&= \sigma^2 v_f \left(\frac{y - \alpha}{\sigma} \right).
\end{aligned}$$

□

En el siguiente ejemplo consideramos una densidad invariante en el eje (l, ∞) , donde $l > -\infty$. En esta situación puede ser más conveniente reescribir la expresión (3.3) de la siguiente forma: sabemos que $1 - F(x)$ también es una función de distribución pero al usarla cambian los límites de integración en $v(x)$ y tenemos que $d(1 - F(x)) = -f(x)$, entonces

$$\begin{aligned}
v(x) &= \frac{2\theta\mu(1 - F(x)) - 2\theta \int_x^\infty yf(y)dy}{-f(x)} \\
&= \frac{-2\theta\mu(1 - F(x)) + 2\theta \int_x^\infty yf(y)dy}{f(x)} \\
&= \frac{-2\theta\mu(1 - F(x)) + 2\theta x(1 - F(x)) - 2\theta x(1 - F(x)) + 2\theta \int_x^\infty yf(y)dy}{f(x)} \\
&= \frac{-2\theta(\mu - x)(1 - F(x)) + 2\theta(-x(1 - F(x)) + \int_x^\infty yf(y)dy)}{f(x)},
\end{aligned}$$

al integrar por partes, tenemos que

$$v(x) = \frac{-2\theta(\mu - x)(1 - F(x)) + 2\theta \int_x^\infty (1 - F(y))dy}{f(x)}.$$

Así tenemos que

$$v(x) = \frac{2\theta \left(\int_x^\infty (1 - F(y)) dy - (\mu - x)(1 - F(x)) \right)}{f(x)}, \quad x > l. \quad (3.15)$$

En el caso de difusiones positivas, es decir $l = 0$, el coeficiente de difusión al cuadrado se puede expresar en términos de la función hazard λ y la función hazard integrada Λ de la siguiente forma: la definición de la tasa hazard es $\lambda(x) = f(x)/(1 - F(x))$ y con esto tenemos la función hazard acumulada pues $\lambda(x) = \Lambda'(x)$ y así $\Lambda(x) = -\log(1 - F(x))$. Entonces tenemos de la expresión (3.15) que

$$\begin{aligned} v(x) &= \frac{2\theta \left(\int_x^\infty (1 - F(y))dy - (\mu - x)(1 - F(x)) \right)}{f(x)} \\ &= \frac{2\theta \left(\frac{1}{1 - F(x)} \int_x^\infty (1 - F(y))dy - (\mu - x) \right)}{\frac{f(x)}{1 - F(x)}} \\ &= \frac{2\theta \left(\exp \{ \log(1 - F(x))^{-1} \} \int_x^\infty \exp \{ \log(1 - F(y)) \} dy - (\mu - x) \right)}{\frac{f(x)}{1 - F(x)}} \\ &= \frac{2\theta \left(e^{\Lambda(x)} \int_x^\infty e^{-\Lambda(y)} dy + x - \mu \right)}{\lambda(x)}. \end{aligned}$$

Por lo tanto,

$$v(x) = \frac{2\theta \left(e^{\Lambda(x)} \int_x^\infty e^{-\Lambda(y)} dy + x - \mu \right)}{\lambda(x)}, \quad x > 0. \quad (3.16)$$

3.4. Modelos de difusión explícitos

Veamos algunos ejemplos de difusiones con una densidad invariante en la línea real completa, es decir $-l = u = \infty$, en la línea real positiva, y con soporte compacto.

■ Distribución Student

En este ejemplo consideramos un proceso de difusión con densidad invariante igual a una distribución $t(\nu)$ con ν grados de libertad, es decir,

$$f(x) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\Gamma\left(\frac{\nu}{2}\right)} \left(1 + \frac{1}{\nu}x^2\right)^{-\frac{\nu+1}{2}}, \quad x \in \mathfrak{R}, \quad \nu > 0.$$

Aquí tomamos $\mu = 0$. Sólo consideramos distribuciones t para las que la varianza existe, entonces suponemos que $\nu > 2$. En este caso

$$\begin{aligned}
 \int_{-\infty}^x yf(y)dy &= \int_{-\infty}^x y \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\nu\pi}\Gamma(\frac{\nu}{2})} (1 + \frac{1}{\nu}y^2)^{-\frac{\nu+1}{2}} dy \\
 &= \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\nu\pi}\Gamma(\frac{\nu}{2})} \int_{-\infty}^x y(1 + \frac{1}{\nu}y^2)^{-\frac{\nu+1}{2}} dy \\
 &= \frac{\nu}{2} \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\nu\pi}\Gamma(\frac{\nu}{2})} \int_{-\infty}^{x^2} u^{-\frac{\nu+1}{2}} du \\
 &= \frac{\nu}{2} \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\nu\pi}\Gamma(\frac{\nu}{2})} \left[\frac{u^{-\frac{\nu+1}{2}+1}}{1 - \frac{\nu+1}{2}} \right]_{-\infty}^{x^2} \\
 &= \frac{\nu}{2} \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\nu\pi}\Gamma(\frac{\nu}{2})} (-2) \left[\frac{(1 + \frac{1}{\nu}y^2)^{-\frac{\nu-1}{2}}}{\nu - 1} \right]_{-\infty}^x \\
 &= -\frac{\nu}{\sqrt{\nu\pi}} \frac{1}{\nu - 1} \left(\frac{1}{\nu} \right)^{-\frac{\nu-1}{2}} \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})} (\nu + x^2)^{-\frac{\nu-1}{2}} \\
 &= -\frac{\nu\nu^{\frac{\nu}{2}}\nu^{-\frac{1}{2}}}{\nu^{\frac{1}{2}}\pi^{\frac{1}{2}}(\nu - 1)} \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})} (\nu + x^2)^{-\frac{\nu-1}{2}} \\
 &= -\frac{\nu^{\frac{\nu}{2}}}{\pi^{\frac{1}{2}}(\nu - 1)} \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})} (\nu + x^2)^{-\frac{\nu-1}{2}}
 \end{aligned}$$

De acuerdo con (3.3) y como $\mu = 0$, tenemos que

$$\begin{aligned}
 v(x) &= \frac{2\theta\mu F(x) - 2\theta \int_{-\infty}^x yf(y)dy}{f(x)} \\
 &= \frac{2\theta \left(-\frac{\nu^{\frac{\nu}{2}}}{\pi^{\frac{1}{2}}(\nu-1)} \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})} (\nu + x^2)^{-\frac{\nu-1}{2}} \right)}{\frac{\nu^{\frac{\nu}{2}+\frac{1}{2}}}{(\nu\pi)^{\frac{1}{2}}} \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})} (\nu + x^2)^{-\frac{\nu+1}{2}}} \\
 &= \frac{2\theta}{(\nu - 1)} (\nu + x^2).
 \end{aligned}$$

Por lo tanto

$$v(x) = \frac{2\theta}{\nu - 1} (\nu + x^2), \quad x \in \mathfrak{R}.$$

La función v también está bien definida para $\nu = 2$ y entonces vemos que se define una difusión ergódica con la distribución t como distribución invariante.

□

■ Distribución Gamma

Consideremos un proceso de difusión con una densidad invariante de la distribución Gamma, es decir,

$$f(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}, \quad x > 0, \quad \alpha > 0, \quad \lambda > 0.$$

Para que la densidad este acotada suponemos que $\alpha \geq 1$. En este caso la esperanza es $\mu = \alpha/\lambda$. La función de distribución está dada por

$$F(x) = \frac{\Gamma(\lambda x; \alpha)}{\Gamma(\alpha)},$$

donde

$$\Gamma(x; \alpha) = \int_0^x y^{\alpha-1} e^{-y} dy \quad (3.17)$$

es una función Gamma incompleta. Ahora resolvamos primero la siguiente integral para la densidad invariante gamma,

$$\begin{aligned} \int_0^x y f(y) dy &= \int_0^x y \frac{\lambda^\alpha}{\Gamma(\alpha)} y^{\alpha-1} e^{-\lambda y} dy \\ &= \frac{1}{\Gamma(\alpha)} \int_0^x (\lambda y)^\alpha e^{-\lambda y} dy \\ &= \frac{1}{\lambda \Gamma(\alpha)} \int_0^{\lambda x} u^\alpha e^{-u} du \\ &= \frac{1}{\lambda \Gamma(\alpha)} \left[-(\lambda x)^\alpha e^{-\lambda x} + \int_0^{\lambda x} \alpha u^{\alpha-1} e^{-u} du \right] \\ &= \frac{\alpha}{\lambda \Gamma(\alpha)} \int_0^{\lambda x} u^{\alpha-1} e^{-u} du - \frac{1}{\lambda \Gamma(\alpha)} (\lambda x)^\alpha e^{-\lambda x} \\ &= \frac{\alpha}{\lambda \Gamma(\alpha)} \Gamma(\lambda x; \alpha) - \frac{x}{\lambda} \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x} \\ &= \frac{\alpha}{\lambda} F(x) - \frac{x}{\lambda} f(x). \end{aligned}$$

De acuerdo con (3.3), tenemos que

$$v(x) = \frac{2\theta \mu F(x) - 2\theta \int_1^x y f(y) dy}{f(x)}$$

$$\begin{aligned}
 &= \frac{2\theta \frac{\alpha}{\lambda} F(x) - 2\theta \frac{\alpha}{\lambda} F(x) + 2\theta \frac{x}{\lambda} f(x)}{f(x)} \\
 &= 2\theta \frac{x}{\lambda}.
 \end{aligned}$$

Por lo tanto,

$$v(x) = \frac{2\theta x}{\lambda}.$$

Este proceso es muy conocido y fue propuesto por Cox, Ingersoll, Jr. y Ross como un modelo para las tasas de interés a corto plazo.

□

El siguiente es un ejemplo de una densidad invariante con soporte compacto.

■ Distribución Beta

Consideremos un proceso de difusión con una densidad invariante correspondiente a la distribución Beta, es decir,

$$f(x) = B(\alpha, \beta)^{-1} x^{\alpha-1} (1-x)^{\beta-1}, \quad 0 < x < 1, \quad \alpha > 0, \quad \beta > 0,$$

donde $B(a, b) = \Gamma(a)\Gamma(b)/\Gamma(a+b)$ es la función Beta. En este caso la función de distribución está dada por

$$F(x) = I_x(\alpha, \beta) = \int_0^x y^{\alpha-1} (1-y)^{\beta-1} dy / B(\alpha, \beta), \quad 0 < x < 1,$$

y la media es $\mu = \alpha / (\alpha + \beta)$. Entonces resolvamos primero la siguiente integral:

$$\begin{aligned}
 \int_0^x y f(y) dy &= \frac{1}{B(\alpha, \beta)} \int_0^x y y^{\alpha-1} (1-y)^{\beta-1} dy \\
 &= \frac{1}{B(\alpha, \beta)} \int_0^x y^\alpha (1-y)^{\beta-1} dy \\
 &= \frac{1}{B(\alpha, \beta)} \int_0^x y^{(\alpha+1)-1} (1-y)^{\beta-1} dy \\
 &= \frac{B(\alpha+1, \beta)}{B(\alpha, \beta)} \int_0^x \frac{1}{B(\alpha+1, \beta)} y^{(\alpha+1)-1} (1-y)^{\beta-1} dy \\
 &= \frac{\alpha}{\alpha + \beta} I_x(\alpha + 1, \beta).
 \end{aligned}$$

De acuerdo con (3.3), tenemos que

$$v(x) = \frac{2\theta \mu F(x) - 2\theta \int_0^x y f(y) dy}{f(x)}$$

$$\begin{aligned}
&= \frac{2\theta \frac{\alpha}{\alpha+\beta} I_x(\alpha, \beta) - 2\theta \frac{\alpha}{\alpha+\beta} I_x(\alpha+1, \beta)}{f(x)} \\
&= \frac{2\theta \frac{\alpha}{\alpha+\beta} [I_x(\alpha, \beta) - I_x(\alpha+1, \beta)]}{f(x)}
\end{aligned}$$

En la página 84 de Jordan [5] se prueba la siguiente igualdad con respecto a la función beta incompleta: $I_x(\alpha, \beta) - I_x(\alpha+1, \beta) = \frac{x^\alpha(1-x)^b}{aB(a,b)}$, entonces sustituyendo esta igualdad tenemos

$$\begin{aligned}
v(x) &= \frac{2\theta \frac{\alpha}{\alpha+\beta} \frac{x^\alpha(1-x)^b}{aB(a,b)}}{f(x)} \\
&= \frac{\frac{2\theta}{\alpha+\beta} x^\alpha(1-x)^b}{B(a,b)f(x)} \\
&= \frac{2\theta x(1-x)}{\alpha+\beta}.
\end{aligned}$$

Entonces el coeficiente de difusión al cuadrado toma la forma

$$v(x) = \frac{2\theta}{\alpha+\beta} x(1-x), \quad 0 < x < 1.$$

□

■ Distribución normal

Consideremos un proceso de difusión con una densidad invariante correspondiente a la distribución normal estándar, es decir

$$f(x) = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}x^2\right\}, \quad -\infty < x < \infty.$$

En este caso la función de distribución está dada por

$$F(x) = \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x \exp\left\{-\frac{1}{2}y^2\right\} dy, \quad -\infty < x < \infty,$$

y la media es $\mu = 0$. Entonces resolvamos primero la siguiente integral:

$$\begin{aligned}
\int_{-\infty}^x y f(y) dy &= \int_{-\infty}^x y \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}y^2\right\} dy \\
&= -\frac{1}{\sqrt{2\pi}} \left[\exp\left\{-\frac{1}{2}y^2\right\} \right]_{-\infty}^x \\
&= -\frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}x^2\right\}.
\end{aligned}$$

De acuerdo con (3.3), tenemos que

$$\begin{aligned} v(x) &= \frac{2\theta\mu F(x) - 2\theta \int_l^x yf(y)dy}{f(x)} \\ &= \frac{2\theta \frac{1}{\sqrt{2\pi}} \exp\{-\frac{1}{2}x^2\}}{\frac{1}{\sqrt{2\pi}} \exp\{-\frac{1}{2}x^2\}} \\ &= 2\theta. \end{aligned}$$

Por lo tanto,

$$v(x) = 2\theta.$$

□

■ **Distribución logística**

Consideremos un proceso de difusión con una densidad invariante correspondiente a la distribución logística, es decir

$$f(x) = \frac{e^x}{(1 + e^x)^2}, \quad -\infty < x < \infty.$$

En este caso la función de distribución está dada por

$$F(x) = \frac{1}{1 + e^x}, \quad -\infty < x < \infty,$$

y la media es $\mu = 0$. Entonces resolvamos primero la siguiente integral:

$$\int_0^x yf(y)dy = \int_{-\infty}^x y \frac{e^y}{(1 + e^y)^2} dy$$

integrando por partes

$$\int_0^x yf(y)dy = 2y \frac{e^y}{1 + e^y} \Big|_{-\infty}^x - \int_{-\infty}^x 2 \frac{e^y}{1 + e^y} dy$$

y haciendo un cambio de variable $u = e^y$ obtenemos

$$\int_0^x yf(y)dy = 2x \frac{e^x}{1 + e^x} - 2 \log(1 + e^x).$$

De acuerdo con (3.3), tenemos que

$$\begin{aligned} v(x) &= \frac{2\theta\mu F(x) - 2\theta \int_l^x yf(y)dy}{f(x)} \\ &= \frac{-2\theta \int_{-\infty}^x yf(y)dy}{\frac{e^x}{(1+e^x)^2}} \end{aligned}$$

$$\begin{aligned}
&= -2\theta \frac{(1+e^x)^2}{e^x} - 2 \left[x \frac{e^x}{1+e^x} - \log(1+e^x) \right] \\
&= 4\theta \left[e^{-x}(1+e^x)^2 \log(1+e^x) - x(1+e^x) \right].
\end{aligned}$$

Por lo tanto,

$$v(x) = 4\theta \left[e^{-x}(1+e^x)^2 \log(1+e^x) - x(1+e^x) \right].$$

□

■ Distribución Valor extremo

Consideremos un proceso de difusión con una densidad invariante de la distribución valor extremo tipo Gumbel, es decir,

$$f(x) = e^{-x-e^{-x}}, \quad -\infty < x < \infty.$$

En este caso la esperanza es $\mu = \gamma$, que es la constante de Euler-Mascheroni y su valor es aproximado a 0.5772. La función de distribución general de la distribución tipo Gumbel está dada por

$$F(x) = e^{-e^{-(x-\mu)/\beta}}, \quad -\infty < x < \infty$$

en nuestro caso usamos la distribución con parámetros $\mu = 0$ y $\beta = 1$. Es decir,

$$F(x) = e^{-e^{-x}}, \quad -\infty < x < \infty.$$

Ahora resolvamos primero la siguiente integral para la densidad invariante,

$$\int_0^x yf(y)dy = \int_{-\infty}^x y \exp \{-y - e^{-y}\} dy$$

haciendo un cambio de variable $u = e^{-y}$, tenemos

$$\int_0^x yf(y)dy = \int_{e^{-x}}^{\infty} \log(y)e^{-y}dy$$

integrando por partes,

$$\begin{aligned}
\int_0^x yf(y)dy &= -\log(e^{-y}) \exp \{-e^{-y}\} \Big|_{-\infty}^x + \int_{e^{-x}}^{\infty} \frac{1}{y} e^{-y} dy \\
&= x \exp \{-e^{-x}\} + \int_{e^{-x}}^{\infty} \frac{1}{y} e^{-y} dy.
\end{aligned}$$

De acuerdo con (3.3), tenemos que

$$v(x) = \frac{2\theta\mu F(x) - 2\theta \int_0^x yf(y)dy}{f(x)}$$

$$\begin{aligned}
 &= \frac{2\theta\gamma \exp\{-e^{-x}\} + 2\theta\left(-x \exp\{-e^{-x}\} - \int_{e^{-x}}^{\infty} \frac{1}{y} e^{-y} dy\right)}{\exp\{-x - e^{-x}\}} \\
 &= 2\theta \exp\{-e^{-x}\} (\gamma - x - \exp\{e^{-x}\} Ei(-e^{-x})),
 \end{aligned}$$

donde $Ei(x) = -\int_{-x}^{\infty} (e^{-y}/y) dy$, $x < \infty$. Por lo tanto,

$$v(x) = 2\theta \exp\{-e^{-x}\} (\gamma - x - \exp\{e^{-x}\} Ei(-e^{-x})).$$

□

■ Distribución Pareto

Consideremos un proceso de difusión con una densidad invariante de la distribución Pareto, es decir,

$$f(x) = \alpha(1+x)^{-\alpha-1}, \quad 0 < x < \infty, \quad \alpha > 1.$$

En este caso la esperanza es $\mu = (\alpha - 1)^{-1}$, $\alpha > 1$. La función de distribución está dada por

$$F(x) = 1 - (1+x)^{-\alpha}, \quad 0 < x < \infty, \quad \alpha > 1.$$

Ahora resolvamos primero la siguiente integral para la densidad invariante,

$$\int_0^x yf(y)dy = \int_0^x y\alpha(1+y)^{-(\alpha+1)}dy$$

integrando por partes,

$$\begin{aligned}
 \int_0^x yf(y)dy &= \alpha \left[-\frac{y}{\alpha(1+y)^\alpha} \Big|_0^x + \int_0^x \frac{1}{\alpha} (1+y)^{-\alpha} dy \right] \\
 &= \alpha \left[-\frac{x}{\alpha(1+x)^\alpha} - \frac{1}{\alpha(\alpha-1)} (1+x)^{-(\alpha-1)} + \frac{1}{\alpha(\alpha-1)} \right] \\
 &= -\frac{x}{(1+x)^\alpha} - \frac{1}{(\alpha-1)(1+x)^{\alpha-1}} + \frac{1}{\alpha-1}.
 \end{aligned}$$

De acuerdo con (3.3), tenemos que

$$\begin{aligned}
 v(x) &= \frac{2\theta\mu F(x) - 2\theta \int_0^x yf(y)dy}{f(x)} \\
 &= \frac{2\theta \frac{1}{\alpha-1} (1 - (1+x)^{-\alpha}) + 2\theta x(1+x)^{-\alpha} + 2\theta \frac{1}{\alpha-1} ((1+x)^{-(\alpha-1)} - 1)}{\alpha(1+x)^{-(\alpha+1)}} \\
 &= \frac{2\theta \frac{1}{\alpha-1} (1+x)^{-\alpha} 2\theta x(1+x)^\alpha}{\alpha(1-x)^{-\alpha-1}}
 \end{aligned}$$

$$\begin{aligned}
&= \frac{2\theta \frac{1}{\alpha-1} x(1+x)^{-\alpha} + 2\theta x(1+x)^{-\alpha}}{\alpha(1+x)^{-\alpha-1}} \\
&= \frac{2\theta x(1+x) \left(\frac{1}{\alpha-1} + 1 \right)}{\alpha} \\
&= 2\theta \frac{1}{\alpha-1} x(1+x).
\end{aligned}$$

Por lo tanto,

$$v(x) = 2\theta \frac{1}{\alpha-1} x(1+x).$$

□

■ Distribución exponencial

Consideremos un proceso de difusión con una densidad invariante de la distribución exponencial, es decir,

$$f(x) = \lambda e^{-\lambda x}, \quad 0 < x < \infty, \quad \lambda > 0.$$

En este caso la esperanza es $\mu = (\lambda)^{-1}$, $\lambda > 0$. La función de distribución está dada por

$$F(x) = 1 - e^{-\lambda x}, \quad 0 < x < \infty, \quad \lambda > 0.$$

Ahora resolvamos primero la siguiente integral para la densidad invariante,

$$\int_0^x y f(y) dy = \int_0^x y \lambda e^{-\lambda y} dy$$

integrando por partes,

$$\begin{aligned}
\int_0^x y f(y) dy &= -y e^{-\lambda y} \Big|_0^x + \int_0^x e^{-\lambda y} dy \\
&= -x e^{-\lambda x} - \frac{1}{\lambda} e^{-\lambda x} + \frac{1}{\lambda}.
\end{aligned}$$

De acuerdo con (3.3), tenemos que

$$\begin{aligned}
v(x) &= \frac{2\theta \mu F(x) - 2\theta \int_0^x y f(y) dy}{f(x)} \\
&= \frac{2\theta \frac{1}{\lambda} (1 - e^{-\lambda x}) + 2\theta \left(x e^{-\lambda x} + \frac{1}{\lambda} (e^{-\lambda x} - 1) \right)}{\lambda e^{-\lambda x}} \\
&= \frac{2\theta x}{\lambda}.
\end{aligned}$$

Por lo tanto,

$$v(x) = \frac{2\theta x}{\lambda}.$$

□

■ Distribución Chi-cuadrada

Consideremos un proceso de difusión con una densidad invariante de la distribución Chi-cuadrada, es decir,

$$f(x) = \frac{1}{2^{\frac{1}{2}\nu}\Gamma(\frac{\nu}{2})} x^{\frac{\nu}{2}-1} e^{-\frac{1}{2}x}, \quad 0 < x < \infty, \quad \nu \geq 2.$$

En este caso la esperanza es $\mu = \nu$, $\nu \geq 2$. La función de distribución está dada por

$$F(x) = \int_0^x \frac{1}{2^{\frac{1}{2}\nu}\Gamma(\frac{\nu}{2})} y^{\frac{\nu}{2}-1} e^{-\frac{1}{2}y} dy, \quad 0 < x < \infty, \quad \nu \geq 2.$$

Ahora resolvamos primero la siguiente integral para la densidad invariante,

$$\int_0^x yf(y)dy = \int_0^x y \frac{1}{2^{\frac{1}{2}\nu}\Gamma(\frac{\nu}{2})} y^{\frac{\nu}{2}-1} e^{-\frac{1}{2}y} dy$$

integrando por partes,

$$\begin{aligned} \int_0^x yf(y)dy &= \frac{1}{2^{\frac{1}{2}\nu}\Gamma(\frac{\nu}{2})} \left[-2y^{\frac{\nu}{2}} e^{-\frac{1}{2}y} \Big|_0^x + \nu \int_0^x y^{\frac{\nu}{2}-1} e^{-\frac{1}{2}y} dy \right] \\ &= \frac{1}{2^{\frac{1}{2}\nu}\Gamma(\frac{\nu}{2})} \left[-2x^{\frac{\nu}{2}} e^{-\frac{1}{2}x} + \nu \int_0^x y^{\frac{\nu}{2}-1} e^{-\frac{1}{2}y} dy \right] \\ &= -\frac{2}{2^{\frac{1}{2}\nu}\Gamma(\frac{\nu}{2})} x^{\frac{\nu}{2}} e^{-\frac{1}{2}x} + \frac{\nu}{2^{\frac{1}{2}\nu}\Gamma(\frac{\nu}{2})} \int_0^x y^{\frac{\nu}{2}-1} e^{-\frac{1}{2}y} dy. \end{aligned}$$

De acuerdo con (3.3), tenemos que

$$\begin{aligned} v(x) &= \frac{2\theta\mu F(x) - 2\theta \int_l^x yf(y)dy}{f(x)} \\ &= \frac{2\theta\nu \int_0^x \frac{1}{2^{\frac{1}{2}\nu}\Gamma(\frac{\nu}{2})} y^{\frac{\nu}{2}-1} e^{-\frac{1}{2}y} dy + 4\theta \frac{1}{2^{\frac{1}{2}\nu}\Gamma(\frac{\nu}{2})} x^{\frac{\nu}{2}} e^{-\frac{1}{2}x}}{f(x)} \\ &\quad - \frac{2\theta \frac{1}{2^{\frac{1}{2}\nu}\Gamma(\frac{\nu}{2})} \nu \int_0^x y^{\frac{\nu}{2}-1} e^{-\frac{1}{2}y} dy}{f(x)} \\ &= \frac{4\theta \frac{1}{2^{\frac{1}{2}\nu}\Gamma(\frac{\nu}{2})} x^{\nu/2} e^{-\frac{1}{2}x}}{\frac{1}{2^{\frac{1}{2}\nu}\Gamma(\frac{\nu}{2})} x^{\frac{\nu}{2}-1} e^{-\frac{1}{2}x}} \\ &= 4\theta x. \end{aligned}$$

Por lo tanto,

$$v(x) = 4\theta x.$$

□

■ **Distribución Gamma inversa**

Consideremos un proceso de difusión con una densidad invariante de la distribución Gamma inversa, es decir,

$$f(x) = \frac{\delta^\lambda}{\Gamma(\lambda)} x^{-\lambda-1} e^{-\frac{\delta}{x}}, \quad 0 < x < \infty, \quad \delta > 0, \quad \lambda > 1.$$

En este caso la esperanza es $\mu = \delta/(\lambda - 1)$, $\delta > 0$ y $\lambda > 1$. La función de distribución está dada por

$$F(x) = \frac{\delta^{\lambda-1}}{\Gamma(\lambda)} \left(x^{-\lambda+1} e^{-\frac{\delta}{x}} + (\lambda - 1) \int_0^x y^{-\lambda} e^{-\frac{\delta}{y}} dy \right),$$

donde $0 < x < \infty$, $\delta > 0$, y $\lambda > 1$.

De acuerdo con (3.3), tenemos que

$$\begin{aligned} v(x) &= \frac{2\theta\mu F(x) - 2\theta \int_l^x y f(y) dy}{f(x)} \\ &= \frac{2\theta \frac{\delta}{\lambda-1} F(x) - 2\theta \int_0^x y f(y) dy}{f(x)} \\ &= \frac{\frac{2\theta}{\lambda-1} \frac{\delta^\lambda}{\Gamma(\lambda)} \left(x^{-\lambda+1} e^{-\frac{\delta}{x}} + (\lambda - 1) \int_0^x y^{-\lambda} e^{-\frac{\delta}{y}} dy \right) - 2\theta \frac{\delta^\lambda}{\Gamma(\lambda)} \int_0^x y^{-\lambda} e^{-\frac{\delta}{y}} dy}{\frac{\delta^\lambda}{\Gamma(\lambda)} x^{-\lambda-1} e^{-\frac{\delta}{x}}} \\ &= \frac{\frac{2\theta}{\lambda-1} x^{-\lambda+1} e^{-\frac{\delta}{x}} + 2\theta \int_0^x y^{-\lambda} e^{-\frac{\delta}{y}} dy - 2\theta \int_0^x y^{-\lambda} e^{-\frac{\delta}{y}} dy}{x^{-\lambda-1} e^{-\frac{\delta}{x}}} \\ &= \frac{2\theta}{\lambda-1} x^2. \end{aligned}$$

Por lo tanto,

$$v(x) = \frac{2\theta}{\lambda-1} x^2.$$

□

■ **Distribución gaussiana inversa**

Consideremos un proceso de difusión con una densidad invariante de la distribución gaussiana inversa, es decir,

$$f(x) = \sqrt{\frac{\lambda}{2\pi x^3}} e^{-\frac{\lambda(x-\delta)^2}{2\delta^2 x}}, \quad 0 < x < \infty, \quad \delta > 0, \quad \lambda > 0.$$

En este caso la esperanza es $\mu = \delta$, $\delta > 0$. La función de distribución está dada por

$$F(x) = \Phi\left(\sqrt{\frac{\lambda}{x}}\left(\frac{x}{\mu} - 1\right)\right) + e^{\frac{2\lambda}{\mu}}\Phi\left(-\sqrt{\frac{\lambda}{x}}\left(\frac{x}{\mu} + 1\right)\right),$$

donde $0 < x < \infty$, $\delta > 0$, y $\lambda > 0$. Ahora resolvamos primero la siguiente integral para la densidad invariante,

$$\begin{aligned} \int_0^x yf(y)dy &= \int_0^x y \frac{\lambda^{1/2}}{\sqrt{2\pi}y^{3/2}} e^{-\frac{\lambda}{2\mu^2y}(y-\mu)^2} dy \\ &= \mu \int_0^x \frac{y}{\mu} \frac{\lambda^{1/2}}{\sqrt{2\pi}y^{3/2}} e^{-\frac{\lambda}{2\mu^2y}(y-\mu)^2} dy + \mu F(x) - \mu F(x) \\ &= \mu \left[\int_0^x \left(\frac{y}{\mu} + 1\right) \frac{\lambda^{1/2}}{\sqrt{2\pi}y^{3/2}} e^{-\frac{\lambda}{2\mu^2y}(y-\mu)^2} dy + F(x) \right] \end{aligned}$$

al hacer un cambio de variable $u = \sqrt{\frac{\lambda}{y}}\left(\frac{x}{\mu} - 1\right)$, tenemos

$$\begin{aligned} \int_0^x yf(y)dy &= \mu \left[\int_{-\infty}^{\sqrt{\frac{\lambda}{x}}\left(\frac{x}{\mu} - 1\right)} \frac{2}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} dy + F(x) \right] \\ &= \mu 2\Phi\left(\sqrt{\frac{\lambda}{x}}\left(\frac{x}{\mu} - 1\right)\right) - \mu F(x). \end{aligned}$$

De acuerdo con (3.3), tenemos que

$$\begin{aligned} v(x) &= \frac{2\theta\mu F(x) - 2\theta \int_0^x yf(y)dy}{f(x)} \\ &= \frac{4\theta\mu F(x) - 4\theta\mu\Phi\left(\sqrt{\frac{\lambda}{x}}\left(\frac{x}{\mu} - 1\right)\right)}{f(x)} \\ &= \frac{4\theta\mu \left[\Phi\left(\sqrt{\frac{\lambda}{x}}\left(\frac{x}{\mu} - 1\right)\right) + e^{\frac{2\lambda}{\mu}}\Phi\left(-\sqrt{\frac{\lambda}{x}}\left(\frac{x}{\mu} + 1\right)\right) \right] - 4\theta\mu\Phi\left(\sqrt{\frac{\lambda}{x}}\left(\frac{x}{\mu} - 1\right)\right)}{f(x)} \\ &= \frac{4\theta\mu e^{\frac{2\lambda}{\mu}}\Phi\left(-\sqrt{\frac{\lambda}{x}}\left(\frac{x}{\mu} + 1\right)\right)}{f(x)}. \end{aligned}$$

Por lo tanto,

$$v(x) = \frac{4\theta\mu e^{\frac{2\lambda}{\mu}}\Phi\left(-\sqrt{\frac{\lambda}{x}}\left(\frac{x}{\mu} + 1\right)\right)}{f(x)}.$$

□

■ Distribución F

Consideremos un proceso de difusión con una densidad invariante de la distribución F, es decir,

$$f(x) = \frac{\alpha^{\alpha/2} \beta^{\beta/2}}{B\left(\frac{\alpha}{2}, \frac{\beta}{2}\right)} \frac{x^{\alpha/2-1}}{(\beta + \alpha x)^{\frac{\alpha+\beta}{2}}}, \quad 0 < x < \infty, \quad \alpha \geq 2, \quad \beta > 2.$$

En este caso la esperanza es $\mu = \beta/(\beta - 2)$, $\beta > 2$. La función de distribución está dada por

$$F(x) = \frac{1}{B\left(\frac{\alpha}{2}, \frac{\beta}{2}\right)} \int_0^{\frac{\alpha x}{\beta + \alpha x}} y^{\alpha/2-1} (1-y)^{\beta/2-1} dy,$$

donde $0 < x < \infty$, $\alpha \geq 2$, y $\beta > 2$. Ahora resolvamos primero la siguiente integral para la densidad invariante,

$$\int_0^x y f(y) dy = \int_0^x \frac{1}{B\left(\frac{\alpha}{2}, \frac{\beta}{2}\right)} \left(\frac{\alpha y}{\beta + \alpha y}\right)^{\alpha/2} \left(1 - \frac{\alpha y}{\beta + \alpha y}\right)^{\beta/2} dy$$

hacemos un cambio de variable y tenemos

$$\begin{aligned} \int_0^x y f(y) dy &= \frac{1}{B\left(\frac{\alpha}{2}, \frac{\beta}{2}\right)} \int_0^{\frac{\alpha x}{\beta + \alpha x}} \frac{\beta}{\alpha} \frac{y}{1-y} y^{\alpha/2-1} (1-y)^{\beta/2-1} dy \\ &= \frac{1}{B\left(\frac{\alpha}{2}, \frac{\beta}{2}\right)} \frac{\beta}{\alpha} \int_0^{\frac{\alpha x}{\beta + \alpha x}} y^{\alpha/2} (1-y)^{\beta/2-2} dy \end{aligned}$$

al integrar por partes obtenemos

$$\begin{aligned} \int_0^x y f(y) dy &= \frac{1}{B\left(\frac{\alpha}{2}, \frac{\beta}{2}\right)} \frac{\beta}{\alpha} \left[-\frac{2}{\beta-2} y^{\alpha/2} (1-y)^{\beta/2-1} \Big|_0^{\frac{\alpha x}{\beta + \alpha x}} \right. \\ &\quad \left. + \frac{\alpha}{\beta-2} \int_0^{\frac{\alpha x}{\beta + \alpha x}} y^{\alpha/2-1} (1-y)^{\beta/2-1} dy \right] \\ &= \frac{1}{B\left(\frac{\alpha}{2}, \frac{\beta}{2}\right)} \frac{2\beta}{\alpha(\beta-2)} \left(\frac{\alpha x}{\beta + \alpha x}\right)^{\alpha/2} \left(1 - \frac{\alpha x}{\beta + \alpha x}\right)^{\beta/2-1} \\ &\quad + \frac{\beta}{B\left(\frac{\alpha}{2}, \frac{\beta}{2}\right) (\beta-2)} \int_0^{\frac{\alpha x}{\beta + \alpha x}} y^{\alpha/2-1} (1-y)^{\beta/2-1} dy. \end{aligned}$$

De acuerdo con (3.3), tenemos que

$$\begin{aligned}
 v(x) &= \frac{2\theta\mu F(x) - 2\theta \int_0^x y f(y) dy}{f(x)} \\
 &= \left[2\theta \frac{\beta}{\beta-2} \frac{1}{B\left(\frac{\alpha}{2}, \frac{\beta}{2}\right)} \int_0^{\frac{\alpha x}{\beta+\alpha x}} y^{\alpha/2-1} (1-y)^{\beta/2-1} dy \right. \\
 &\quad \left. + 2\theta \frac{2\beta}{\alpha(\beta-2)B\left(\frac{\alpha}{2}, \frac{\beta}{2}\right)} y^{\alpha/2} (1-y)^{\beta/2-1} \right. \\
 &\quad \left. - 2\theta \frac{\beta}{(\beta-2)B\left(\frac{\alpha}{2}, \frac{\beta}{2}\right)} \int_0^{\frac{\alpha x}{\beta+\alpha x}} y^{\alpha/2-1} (1-y)^{\beta/2-1} dy \right] \frac{1}{f(x)} \\
 &= \frac{4\theta\beta}{\alpha(\beta-2)B\left(\frac{\alpha}{2}, \frac{\beta}{2}\right)} \frac{\left(\frac{\alpha x}{\beta+\alpha x}\right)^{\alpha/2} \left(1 - \frac{\alpha x}{\beta+\alpha x}\right)^{\beta/2-1}}{\frac{1}{B\left(\frac{\alpha}{2}, \frac{\beta}{2}\right)} \left(\frac{\alpha x}{\beta+\alpha x}\right)^{\alpha/2} \left(1 - \frac{\alpha x}{\beta+\alpha x}\right)^{\beta/2} x^{-1}} \\
 &= \frac{4\theta\beta}{\alpha(\beta-2)} \frac{x}{1 - \frac{\alpha x}{\beta+\alpha x}} \\
 &= \frac{4\theta}{\alpha(\beta-2)} x(\beta + \alpha x).
 \end{aligned}$$

Por lo tanto,

$$v(x) = \frac{4\theta}{\alpha(\beta-2)} x(\beta + \alpha x).$$

□

■ Distribución log-normal

Consideremos un proceso de difusión con una densidad invariante de la distribución log-normal, es decir,

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2 x}} e^{-\frac{1}{2\sigma^2}(\log x - \delta)^2}, \quad 0 < x < \infty, \quad \sigma^2 > 0.$$

En este caso la esperanza es $\mu = e^{\delta + \frac{1}{2}\sigma^2}$, $\sigma > 0$. La función de distribución está dada por

$$F(x) = \Phi\left(\frac{\log(x) - \mu}{\sigma}\right), \quad 0 < x < \infty, \quad \sigma^2 > 0.$$

Ahora resolvamos primero la siguiente integral para la densidad invariante,

$$\int_0^x y f(y) dy = \int_0^x \frac{y}{\sqrt{2\pi\sigma y}} e^{-\frac{1}{2\sigma^2}(\log y - \delta)^2} dy$$

$$= \int_0^x \frac{1}{\sqrt{2\pi\sigma y}} e^{-\frac{1}{2}\left(\frac{\log y - \delta}{\sigma} - \sigma\right)^2} e^{\delta + \frac{\sigma^2}{2}} dy$$

haciendo un cambio de variable

$$\begin{aligned} \int_0^x y f(y) dy &= \mu \int_{-\infty}^{\frac{\log x - \delta}{\sigma} - \sigma} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy \\ &= \mu \Phi\left(\frac{\log x - \delta}{\sigma} - \sigma\right). \end{aligned}$$

De acuerdo con (3.3), tenemos que

$$\begin{aligned} v(x) &= \frac{2\theta\mu F(x) - 2\theta \int_l^x y f(y) dy}{f(x)} \\ &= \frac{2\theta\mu}{f(x)} \left[\Phi\left(\frac{\log(x) - \mu}{\sigma}\right) - \Phi\left(\frac{\log(x) - \delta}{\sigma} - \sigma\right) \right]. \end{aligned}$$

Por lo tanto,

$$v(x) = \frac{2\theta\mu}{f(x)} \left[\Phi\left(\frac{\log(x) - \mu}{\sigma}\right) - \Phi\left(\frac{\log(x) - \delta}{\sigma} - \sigma\right) \right].$$

□

■ Distribución Weibull

Consideremos un proceso de difusión con una densidad invariante correspondiente a la distribución Weibull, es decir,

$$f(x) = cx^{c-1}e^{-x^c}, \quad x \in (0, \infty), \quad c > 0.$$

En este caso la esperanza es $\mu = \Gamma(c^{-1} + 1)$, $c > 0$. La función de distribución está dada por

$$F(x) = 1 - e^{-x^c}, \quad x \in (0, \infty), \quad c > 0.$$

Resolvamos primero la siguiente integral:

$$\begin{aligned} \int_0^x y f(y) dy &= \int_0^x cy y^{c-1} e^{-y^c} dy \\ &= \int_0^x cy^c e^{-y^c} dy \\ &= \int_0^{x^c} u^{\frac{1}{c}} e^{-u} du \\ &= \int_0^{x^c} u^{(\frac{1}{c}+1)-1} e^{-u} du \end{aligned}$$

$$= \Gamma(x^c; \frac{1}{c} + 1).$$

De acuerdo con (3.3), tenemos que

$$\begin{aligned} v(x) &= \frac{2\theta\mu F(x) - 2\theta \int_0^x yf(y)dy}{f(x)} \\ &= \frac{2\theta\Gamma(\frac{1}{c} + 1)(1 - e^{-x^c}) - 2\theta\Gamma(x^c; \frac{1}{c} + 1)}{f(x)}. \end{aligned}$$

Por lo tanto,

$$v(x) = \frac{2\theta\Gamma(\frac{1}{c} + 1)(1 - e^{-x^c}) - 2\theta\Gamma(x^c; \frac{1}{c} + 1)}{f(x)}, \quad x \in (0, \infty).$$

□

■ **Distribución uniforme**

Consideremos un proceso de difusión con una densidad invariante correspondiente a la distribución uniforme, es decir,

$$f(x) = 1, \quad x \in (0, 1).$$

En este caso la esperanza es $\mu = 1/2$. La función de distribución está dada por

$$F(x) = x, \quad x \in (0, 1).$$

Resolvamos primero la siguiente integral:

$$\int_0^x yf(y)dy = \int_0^x ydy = \left[\frac{y^2}{2} \right]_0^x = \frac{x^2}{2}.$$

De acuerdo con (3.3), tenemos que

$$\begin{aligned} v(x) &= \frac{2\theta\mu F(x) - 2\theta \int_0^x yf(y)dy}{f(x)} \\ &= 2\theta \frac{1}{2}x - 2\theta \frac{x^2}{2} \\ &= \theta x - \theta x^2 \\ &= \theta x(1 - x). \end{aligned}$$

Por lo tanto,

$$v(x) = \theta x(1 - x), \quad x \in (0, 1).$$

□

Cuadro 3.1: El coeficiente de difusión al cuadrado para las distribuciones más comunes

Nombre de la distribución	Función de densidad $f(x)$	Espacio de estados (l, u)	Media μ	Espacio parametral	Coeficiente de difusión $v(x)$
Normal	$\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$	$(-\infty, \infty)$	0	-	2θ
Student	$\frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\nu\pi}\Gamma(\frac{\nu}{2})} (1 + \frac{1}{\nu}x^2)^{-\frac{\nu+1}{2}}$	$(-\infty, \infty)$	0	$\nu > 1$	$\frac{2\theta}{\nu-1}(\nu + x^2)$
Logística	$\frac{e^x}{(1+e^x)^2}$	$(-\infty, \infty)$	0	-	$4\theta \left[e^{-x}(1+e^x)^2 \log(1+e^x) - x(1+e^x) \right]$
Valor extremo	$e^{-x-e^{-x}}$	$(-\infty, \infty)$	γ	-	$2\theta e^{-e^{-x}} \left(\gamma - x + e^{e^{-x}} Ei(-e^{-x}) \right)$
Pareto	$\alpha(1+x)^{-\alpha-1}$	$(0, \infty)$	$\frac{1}{\alpha-1}$	$\alpha > 1$	$2\theta\mu x(1+x)$
Exponencial	$\lambda e^{-\lambda x}$	$(0, \infty)$	$\frac{1}{\lambda}$	$\lambda > 0$	$\frac{2\theta}{\lambda} x$
Gamma	$\frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}$	$(0, \infty)$	$\frac{\alpha}{\lambda}$	$\alpha \geq 1, \lambda > 0$	$\frac{2\theta}{\lambda} x$
χ^2	$\frac{1}{2^{\frac{\nu}{2}} \Gamma(\frac{\nu}{2})} x^{\frac{\nu}{2}-1} e^{-\frac{1}{2}x}$	$(0, \infty)$	ν	$\nu \geq 2$	$4\theta x$
Gamma inversa	$\frac{\delta^\lambda}{\Gamma(\lambda)} x^{-\lambda-1} e^{-\delta/x}$	$(0, \infty)$	$\frac{\delta}{\lambda-1}$	$\delta > 0, \lambda > 1$	$\frac{2\theta}{\lambda-1} x^2$
Gausiana inversa	$\sqrt{\frac{\lambda}{2\pi x^3}} e^{-\frac{\lambda(x-\delta)^2}{2\delta^2 x}}$	$(0, \infty)$	δ	$\lambda > 0, \delta > 0$	$\frac{4\theta\delta}{f(x)} e^{\frac{2\lambda}{\delta}} \Phi \left(-\sqrt{\frac{\lambda}{\delta}} \left(\frac{x}{\delta} + 1 \right) \right)$
F	$\frac{\alpha^{\frac{\alpha}{2}} \beta^{\frac{\beta}{2}}}{B(\frac{\alpha}{2}, \frac{\beta}{2})} \frac{x^{\frac{\alpha}{2}-1}}{(\beta+\alpha x)^{\frac{\alpha+\beta}{2}}}$	$(0, \infty)$	$\frac{\beta}{\beta-2}$	$\alpha \geq 2, \beta > 2$	$\frac{4\theta}{\alpha(\beta-2)} x(\beta + \alpha x)$
Log-normal	$\frac{1}{\sqrt{2\pi\sigma^2 x}} e^{-\frac{1}{2\sigma^2}(\log x - \delta)^2}$	$(0, \infty)$	$e^{\delta + \frac{1}{2}\sigma^2}$	$\sigma^2 > 0$	$\frac{2\theta\mu}{f(x)} \left(\Phi \left(\frac{\log x - \delta}{\sigma} \right) - \Phi \left(\frac{\log x - \delta}{\sigma} - \sigma \right) \right)$
Weibull	$c x^{c-1} e^{-x^c}$	$(0, \infty)$	$\Gamma \left(\frac{1}{c} + 1 \right)$	$c > 0$	$\frac{2\theta}{f(x)} \left(\Gamma \left(\frac{1}{c} + 1 \right) \left(1 - e^{-x^c} \right) - \Gamma \left(x^c; \frac{1}{c} + 1 \right) \right)$
Uniforme	1	$(0, 1)$	$\frac{1}{2}$	-	$\theta x(1-x)$
Beta	$\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1} (1-x)^{\beta-1}$	$(0, 1)$	$\frac{\alpha}{\alpha+\beta}$	$\alpha > 0, \beta > 0$	$\frac{2\theta}{\alpha+\beta} x(1-x)$

Nota: Las distribuciones Student, Pareto, Gamma inversa y F no tienen varianza finita.

En el cuadro 3.1 presentamos el coeficiente de difusión al cuadrado de las funciones de distribución más comunes. Denotamos por Φ a la función de distribución normal estándar, por $\Gamma(x; \alpha)$ a la función Gamma incompleta dada por (3.17), y por Ei a la función integral exponencial dada por

$$Ei(x) = - \int_{-x}^{\infty} \frac{1}{y} e^{-y} dy, \quad x < 0.$$

Finalmente, γ denota la constante de Euler, $\gamma = 0.57722$.

3.5. Conclusiones

Una vez demostrado el teorema 5, que es uno de los resultados más relevantes de este capítulo, pudimos concretar el cálculo explícito del coeficiente de difusión de la ecuación (3.1) para diferentes distribuciones conocidas que se resumen en el cuadro 3.1. Dentro de esas distribuciones conocidas se calculó el coeficiente de difusión para el caso de la distribución Gamma. Esta distribución la vamos a considerar en el siguiente capítulo para construir un proceso de difusión a partir de probabilidades de transición.

Capítulo 4

Construcción alternativa de un proceso de difusión

Como el título de este capítulo lo dice se estudia una construcción de procesos de difusión. Esta construcción se basa en iteraciones que involucran probabilidades de transición bajo el supuesto de que el proceso tiene una distribución marginal conocida. La construcción expuesta en este capítulo se fundamenta en el artículo de Mena y Walker [12] y se aplica a una base de datos referente a los precios de electricidad de Alberta, Canadá.

4.1. Construcción de difusiones con probabilidades de transición

Lo que se pretende en esta sección es establecer la construcción de un modelo de Markov, $\{X_t, t \geq 0\}$, que tenga una distribución marginal perteneciente a una familia paramétrica $\Pi_X(x)$ dada. Para hacer la aproximación de tal proceso lo que primero haremos es construir las probabilidades de transición haciendo que la marginal sea invariante en el tiempo. Para la construcción de las probabilidades de transición usamos una densidad condicional, $f_{Y|X}(y|x)$, y así obtenemos la densidad de transición dada de la siguiente forma

$$p(x_{n-1}, x_n) = \int f_{X|Y}(x_n|y) f_{Y|X}(y|x_{n-1}) \eta_1(dy), \quad (4.1)$$

donde $f_{X|Y}(x|y) \sim f_{Y|X}(y|x) \pi_X(x)$, donde π_X es una densidad priori que es invariante para la densidad de transición (4.1), es decir,

$$\pi_X(x_n) = \int p(x_{n-1}, x_n) \pi_X(x_{n-1}) \eta_2(dx_{n-1}). \quad (4.2)$$

Ambas medidas, η_1 y η_2 , denotan medidas que se usan en la práctica como son las medidas de Lebesgue o de conteo.

Consideremos una densidad condicional $f_{Y|X}$ tal que la densidad de transición resultante de (4.1) satisfaga las ecuaciones de Chapman-Kolmogorov, esto es,

$$p(x_0, x_{t+s}) = \int p(x_s, x_{t+s})p(x_0, x_s)\eta_2(dx_s). \quad (4.3)$$

Un proceso de Markov construído a través de la probabilidad de transición con densidad dada por (4.1) hereda algunas características de una cadena de Markov generada a través de un algoritmo de muestreo de Gibbs.

Veamos un ejemplo sobre esta construcción.

1. Supongamos que X_0 se distribuye como una **Gamma**(α, λ), para $\alpha, \lambda > 0$.
2. También supongamos que $f_{Y|X}$ es la distribución Poisson. Entonces si $Y_1|X_0 \sim \mathbf{Poisson}(\phi X_0)$, donde $\phi > 0$, por lo tanto $X_1|Y_1 \sim \mathbf{Gamma}(\alpha + Y_1, \lambda + \phi)$, entonces la densidad marginal de X_1 también tiene una densidad **Gamma**(α, λ).
3. Después tomamos $Y_2|X_1 \sim \mathbf{Poisson}(\phi X_1)$ y $X_2|Y_2 \sim \mathbf{Gamma}(\alpha + Y_2, \lambda + \phi)$, y así sucesivamente hasta obtener la muestra deseada.

El parámetro ϕ es el parámetro que controla la correlación del proceso $\{X_t, t \geq 0\}$.

Al tomar en cuenta la ecuación en (4.1) se obtiene la densidad de transición para el proceso $\{X_t, t \geq 0\}$, donde $f_{X|Y}(x_n|y)$ es la función de densidad **Gamma**($x_n; y + \alpha, \phi + \lambda$) y $f_{Y|X}(y|x_{n-1})$ es la función de densidad **Poisson**($y; x_{n-1}\phi$),

$$\begin{aligned} p(x_{n-1}, x_n) &= \sum_{y=0}^{\infty} \frac{(\phi + \lambda)^{y+\alpha}}{\Gamma(y + \alpha)} x_n^{y+\alpha-1} \exp\{-(\phi + \lambda)x_n\} (x_{n-1}\phi)^y \frac{\exp\{-x_{n-1}\phi\}}{y!} \\ &= \sum_{y=0}^{\infty} \exp\{-\phi(x_n + x_{n-1}) - \lambda x_n\} (\phi + \lambda)^\alpha (x_{n-1}x_n)^y x_n^{\alpha-1} \frac{(\phi(\phi + \lambda))^y}{\Gamma(y + \alpha) y!} \\ &= \exp\{-\phi(x_n + x_{n-1}) - \lambda x_n\} (\phi + \lambda)^\alpha x_n^{\alpha-1} \sum_{y=0}^{\infty} \frac{(\phi(\phi + \lambda)x_{n-1}x_n)^y}{\Gamma(y + \alpha) y!} \\ &= \exp\{-\phi(x_n + x_{n-1}) - \lambda x_n\} (\phi + \lambda)^\alpha x_n^{\alpha-1} (x_n x_{n-1} \phi(\phi + \lambda))^{-\frac{\alpha-1}{2}} \\ &\quad \times \sum_{y=0}^{\infty} \frac{(\phi(\phi + \lambda)x_{n-1}x_n)^{y+\frac{\alpha-1}{2}}}{\Gamma(y + \alpha) y!} \\ &= \frac{\exp\{-\phi(x_n + x_{n-1}) - \lambda x_n\}}{(\phi + \lambda)^{-(\alpha+1)/2} \phi^{(\alpha-1)/2}} \end{aligned}$$

$$\times \left(\frac{x_n}{x_{n-1}} \right)^{\frac{\alpha-1}{2}} I_{\alpha-1} \left(2\sqrt{x_n x_{n-1} \phi(\phi + \lambda)} \right), \quad (4.4)$$

donde $I_\nu(\cdot)$ denota la función de Bessel modificada de la primera clase con índice ν .¹

Entonces si se quiere introducir dependencia en tiempo continuo en el modelo anterior podemos dejar que el parámetro ϕ controle la correlación, y que varíe en el tiempo t , denotándolo como ϕ_t . Entonces al encontrar explícitamente la forma de esta función se espera que resulte un proceso, $\{X_t, t \geq 0\}$, que exista y sea de Markov.

El espacio correspondiente del proceso presentado es completo y separable, por lo tanto la existencia de un proceso de Markov, con las distribuciones establecidas, se puede asegurar cuando se cumplen las ecuaciones de Chapman-Kolmogorov.

Con respecto a la función ϕ_t , supongamos que es una función determinista estrictamente positiva, entonces la densidad de transición en tiempo homogéneo que se especificó en (4.4) está dada ahora con la función de densidad **Gamma**($x_t; y + \alpha, \phi_t + \lambda$) y con la función de densidad **Poisson**($y; x_0 \phi_t$), y obtenemos que

$$p(x_0, x_t)$$

$$= \frac{\exp \{ -(\phi_t(x_t + x_0) + \lambda x_t) \}}{(\phi_t + \lambda)^{-(\alpha+1)/2} \phi_t^{(\alpha-1)/2}} \cdot \left(\frac{x_t}{x_0} \right)^{\frac{\alpha-1}{2}} I_{\alpha-1} \left(2\sqrt{x_t x_0 \phi_t(\phi_t + \lambda)} \right), \quad (4.5)$$

donde $I_\nu(\cdot)$ denota la función de Bessel modificada de la primera clase con índice ν .

Entonces para ver cuáles son los valores de ϕ_t con los que la expresión (4.5) satisface las ecuaciones de Chapman-Kolmogorov lo que se usa es la transformada de Laplace en vez de la densidad de transición.

Denotemos la transformada de Laplace de la variable aleatoria Z como $\mathcal{L}_Z(\xi) := \mathbf{E}[\exp\{\xi Z\}]$, donde $\xi \in \mathcal{C}$. Por lo tanto, si Z tiene una distribución **Gamma**(α, λ), entonces $\mathcal{L}_Z(\xi) = (1 - \lambda^{-1}\xi)^{-\alpha}$ y si Z se distribuye como una **Poisson**(η), entonces $\mathcal{L}_Z(\xi) = \exp\{\eta(e^\xi - 1)\}$, donde en nuestro caso $\eta = x_0 \phi_t$.

La transformada de Laplace para la densidad de transición (4.5) se obtiene como sigue:

$$\begin{aligned} \mathcal{L}_{X_t|X_0=x_0}(\xi) &= \mathbf{E}[\mathcal{L}_{X_t|Y_t} | X_0 = x_0] \\ &= \mathbf{E}[\mathbf{E}[\exp\{\xi X_t\} | Y_t] | X_0 = x_0] \end{aligned}$$

¹Función de Bessel modificada de la primera clase con índice ν :
 $I_\nu(x) = \sum_{y=0}^{\infty} \frac{1}{\Gamma(y+\nu+1)y!} \left(\frac{x}{2}\right)^{2y+\nu}$.

$$\begin{aligned}
&= \mathbf{E} \left[\left(1 - (\phi_t + \lambda)^{-1} \xi \right)^{(Y_t + \alpha)} \mid X_0 = x_0 \right] \\
&= \mathbf{E} \left[\left(1 - (\phi_t + \lambda)^{-1} \xi \right)^{-\alpha} \exp \left\{ \ln \left(1 - (\phi_t + \lambda)^{-1} \xi \right)^{-y} \right\} \mid X_0 = x_0 \right] \\
&= \left(1 - (\phi_t + \lambda)^{-1} \xi \right)^{-\alpha} \mathbf{E} \left[\exp \left\{ -y \ln \left(1 - (\phi_t + \lambda)^{-1} \xi \right) \right\} \mid X_0 = x_0 \right] \\
&= \left(1 - (\phi_t + \lambda)^{-1} \xi \right)^{-\alpha} \mathcal{L}_{Y_t \mid X_0} \left(-\ln \left(1 - (\phi_t + \lambda)^{-1} \xi \right) \right) \\
&= \left(1 - (\phi_t + \lambda)^{-1} \xi \right)^{-\alpha} \exp \left\{ x_0 \phi_t \left[\exp \left\{ -\ln \left(1 - (\phi_t + \lambda)^{-1} \xi \right) \right\} - 1 \right] \right\} \\
&= \left(1 - (\phi_t + \lambda)^{-1} \xi \right)^{-\alpha} \exp \left\{ x_0 \phi_t \left(\frac{1}{1 - (\phi_t + \lambda)^{-1} \xi} - 1 \right) \right\} \\
&= \left(1 - (\phi_t + \lambda)^{-1} \xi \right)^{-\alpha} \exp \left\{ x_0 \phi_t \left(\frac{\xi}{\phi_t + \lambda - \xi} \right) \right\}
\end{aligned}$$

Usemos esto para comprobar que se satisfacen las ecuaciones de Chapman-Kolmogorov para cierta función ϕ_t . Entonces

$$\begin{aligned}
\mathbf{E} [\mathcal{L}_{X_{t+s} \mid X_s}(\xi) \mid X_0] &= \left\{ 1 - (\phi_t + \lambda)^{-1} \xi \right\}^{-\alpha} \mathcal{L}_{X_s \mid X_0} \left(\frac{\phi_t \xi}{\phi_t + \lambda - \xi} \right) \\
&= \left\{ 1 - \frac{\xi(\phi_t + \phi_s + \lambda)}{(\phi_t + \lambda)(\phi_s + \lambda)} \right\}^{\alpha} \exp \left\{ \frac{x_0 \xi \phi_t \phi_s}{(\phi_t + \lambda)(\phi_s + \lambda) - \xi(\phi_t + \phi_s + \lambda)} \right\}.
\end{aligned}$$

La cual es igual a $\mathcal{L}_{X_{t+s} \mid X_0}(\xi)$ si y sólo si ϕ satisface la siguiente igualdad

$$\phi_{t+s} = \frac{\phi_t \phi_s}{\phi_t + \phi_s + \lambda}. \quad (4.6)$$

Multipliquemos la ecuación (4.6) por $\frac{\lambda}{\lambda}$ del lado derecho,

$$\begin{aligned}
\phi_{t+s} &= \frac{\lambda}{\lambda} \frac{\phi_t \phi_s}{\phi_t + \phi_s + \lambda} \\
&= \frac{\lambda \phi_t \phi_s}{(\phi_t + \lambda)(\phi_s + \lambda) - \phi_t \phi_s},
\end{aligned}$$

ahora multipliquemos por $\frac{(\phi_t + \lambda)(\phi_s + \lambda) - \phi_t \phi_s}{\phi_t \phi_s}$, entonces tenemos que

$$\begin{aligned}
\lambda &= \phi_{t+s} \frac{(\phi_t + \lambda)(\phi_s + \lambda) - \phi_t \phi_s}{\phi_t \phi_s} \\
&= \phi_{t+s} \left(\frac{(\phi_t + \lambda)(\phi_s + \lambda)}{\phi_t \phi_s} - 1 \right)
\end{aligned}$$

$$= \frac{\phi_{t+s}}{\phi_t \phi_s} (\phi_t + \lambda)(\phi_s + \lambda) - \phi_{t+s}.$$

Por lo que obtenemos

$$\frac{\phi_{t+s} + \lambda}{\phi_{t+s}} = \frac{(\phi_t + \lambda)(\phi_s + \lambda)}{\phi_t \phi_s}. \quad (4.7)$$

Ahora definamos $\varphi_t := (\lambda + \phi_t) / \phi_t$ entonces llegamos a que

$$\varphi_{t+s} = \varphi_t \varphi_s, \quad (4.8)$$

que se conoce como la ecuación de Cauchy y tiene como solución positiva $\varphi_t = e^{\theta t}$. Entonces sustituyendo esto en la definición de φ_t logramos obtener que

$$\phi_t = \frac{\lambda}{e^{\theta t} - 1}, \quad (4.9)$$

donde $\theta > 0$.

4.2. El proceso de difusión Gamma-Poisson

Supongamos que tenemos un proceso de Markov en tiempo homogéneo con densidad de transición $p(x_0, x_t)$, entonces podríamos probar si efectivamente es un proceso de difusión. Esto lo podemos comprobar, como ya lo hemos mencionado, con la condición de Dynkin, y se verifica cuando se satisface que

$$\lim_{t \downarrow 0} \frac{1}{t} \int_{|x_t - x_0| > \epsilon} p(x_0, x_t) dx_t = 0, \quad (4.10)$$

para $\epsilon > 0$. Recordemos que la condición (4.10) esencialmente impide que un proceso tenga saltos instantáneos. Si usamos la desigualdad de Chebyshev y de acuerdo con el lema 2 del capítulo 2 podemos asegurar que (4.10) se satisface si

$$\lim_{t \downarrow 0} \frac{1}{t} \mathbf{E} \left[|X_t - X_0|^h | X_0 = x_0 \right] = 0, \quad (4.11)$$

para $h > 2$. Al satisfacerse esta condición podemos establecer las ecuaciones de diferencias estocásticas con coeficiente de tendencia $\mu(x)$ y coeficiente de difusión $\sigma(x)$ como

$$\mu(x_0) := \lim_{t \downarrow 0} \frac{1}{t} \mathbf{E} [X_t - X_0 | X_0 = x_0], \quad (4.12)$$

$$\sigma(x_0) := \lim_{t \downarrow 0} \frac{1}{t} \mathbf{E} \left[|X_t - X_0|^2 | X_0 = x_0 \right]. \quad (4.13)$$

Para comprobar que existen estos límites para el proceso Gamma-Poisson, primero definimos $E_{x_0}(\cdot) := \mathbf{E}(\cdot | X_0 = x_0)$. Si $Z \sim \mathbf{Gamma}(\alpha, \lambda)$ entonces $\mathbf{E}[Z^j] = (\alpha)_j / \lambda^j$, donde $(\alpha)_j := \alpha(\alpha + 1) \cdots (\alpha + j - 1)$ denota el factorial ascendente. Por lo tanto, para el proceso Gamma-Poisson tenemos

$$E_{x_0} [X_t^j] = \frac{E_{x_0} [(y + \alpha)_j]}{(\lambda + \phi_t)^j}, \quad (4.14)$$

donde la esperanza del lado derecho se obtiene por medio de una distribución **Poisson**($x_0 \phi_t$). Por lo tanto, para verificar la condición (4.11), es suficiente considerar $h = 4$, en cuyo caso obtenemos

$$\begin{aligned} E_{x_0} (|X_t - X_0|^4) &= E_{x_0} (X_t^4) - 4x_0 E_{x_0} (X_t^3) \\ &\quad + 6x_0^2 E_{x_0} (X_t^2) - 4x_0^3 E_{x_0} (X_t) + x_0^4 \\ &= \frac{12x_0^2 \phi_t^2 + \phi_t (12x_0^3 \lambda^2 - x_0^2 \lambda (24 + 24\alpha) + x_0 (24 + 36\alpha + 12\alpha^2))}{(\lambda + \phi_t)^4} \\ &\quad + x_0^4 \lambda^4 - 4x_0^3 \lambda^3 \alpha + x_0^2 \lambda^2 (6\alpha + 6\alpha^2) + x_0 \lambda (8\alpha + 12\alpha^2 - 4\alpha^3) + (\alpha)_4. \end{aligned}$$

Usemos $\phi_t = \lambda (e^{\theta t} - 1)^{-1}$, entonces

$$\lim_{t \downarrow 0} \frac{1}{t} (\lambda + \phi_t)^{-4} = \lim_{t \downarrow 0} \frac{1}{t} \phi_t^2 (\lambda + \phi_t)^{-4} = \lim_{t \downarrow 0} \frac{1}{t} \phi_t (\lambda + \phi_t)^{-4} = 0. \quad (4.15)$$

Por lo tanto, la condición (4.11) se cumple. Análogamente, si aplicamos (4.12) y (4.13), vemos que

$$\mu(x) = \theta \left(\frac{\alpha}{\lambda} - x \right)$$

y

$$\sigma(x) = \sqrt{\frac{2\theta}{\lambda}} x. \quad (4.16)$$

En este caso, la expresión de $\sigma(x)$ podemos verificar que es la que tenemos en el cuadro 3.1 para la distribución Gamma.

Por lo tanto, el proceso Gamma-Poisson se puede ver como la distribución de la solución de una ecuación diferencial estocástica dada por

$$dX_t = \theta \left(\frac{\alpha}{\lambda} - X_t \right) dt + \sqrt{\frac{2\theta}{\lambda}} X_t dW_t, \quad (4.17)$$

donde $\{W_t, t \geq 0\}$ denota un movimiento browniano. Este proceso es una reparametrización del modelo CIR que es ampliamente usado como un modelo de tasas de interés. La forma de la función de tendencia muestra el efecto de la reversión a la media hacia su valor de equilibrio $\frac{\alpha}{\lambda}$, lo cual es precisamente la media de la densidad invariante.

4.3. Resultados

Para ejemplificar el método descrito en este capítulo tomamos observaciones del mercado de electricidad de Alberta, Canadá¹. El centro de mercado de electricidad de Alberta es el System Coordination Centre (SCC), abierto las 24 horas del día, siete días a la semana y soporta las operaciones eléctricas en tiempo real del sistema. Por su parte, en el Energy Trading System (ETS) se vinculan en tiempo real las ofertas de suministro de electricidad y las ofertas de la demanda de los participantes en el mercado, además de que publican esta información de mercado en la página web de Alberta Electric System Operator (AESO). También, el ETS recibe los datos de medición eléctrica y realiza la liquidación financiera y las funciones de facturación para el mercado de electricidad. Los controladores del sistema envían electricidad para satisfacer la demanda, lo cual establece el precio del mercado en cada hora en tiempo real, denominado "precio del pool-eléctrico" para referirse a los precios de electricidad cuya medida son para el caso de Alberta: dólares canadienses por kilowatio-hora. Este precio se publica por hora en los informes AESO y se utiliza para calcular los pagos a proveedores y cobrar a los consumidores.

El mercado de energía de Alberta es un mercado liberalizado desde 1996, y se considera que por esto los precios del pool-eléctrico están determinados por un mercado altamente competitivo. Las importaciones y las exportaciones también influyen en los precios de la electricidad en Alberta.

El periodo de demanda considerado para este ejemplo es del 1 de enero de 1996 hasta el 7 de diciembre de 2003, que son un total de 69,523 observaciones. En la figure 4.1 vemos la gráfica de los precios del pool-eléctrico de todas las observaciones.

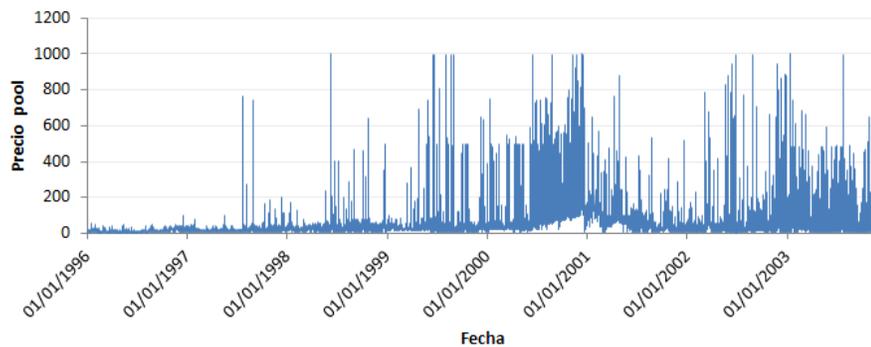


Figura 4.1: Precios del pool-eléctrico de 1996 a 2003 en Alberta, Canadá.

Como podemos apreciar en la figura 4.1 en el año 1996 permanecen los precios por debajo de 100, y es a partir de mediados del año 1997 cuando se presentan los primeros saltos drásticamente grandes, conocidos como *spikes* en

¹Alberta Electric System Operator (AESO), www.aeso.ca.

Cuadro 4.1: Estadísticas de los precios del pool-eléctrico de 1996 a 2003 por año.

Estadística	Precios del pool-eléctrico por año								
	1996	1997	1998	1999	2000	2001	2002	2003	1996 a 2003
Máximo	100.00	765.50	999.50	998.00	999.99	879.20	999.00	999.90	999.99
Percentil 0.75	19.75	23.61	39.73	38.93	146.60	97.76	54.45	68.86	53.97
Media	14.42	20.39	33.03	42.74	133.23	71.29	43.98	64.47	52.86
Mediana	14.30	19.89	27.76	32.26	68.73	53.50	32.88	47.68	31.84
Percentil 0.25	5.53	15.99	21.97	26.71	43.86	35.49	12.92	29.83	19.08
Mínimo	2.74	3.68	5.15	5.76	5.84	5.82	0.01	7.07	0.01
Desviación estándar	8.97	15.96	28.08	70.93	153.54	56.78	64.84	72.05	81.03
Asimetría	0.90	24.58	11.84	8.84	2.12	3.25	7.76	4.37	5.11
Curtosis	1.12	1053.49	266.34	93.32	4.22	20.88	80.64	29.65	33.74
Núm. observaciones	8783	8760	8760	8759	8783	8759	8736	8183	69523

inglés. Al comparar las estadísticas descriptivas de cada año, ver cuadro 4.1, vemos que en el año 2000 es cuando se presenta el máximo valor de todas las observaciones que es de 999.99, el cual se produjo en dos fechas de ese mismo año y a la misma hora, en el 11 y 12 de diciembre a las 18:00 hrs. Por su parte, el mínimo valor que se alcanza es de 0.01, que se produjo el 30 de junio de 2002 a las 05:00 y 07:00 hrs. de ese mismo día.

El año de 1996 es el año que presenta menor variación en los precios del pool-eléctrico en comparación con el resto de los años, ver cuadro 4.1. Debido a esto y a que se tiene el comportamiento de reversión a la media, consideramos estas observaciones de 1996 para estimar los parámetros del modelo Gamma-Poisson, los cuales sí se obtuvieron, ver cuadro 4.4, pero cuando se les hizo la prueba de Dickey-Fuller aumentada no se rechazó la hipótesis alternativa, por lo que las observaciones no son estacionarias, ver cuadro 4.6. También podemos ver en la figura 4.2 el histograma de las observaciones de 1996, donde se comprueba lo que tenemos en el cuadro 4.1 respecto al coeficiente de asimetría, pues presenta asimetría a la derecha. También vemos que la curtosis es leptocúrtica. Aunque tanto la asimetría a la derecha como la curtosis leptocúrtica también se presenta en los otros años, ver cuadro 4.1. Cabe mencionar que estas características de los precios de la electricidad en mercados liberalizados es común encontrarlas, debido a que los precios pueden alcanzar niveles enormes que muchas veces son imprevisibles, y en las siguientes horas volver a los niveles medios en los que se encontraba.

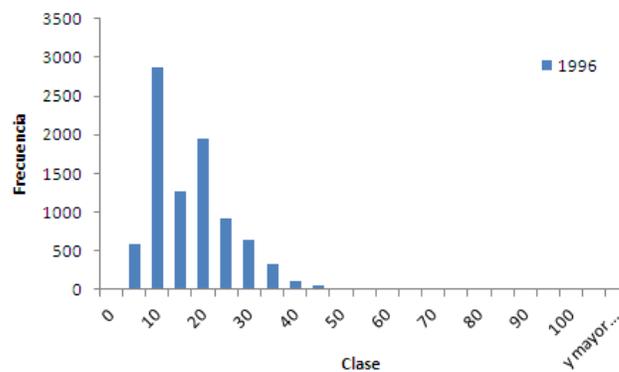


Figura 4.2: Histograma de los precios del pool-eléctrico en 1996.

Se hizo otro tipo de análisis a los precios del pool-eléctrico tomando un corte de los datos. Elegimos todas las observaciones de los días lunes desde 1996 hasta 2003, y después comparamos el comportamiento en cada una de las 24 horas del día. Del mismo modo que se hizo para las submuestras por años, hicimos un análisis descriptivo de los datos por hora, ver cuadros 4.2 y 4.3.

Los precios del pool-eléctrico presentan mayores variaciones en ciertas horas que van desde las 07:00 hasta las 22:00, que son las horas de mayor consumo

Cuadro 4.2: Estadísticas de los precios del pool-eléctrico de los lunes de 1996 a 2003 en cada hora del día.

Estadística	Precios del pool-eléctrico por hora											
	01	02	03	04	05	06	07	08	09	10	11	12
Máximo	433.10	449.40	194.20	177.00	480.80	468.90	499.99	499.99	639.48	639.97	725.39	998.00
Percentil 0.75	33.87	31.99	30.13	29.04	28.42	32.61	44.89	53.77	58.34	61.11	67.50	67.67
Media	28.63	26.63	24.77	23.54	25.03	28.49	39.72	49.30	55.26	61.28	67.77	71.69
Mediana	20.73	17.75	17.01	14.98	15.23	19.52	25.60	32.62	38.44	39.76	43.00	44.08
Percentil 0.25	10.31	9.84	9.69	9.04	8.50	10.13	15.76	20.45	23.20	24.91	27.44	28.05
Mínimo	3.08	3.08	3.05	3.02	2.99	2.74	2.75	1.93	4.63	5.37	5.40	5.42
Desviación estándar	32.89	32.06	24.84	24.36	33.02	34.16	49.06	62.57	65.19	72.74	81.62	94.28
Asimetría	5.62	6.45	2.60	2.61	7.25	6.12	4.56	4.37	4.27	3.79	3.85	4.51
Curtosis	56.88	73.95	9.39	8.78	88.69	67.53	29.40	23.97	25.01	18.53	19.09	29.11
Núm. observaciones	414	414	414	414	414	414	414	414	414	414	414	414

Cuadro 4.3: Estadísticas de los precios del pool-eléctrico de los lunes de 1996 a 2003 en cada hora del día (continuación).

Estadística	Precios del pool-eléctrico por hora												
	13	14	15	16	17	18	19	20	21	22	23	24	Lunes
Máximo	998.00	998.00	998.00	998.00	755.32	999.99	645.20	505.00	612.80	499.99	368.76	281.60	999.99
Percentil 0.75	69.00	69.02	68.90	71.65	79.32	87.68	68.29	62.19	64.72	60.02	44.18	40.05	56.20
Media	74.24	75.42	71.38	71.02	79.98	94.59	72.23	67.41	69.51	62.55	41.06	35.02	54.86
Mediana	43.54	41.47	42.08	41.73	45.58	48.42	40.94	39.97	40.16	38.72	30.94	25.16	33.69
Percentil 0.25	27.93	26.12	25.85	26.21	29.17	31.92	29.77	28.75	27.46	26.15	21.19	16.53	19.36
Mínimo	5.42	5.42	5.42	5.43	5.44	5.44	5.44	5.40	5.38	5.39	5.02	3.08	1.93
Desviación estándar	106.45	110.19	104.20	99.27	104.58	136.64	95.88	84.58	96.57	77.99	36.59	33.32	80.42
Asimetría	4.36	4.17	4.57	4.59	3.27	3.45	3.49	3.34	3.55	3.48	3.26	2.81	4.87
Curtosis	24.04	21.96	26.31	28.32	11.76	13.26	12.95	12.08	12.89	13.51	18.25	11.24	31.67
Núm. observaciones	414	414	414	414	414	414	414	414	414	414	414	414	9936

Cuadro 4.4: Parámetros estimados de los precios de electricidad en Alberta usando el modelo Gamma-Poisson.

SUBMUESTRA	PARÁMETROS ESTIMADOS		
	α	λ	θ
Año 1996	2.8069	0.1940	0.05534
Lunes 01:00	1.6567	0.0576	0.42260
Lunes 06:00	1.5645	0.0544	0.45662
Lunes 08:00	1.5378	0.0310	0.78118
Lunes 23:00	2.2327	0.0542	0.32918

Nota: El método BFGS (Broyden-Fletcher-Goldfarb-Shanno, ver Apéndice A) es el método que se usó para la maximización de la verosimilitud en la construcción Gamma-Poisson.

Cuadro 4.5: Parámetros estimados de tendencia y de difusión de los precios de electricidad en Alberta.

SUBMUESTRA	PARÁMETROS ESTIMADOS	
	$\mu(x)$	$\sigma^2(x)$
Lunes 01:00	$12.155 - 0.423 * x$	$14.674 * x$
Lunes 06:00	$13.132 - 0.457 * x$	$16.788 * x$
Lunes 08:00	$38.752 - 0.781 * x$	$50.399 * x$
Lunes 23:00	$13.560 - 0.329 * x$	$12.147 * x$

durante el día. Se aplicó la prueba de Dickey-Fuller aumentada a cada grupo de observaciones por hora para comprobar si son estacionarias o no, los resultados los podemos ver en el cuadro 4.7, donde sólo seis grupos de observaciones resultaron estacionarias (al 5 por ciento de significancia). Sin embargo, para dos de estos grupos fue imposible lograr la convergencia en el algoritmo del modelo Gamma-Poisson. Al resto de los grupos de observaciones se les logró estimar los parámetros de dicho modelo, ver el cuadro 4.4.

Entonces, para los casos en que los grupos de observaciones resultaron ser estacionarios calculamos sus respectivos parámetros de tendencia y de difusión tomando en cuenta lo obtenido en el cuadro 4.4. Sus respectivas funciones de los parámetros las podemos ver en el cuadro 4.5.

Se intentó ajustar el modelo Gamma-Poisson a los precios del pool-eléctrico en toda la muestra, pero no se consiguió la estimación de los parámetros. Suponemos que esto se debe a que este modelo no es el mejor debido a que a lo largo de las observaciones se presentan valores que son excesivamente mayores

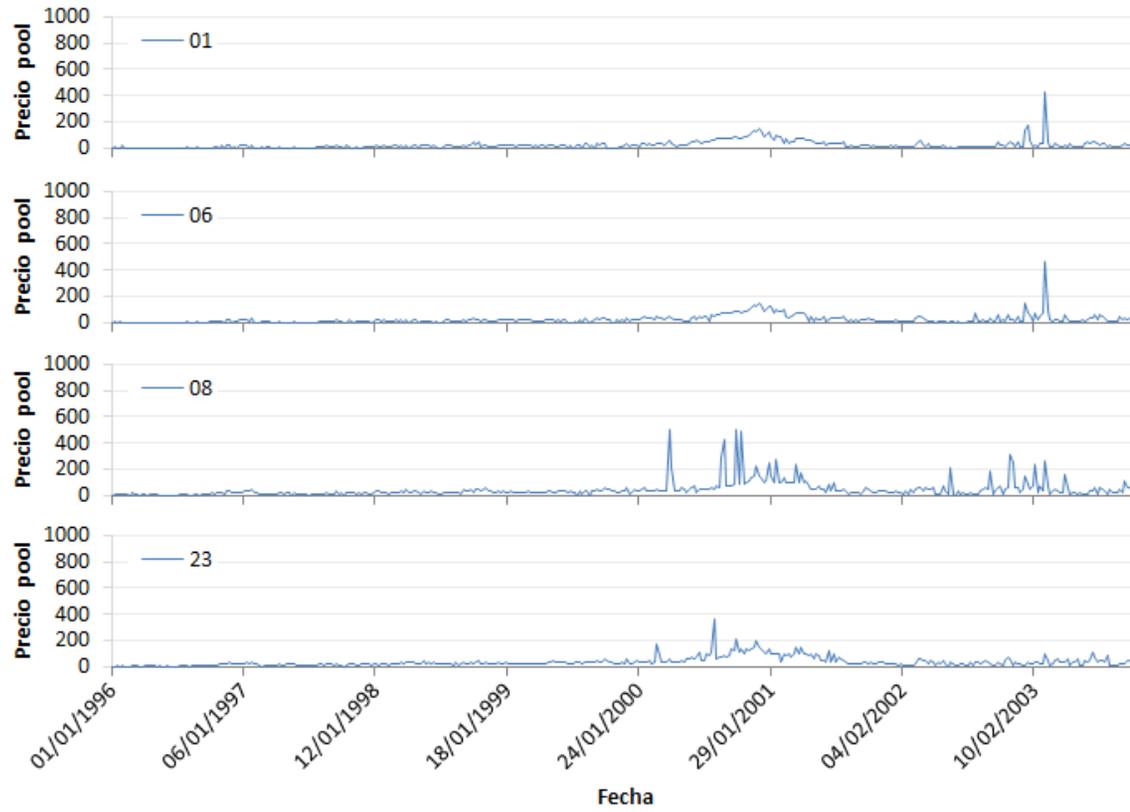


Figura 4.3: Precios del pool-eléctrico de los lunes 1996 a 2003 en ciertas horas del día.

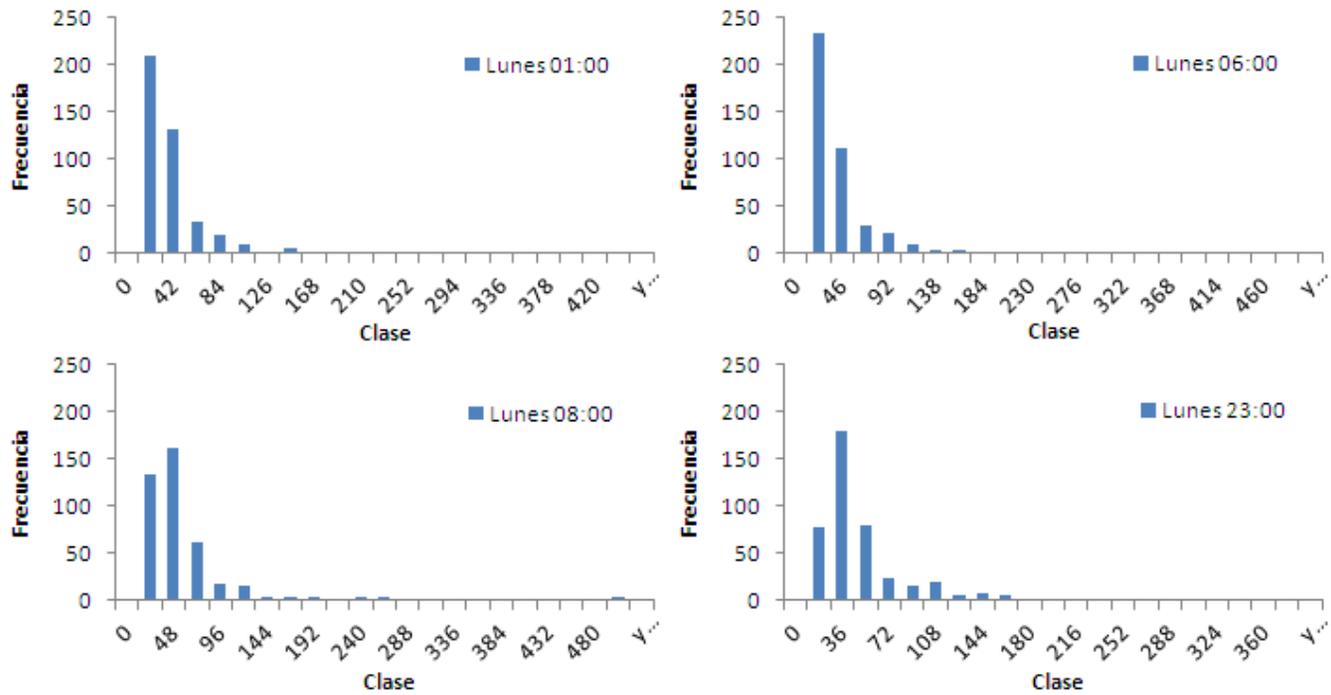


Figura 4.4: Histogramas de los precios del pool-eléctrico de los lunes de 1996 a 2003 en las horas 01:00, 06:00, 08:00 y 23:00.

Cuadro 4.6: Prueba de Dickey-Fuller aumentada, por año.

Año	1996	1997	1998	1999	2000	2001	2002	2003
Estadístico de Dickey-Fuller aumentado	-5.01 [0.000]	-9.74 [0.000]	-9.48 [0.000]	-13.14 [0.000]	-7.34 [0.000]	-5.24 [0.000]	-9.72 [0.000]	-9.7 [0.000]

Nota: ***, ** y * equivalen al 1, 5 y 10 por ciento de significancia, respectivamente. Los valores en corchetes son su respectivo p-valor.

Cuadro 4.7: Prueba de Dickey-Fuller aumentada por hora.

Hora	01	02	03	04	05	06	07	08
Estadístico de Dickey-Fuller aumentado	-2.79** [0.060]	-4.96 [0.000]	-4.08 [0.001]	-3.87 [0.003]	-5.32 [0.000]	-2.67** [0.080]	-4.64 [0.000]	-2.8** [0.059]
Hora (continuación)	09	10	11	12	13	14	15	16
Estadístico de Dickey-Fuller aumentado	-4.34 [0.000]	-4.27 [0.001]	-2.23*** [0.196]	-2.78** [0.063]	-3.09* [0.028]	-4.09 [0.001]	-6.44 [0.000]	-3.51 [0.008]
Hora (continuación)	17	18	19	20	21	22	23	24
Estadístico de Dickey-Fuller aumentado	-4.76 [0.000]	-3.59 [0.006]	-4.98 [0.000]	-6.06 [0.000]	-5.88 [0.000]	-3.89 [0.002]	-1.97*** [0.300]	-3.66 [0.005]

Nota: ***, ** y * equivalen al 1, 5 y 10 por ciento de significancia, respectivamente. Los valores en corchetes son su respectivo p-valor.

a las restantes y que sería mejor usar otra distribución con colas pesadas que esté más acorde con la modelación de este tipo de comportamiento.

Por otro lado, también se intentó estimar los parámetros del modelo Gamma-Poisson al usar el logaritmo natural en toda la muestra, pero tampoco se tuvo éxito. Sin embargo, al tomar las submuestras presentadas arriba se lograron las estimaciones de los parámetros de algunos grupos de observaciones. El siguiente paso, para un trabajo posterior, es encontrar un modelo de difusión con la distribución adecuada para la modelación de los precios del pool-eléctrico en toda la muestra.

4.4. Conclusiones

El objetivo de este capítulo es la construcción de un proceso de difusión con probabilidades de transición a partir de datos reales. Los datos que se tomaron fueron los precios del pool-eléctrico y con estos datos se intentó construir un proceso de difusión Gamma-Poisson. En un primer intento, es decir usando todas las observaciones como se iban presentando hora tras hora desde el 1 de enero de 1996 al 7 de diciembre de 2003, no se logró estimar los parámetros de tendencia y de difusión, porque las observaciones no eran estacionarias. Entonces se hicieron cortes en las observaciones, y en la mayoría de estos cortes, donde sus observaciones resultaron ser estacionarias, se logró estimar los parámetros de tendencia y de difusión. Pero hubieron dos de estos casos, donde sí resultaron ser estacionarias las observaciones, en los que no se logró la estimación de los parámetros infinitesimales, aún no hemos logrado determinar por qué no se logró la estimación. Aunque esto da motivo para intentar usar otras alternativas para estimar los parámetros infinitesimales, ya sea transformando los datos, como puede ser aplicarles algún suavizamiento conveniente debido a la volatilidad que presentan, o buscar otra construcción de proceso de difusión en la que se usen otras distribuciones.

Apéndice A

Método de Broyden-Fletcher-Goldfarb-Shanno

El *Método de Broyden-Fletcher-Goldfarb-Shanno* (BFGS) es un método que se propuso en 1970 por Broyden. Después Fletcher, Goldfarb y Shanno lo retomaron y lo mejoraron. Este es un método que tiene como fin resolver un problema de optimización no lineal sin restricciones. El método BFGS se deriva del método de optimización de Newton, que supone que la función que se quiere optimizar se puede aproximar localmente con una función de segundo grado alrededor del punto óptimo, y para ello hace uso de la primera y segunda derivadas. Su mayor desventaja radica en que se necesita almacenar información en una matriz \mathbf{A} de $N \times N$.

Supongamos que tenemos una función g que se puede aproximar localmente por una forma cuadrática,

$$g(P) \approx c - b \cdot P + \frac{1}{2}P \cdot \mathbf{A} \cdot P.$$

La idea del método es construir una buena aproximación de la matriz Hessiana inversa \mathbf{A}^{-1} , es decir, construir una sucesión de matrices \mathbf{H}_i tales que

$$\lim_{i \rightarrow \infty} H_i = \mathbf{A}^{-1}.$$

Lo ideal es que el límite se alcance en N iteraciones en vez de infinitas iteraciones. Lo que se quiere es encontrar un mínimo al utilizar el método de optimización de Newton para buscar un cero del gradiente de la función. Entonces, alrededor del punto P_i tenemos que

$$g(P) = g(P_i) + (P - P_i) \cdot \nabla g(P_i) + \frac{1}{2}(P - P_i) \cdot \mathbf{A} \cdot (P - P_i),$$

es decir,

$$\nabla g(P) = \nabla g(P_i) + \mathbf{A} \cdot (P - P_i),$$

donde $\nabla g(P)$ es el gradiente (vector de las primeras derivadas parciales).

En los métodos de Newton se toma a $\nabla g(P) = 0$ para determinar el siguiente punto de iteración

$$P - P_i = -\mathbf{A}^{-1} \cdot \nabla g(P_i). \quad (\text{A.1})$$

En el lado izquierdo está la medida finita que necesitamos para llegar al mínimo necesario, mientras que el lado derecho se conoce una vez que se ha acumulado una precisión de $\mathbf{H} \approx \mathbf{A}^{-1}$.

Los métodos quasi-Newton llevan el «quasi» en el nombre porque no se usa la matriz Hessiana actual de g , sino que se usa la aproximación actual de ella. A menudo es mejor usar esta aproximación que la matriz Hessiana verdadera. Ahora, para que la dirección de Newton en la última ecuación sea una dirección descendente, se debe satisfacer

$$\nabla g(P_i) \cdot (P - P_i) = -(P - P_i) \cdot \mathbf{A} \cdot (P - P_i) < 0,$$

lo cual se cumple si \mathbf{A} es definida positiva. La idea detrás de los métodos quasi-Newton es comenzar con una aproximación positiva definida, simétrica a la \mathbf{A} (por lo general la matriz identidad) y construir la aproximación de \mathbf{H}_i de tal forma que la matriz \mathbf{H}_i sigue siendo definida positiva y simétrica.

Entonces, si restamos la ecuación (A.1) a P_{i+1} , tenemos

$$P_{i+1} - P_i = \mathbf{A}^{-1} \cdot (\nabla g_{i+1} - \nabla g_i).$$

Entonces la fórmula que actualiza la aproximación a la matriz \mathbf{H}_i es:

$$\begin{aligned} H_{i+1} = H_i + & \frac{(P_{i+1} - P_i) \otimes (P_{i+1} - P_i)}{(P_{i+1} - P_i) \cdot (\nabla g_{i+1} - \nabla g_i)} \\ & - \frac{[H_i \cdot (\nabla g_{i+1} - \nabla g_i)] \otimes [H_i \cdot (\nabla g_{i+1} - \nabla g_i)]}{(\nabla g_{i+1} - \nabla g_i) \cdot H_i \cdot (\nabla g_{i+1} - \nabla g_i)} \\ & + [(\nabla g_{i+1} - \nabla g_i) \cdot H_i \cdot (\nabla g_{i+1} - \nabla g_i)] u \otimes u, \end{aligned}$$

donde u es el vector

$$u \equiv \frac{(P_{i+1} - P_i)}{(P_{i+1} - P_i) \cdot (\nabla g_{i+1} - \nabla g_i)} - \frac{H_i \cdot (\nabla g_{i+1} - \nabla g_i)}{(\nabla g_{i+1} - \nabla g_i) \cdot H_i \cdot (\nabla g_{i+1} - \nabla g_i)}.$$

Apéndice B

Prueba aumentada de Dickey Fuller

Una serie se dice que es estacionaria si la media y las autocovarianzas de la serie no dependen del tiempo, ver definición 4. Cualquier serie que no es estacionaria se dice que es *no estacionaria*. Un ejemplo común de una serie no estacionaria es la caminata aleatoria:

$$y_t = y_{t-1} + \epsilon_t,$$

donde ϵ es un término aleatorio estacionario. La caminata aleatoria es una serie estacionaria diferenciada pues la primera diferencia de y es estacionaria:

$$y_t - y_{t-1} = (1 - L)y_t = \epsilon_t. \quad (\text{B.1})$$

Una serie estacionaria y diferenciada se dice ser *integrada* y se denota como $I(d)$, donde d es el orden de integración. El orden de integración es el número de raíces unitarias en la serie, o el número de operaciones de diferencias que se hicieron para que la serie sea estacionaria. Para la caminata aleatoria anterior, hay una raíz unitaria, entonces es una serie $I(1)$. Similarmente, una serie estacionaria es $I(0)$. El método formal para probar la estacionariedad de una serie es la prueba de raíz unitaria. EViews tiene una variedad de herramientas para probar si una serie (o las primeras o segundas diferencias de la serie) presenta una raíz unitaria. La siguiente discusión desglosa una prueba de raíz unitaria. Consideremos un proceso AR(1) simple:

$$y_t = \rho y_{t-1} + x_t' \delta + \epsilon_t, \quad (\text{B.2})$$

donde x_t son regresores exógenos opcionales que pueden consistir de constantes, o una constante y tendencia, ρ y δ son parámetros por estimar, y ϵ_t es ruido blanco. Si $|\rho| \geq 1$, y es una serie no estacionaria, la varianza de y aumenta con el tiempo y se aproxima a infinito. Si $|\rho| < 1$, y es una serie estacionaria (en tendencia). Así, la hipótesis de estacionariedad (de tendencia) se puede evaluar

al probar si el valor absoluto de ρ es estrictamente menor que uno. Las pruebas de raíz unitaria que se incluyen en EViews prueban la hipótesis nula $H_0 : \rho = 1$ contra la alternativa de una cola $H_1 : \rho < 1$. En algunos casos, la hipótesis nula se prueba contra una prueba alternativa puntual.

La prueba aumentada de Dickey-Fuller (ADF)

La prueba de Dickey-Fuller estándar estima la ecuación (B.2) después de restar y_{t-1} en ambos lados de la ecuación:

$$\Delta y_t = \alpha y_{t-1} + x_t' \delta + \epsilon_t,$$

donde $\alpha = \rho - 1$. La hipótesis nula y alternativa se pueden escribir como

$$H_0 : \alpha = 0 \text{ vs. } H_1 : \alpha < 0 \quad (\text{B.3})$$

y se evalúa utilizando el radio t convencional para α :

$$t_\alpha = \frac{\hat{\alpha}}{(\sigma(\hat{\alpha}))} \quad (\text{B.4})$$

donde $\hat{\alpha}$ es la estimación de α , y $\sigma(\hat{\alpha})$ es el error estándar del coeficiente. Dickey y Fuller mostraron que bajo la hipótesis nula de raíz unitaria esta estadística no se ajusta a la distribución t de Student, y también obtuvieron resultados asintóticos, además de que simularon los valores críticos de varias pruebas de diversos tamaños de la muestra. La prueba de raíz unitaria Dickey-Fuller estándar descrita anteriormente es válida sólo si la serie es un proceso AR(1). Si la serie se correlaciona en los atrasos de orden superior, la suposición de perturbaciones de ruido blanco ϵ_t es violada. En la prueba aumentada de Dickey-Fuller (ADF) se construye una corrección paramétrica relativa a la correlación de orden superior al suponer que la serie y sigue un proceso AR(p) y suma p términos de diferencias atrasados de la variable dependiente y en el lado derecho de la regresión de prueba:

$$\Delta y_t = \alpha y_{t-1} + x_t' \delta + \beta_1 \Delta y_{t-1} + \beta_2 \Delta y_{t-2} + \cdots + \beta_p \Delta y_{t-p} + v_t. \quad (\text{B.5})$$

Esta especificación aumentada se utiliza para probar (B.3) usando la relación en (B.4). Un resultado obtenido por Fuller es que la distribución asintótica del radio t para α es independiente del número de primeras diferencias atrasadas incluidas en la regresión ADF.

Bibliografía

- [1] Bibby, B. M.; Skovgaard, I. M. y Sorensen, M. (2005) *Diffusion-type Models with given Marginal Distribution and Autocorrelation Function*. Bernoulli 11(2), 191 - 220.
- [2] Billingsley, P. (1995) *Probability and Measure*. Tercera edición. John Wiley & Sons.
- [3] Engelbert, H. J. y Schmidt, W. (1985) *On Solutions of One-Dimensional Stochastic Differential Equations Without Drift*. Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete, 68, 28 - 314.
- [4] Genon-Catalot, V.; Jeantheau, T. y Larédo, C. (2000) *Stochastic Volatility Models as hidden Markov Models and Statistical Applications*. Bernoulli 6(6), 1051 - 1079.
- [5] Jordan, C. (1950) *Calculus of Finite Differences*. Segunda edición. Chelsea Publishing Company.
- [6] Karlin, S. y Taylor, H. M. (1975) *A First Course in Stochastic Processes*. Segunda edición. Academic Press.
- [7] Karlin, S. y Taylor, H. M. (1981) *A Second Course in Stochastic Processes*. Academic Press.
- [8] Kloeden, P. E. y Platen, E. (1992) *Numerical Solution of Stochastic Differential Equations*. Springer.
- [9] Lefebvre, M. (2007) *Applied Stochastic Processes*. Springer.
- [10] Liptser, R. S. y Shiryaev, A. N. (2001) *Statistics of Random Processes. I General Theory*. Springer.
- [11] Mikosch, T. (2004) *Elementary Stochastic Calculus with Finance in View*. World Scientific Publishing Co.
- [12] Mena, R. H. y Walker, S. G. (2009) *On a Construction of Markov Models in Continuous Time*. Metron - International Journal of Statistics, vol. LXVII, issue 3, 303 - 323.

- [13] Morris, C. N. (1982) *Natural Exponential Families with Quadratic Variance Functions*. The Annals of Statistics, vol. 10, No. 1, 65 - 80.
- [14] Press, W. H.; Teukolsky, S. A.; Vetterling, W. T. y Flannery, B. P. (2007) *Numerical Recipes. The Art of Scientific Computing*. Tercera edición. Cambridge University Press.
- [15] Revuz, D. y Yor, M. (1999) *Continuous Martingales and Brownian Motion*. Tercera edición. Springer.
- [16] Skorokhod, A. V. (1989) *Asymptotic Methods in the Theory of Stochastic Differential Equation*. Providence, RI: American Mathematical Society.
- [17] Skorokhod, A. V. (1996) *Lectures on the Theory of Stochastic Processes*. V. S. P. Intl. Science.
- [18] Quantitative Micro Software, LLC. (1994-2009) *EViews 7 User's Guide II*.