



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE CIENCIAS

ESTIMACIÓN PARAMÉTRICA PARA PROCESOS DE
DIFUSIÓN Y APLICACIONES EN FINANZAS

T E S I S

QUE PARA OBTENER EL TÍTULO DE:

MATEMÁTICA APLICADA

P R E S E N T A :

LOLA SERRANO CARRASCO

TUTOR

DR. FERNANDO BALTAZAR LARIOS



CIUDAD DE MÉXICO, 2024



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.

Agradecimientos

Agradezco a mis papás por la red de apoyo e impulso que han sido toda mi vida. A mi mamá, por siempre cuidar mi felicidad y acompañarme en cada paso que doy. A mi papá, por sus porras y ánimos que me hacen creer en mí.

He disfrutado mucho el trayecto hacia la realización de este trabajo. Por todo el aprendizaje y crecimiento que he vivido estoy verdaderamente agradecida con mi asesor, Fernando Baltazar. A Sergio López, agradezco enormemente su guía y acompañamiento a lo largo de mi licenciatura.

Agradezco a mis sinodales Lizbeth Naranjo, Luis Eduardo Pavón, Ramsés Mena y Sergio López. Los comentarios que hacen desde su experiencia y áreas de conocimiento son muy valiosos y enriquecedores.

Índice general

Introducción	4
1. Definiciones y notación	5
1.1. Teoría de probabilidad	5
1.2. Procesos estocásticos	6
1.3. Estimadores máximo verosímiles	8
2. Cálculo estocástico	11
2.1. Construcción integral estocástica	11
2.1.1. Integral estocástica para procesos simples	12
2.1.2. Integral estocástica para procesos continuos	13
2.1.3. Variación y covariación cuadrática	14
2.2. Fórmulas de Itô	15
3. Ecuaciones diferenciales estocásticas y procesos de difusión	19
3.1. Existencia y unicidad de soluciones fuertes y débiles	19
3.2. Propiedades de los procesos de difusión	23
3.3. Ejemplos ecuaciones diferenciales estocásticas	24
3.3.1. Ornstein-Uhlenbeck y Vasicek	24
3.3.2. Movimiento browniano geométrico	26
3.3.3. Puentes brownianos	27
3.3.4. Puentes de difusión	28
4. Simulación de difusiones mediante ecuaciones diferenciales estocásticas	30
4.1. Orden de convergencia	30
4.1.1. Orden fuerte de convergencia	31
4.1.2. Orden débil de convergencia	31
4.2. Método de Euler	31
4.3. Método de Milstein	32
4.4. Transformación de Lamperti	32
4.5. Relación Euler y Milstein	33
4.6. Simulación de algunas difusiones	34
4.6.1. Simulación movimiento browniano	34
4.6.2. Simulación Vasicek	35
4.6.3. Simulación movimiento browniano geométrico	37
4.6.4. Simulación puente browniano	41
4.6.5. Simulación puente de difusión	42
4.6.6. Simulación puente Vasicek	43
4.6.7. Simulación puente browniano geométrico	46

5. Estimación de parámetros	47
5.1. Estimadores basados en verosimilitud	47
5.1.1. Verosimilitud del proceso Vasicek	48
5.1.2. Verosimilitud del movimiento browniano geométrico	54
5.2. Algoritmo EM	59
5.2.1. Algoritmo EMMC	61
5.2.2. Algoritmo EME	62
5.2.3. Algoritmo EME para procesos de difusión	62
6. Aplicaciones de procesos de difusión en finanzas	71
6.1. Proceso de Vasicek y tasas de interés	71
6.1.1. Modelos de tasa corta	71
6.1.2. Modelo de Vasicek	72
6.1.3. Valuación bonos cupón cero	73
6.1.4. Implementación computacional	75
6.1.5. Estimación paramétrica	77
6.2. Movimiento browniano geométrico y valuación de opciones	84
6.2.1. Opciones financieras	84
6.2.2. Valuación libre de arbitraje y Black-Scholes	84
6.2.3. Implementación computacional	87
6.2.4. Estimación paramétrica	88
Conclusiones	92

Introducción

Este trabajo se enfoca en el estudio de procesos de difusión como solución a ecuaciones diferenciales estocásticas. También en su simulación, estimación de sus parámetros y aplicaciones. Como primer objetivo se busca dar un contexto amplio, tanto teórico como computacional sobre este tipo de procesos y su estimación paramétrica.

Las ecuaciones diferenciales estocásticas sirven para modelar sistemas dinámicos aleatorios. Por ello, una de sus principales aplicaciones es en finanzas. Son usadas por ejemplo para modelar el precio de acciones o índices, para la valuación de instrumentos financieros derivados o para modelar tasas de interés. Para aplicar los modelos teóricos es necesario tener datos. Por lo tanto, la falta de información resulta un problema. Un segundo objetivo del trabajo es utilizar el contexto teórico para trabajar con datos de bases reales, tanto de datos completos como incompletos, y poder ajustar modelos de finanzas.

En el Capítulo 1 se enuncian algunas definiciones y notaciones necesarias como base para los temas a tratar en los siguientes capítulos. El Capítulo 2 se enfoca en cálculo estocástico, con el objetivo de construir la integral estocástica. Se presentan algunas propiedades de ésta y finalmente se enuncian las fórmulas de Itô.

Estas fórmulas serán esenciales para el Capítulo 3, enfocado en ecuaciones diferenciales estocásticas y en particular en procesos de difusión, los cuales son definidos dentro del mismo. Se enuncia el teorema de existencia y unicidad, así como propiedades y ejemplos de procesos de difusión. En el Capítulo 4 se trabaja con la simulación de los procesos de difusión presentados en el capítulo anterior. Para ello se utilizan los métodos de aproximación de Euler y Milstein. Los diferentes procesos son simulados y graficados utilizando R.

En el Capítulo 5 se habla sobre estimación paramétrica para procesos de difusión, enfocado en el proceso de Vasicek y movimiento browniano geométrico. Primero, se presentan los estimadores máximo verosímiles de ambos procesos. Estos pueden ser calculados fácilmente ya que la probabilidad de transición de los dos procesos está bien definida. Después se presentan los algoritmos Esperanza-Maximización (EM), EM Monte Carlo (EMMC) y EM estocástico (EME). Con este último también se estiman parámetros para ambos procesos. Los estimadores son implementados en R.

Por último, en el Capítulo 6 se mencionan aplicaciones en finanzas para ambas difusiones. Para el proceso de Vasicek se trabaja con modelos de tasa corta para la valuación de bonos cupón cero. El movimiento browniano geométrico se aplica para la valuación de opciones europeas, utilizando la fórmula de Black y Scholes. Se da una breve introducción teórica sobre las dos aplicaciones. Finalmente se ajusta una base de datos a los procesos correspondientes y se estiman parámetros con los métodos del Capítulo 5.

Capítulo 1

Definiciones y notación

Se enuncian algunas definiciones, resultados y notación de teoría de probabilidad, procesos estocásticos e inferencia estadística. Serán importantes para la construcción de la integral estocástica, las ecuaciones diferenciales estocásticas y la estimación de parámetros. Para consultar detalles de los conceptos presentados en la Sección 1.1 y Sección 1.2 ver [7], [12],[19] y [20]. Para detalles de conceptos abordados en la Sección 1.3 consultar [13] y [26].

1.1. Teoría de probabilidad

Definición 1.1.1. Un *espacio de probabilidad* es una terna $(\Omega, \mathcal{F}, \mathbb{P})$ donde:

- (i) Ω es un conjunto.
- (ii) \mathcal{F} es una σ -álgebra de Ω .
- (iii) $\mathbb{P} : \mathcal{F} \rightarrow [0, 1]$ es una medida de probabilidad.

Para las siguientes definiciones se asume el espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$.

Definición 1.1.2. Una *variable aleatoria* X es una función medible de Ω a \mathbb{R} , es decir:

$$X : \Omega \rightarrow \mathbb{R}, \\ \forall B \in \mathcal{B}(\mathbb{R}), X^{-1}(B) \in \mathcal{F}.$$

$\mathcal{B}(\mathbb{R})$ es la σ -álgebra de Borel en \mathbb{R} .

Definición 1.1.3 (*Espacio $L^p(\mathbb{P})$*). Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad y $1 \leq p < \infty$. La variable aleatoria $X : \Omega \rightarrow \mathbb{R}$ está en L^p si $\mathbb{E}|X|^p < \infty$. La norma en el espacio es

$$\|X\|_p = (\mathbb{E}|X|^p)^{1/p}.$$

En particular, si $p = 2$ entonces L^2 es el conjunto de variables aleatorias cuadrado integrables.

Definición 1.1.4. Una sucesión de variables aleatorias $\{X_n\}_{n \in \mathbb{N}}$ *converge en L^p* a la variable aleatoria X si

$$\lim_{n \rightarrow \infty} \mathbb{E}|X_n - X|^p = 0.$$

Si $p = 2$ se dice que converge en *media cuadrática*.

Observación 1.1.1. Si $1 \leq q < p$ entonces $L^q(\mathbb{P}) \subset L^p(\mathbb{P})$. Por lo tanto, si $X : \Omega \rightarrow \mathbb{R}$

$$\mathbb{E}|X|^2 < \infty \Rightarrow \mathbb{E}|X| < \infty.$$

Observación 1.1.2. El espacio $L^2(\mathbb{P})$ es un espacio completo. Por lo tanto, toda sucesión de Cauchy definida en el espacio converge dentro del mismo.

1.2. Procesos estocásticos

Definición 1.2.1. Un *proceso estocástico* es una colección parametrizada de variables aleatorias $\{X_t : t \in \mathbf{T}\}$. El conjunto \mathbf{T} es llamado *espacio parametral*. Se utilizará $\mathbf{T} = [0, T]$ donde $0 < T \leq \infty$.

Definición 1.2.2. Una *filtración* es una familia creciente de sub- σ -álgebras de \mathcal{F} . La *filtración natural* del proceso estocástico $\{X_t : t \in [0, T]\}$ se define como $\{\mathcal{F}_t : t \geq 0\}$ donde \mathcal{F}_t es la σ -álgebra generada por el proceso hasta el tiempo t , es decir:

$$\mathcal{F}_t = \sigma\{X_s : 0 \leq s \leq t\}.$$

$\{\mathcal{F}_t\}$ se interpreta como la información de $\{X_t : t \in [0, T]\}$ hasta el tiempo t . Las filtraciones usadas a lo largo del texto serán la filtración natural, a menos que se indique lo contrario.

Notación. Para procesos estocásticos se utilizará la notación:

$$\mathbf{X} = \{X_t : t \in [0, T]\} = \{X_t\}.$$

Sin corchetes X_t representa la variable aleatoria que toma el proceso al tiempo t . Se utilizará como filtración la natural con la notación

$$\{\mathcal{F}_t\} = \{\mathcal{F}_t : t \geq 0\}.$$

Definición 1.2.3. Un *proceso de Markov* es un proceso estocástico $\{X_t\}$ adaptado a la filtración $\{\mathcal{F}_t\}$ que satisface la siguiente propiedad:

$$\mathbb{P}(X_{s+t} \in A | \mathcal{F}_s) = \mathbb{P}(X_{s+t} \in A | X_s), \quad \forall s, t \in T, A \in \mathcal{F}.$$

A esta propiedad se le conoce como *propiedad de Markov*. Una definición equivalente es si para cada $0 \leq s < t$ y $f : \mathbb{R} \rightarrow \mathbb{R}$ medible tal que $\mathbb{E}[f(X_t)] < \infty$ se tiene

$$\mathbb{E}[f(X_t) | \mathcal{F}_s] = \mathbb{E}[f(X_t) | X_s].$$

Definición 1.2.4. El proceso $\{X_t\}$ es *gaussiano* si y sólo si para cualesquiera $t_1, \dots, t_k \in [0, T]$ se tiene que $(X_{t_1}, \dots, X_{t_k})$ es una variable aleatoria gaussiana multivariada.

A continuación se definen dos procesos estocásticos necesarios para la construcción de la integral estocástica: las martingalas y el movimiento browniano. Se recuerda la definición de esperanza condicional como variable aleatoria y algunas de sus propiedades. Ésta es fundamental para la definición de los procesos Martingala.

Definición 1.2.5. La *Esperanza Condicional* de la variable aleatoria Y , dada la σ -álgebra \mathcal{F}_t se denota como $\mathbb{E}[Y | \mathcal{F}_t]$. Es la única variable aleatoria que satisface las siguientes propiedades:

(i) $\mathbb{E}[Y|\mathcal{F}_t]$ es \mathcal{F}_t -medible.

(ii) Para cada evento $A \in \mathcal{F}$ que sea \mathcal{F}_t -medible,

$$\mathbb{E}[\mathbb{E}[Y|\mathcal{F}_t]\mathbb{1}_A] = \mathbb{E}[Y\mathbb{1}_A].$$

Proposición 1.2.1. ¹ Sea $\{X_t\}$ un proceso adaptado a la filtración $\{\mathcal{F}_t\}$. La esperanza condicional $\mathbb{E}[Y|\mathcal{F}_t]$ satisface las siguientes propiedades:

- Si la variable Y es \mathcal{F}_t -medible, entonces $\mathbb{E}[Y|\mathcal{F}_t] = Y$.
- Si $A \in \mathcal{F}$ es \mathcal{F}_t -medible, por definición

$$\mathbb{E}[\mathbb{E}[Y|\mathcal{F}_t]\mathbb{1}_A] = \mathbb{E}[Y\mathbb{1}_A].$$

En particular

$$\mathbb{E}[\mathbb{E}[Y|\mathcal{F}_t]] = \mathbb{E}[Y].$$

- Si la variable aleatoria Y es independiente de las variables X_s , para toda $0 \leq s \leq t$ entonces

$$\mathbb{E}[Y|\mathcal{F}_t] = \mathbb{E}[Y].$$

La independencia de Y con X_s , $0 \leq s \leq t$ implica que \mathcal{F}_t no contiene información sobre Y . Por eso es igual condicionar o no la esperanza de Y con respecto a \mathcal{F}_t .

- Si Z es una variable aleatoria \mathcal{F}_t -medible, entonces

$$\mathbb{E}[YZ|\mathcal{F}_t] = Z\mathbb{E}[Y|\mathcal{F}_t].$$

- La esperanza condicional es **lineal**. Si Z es otra variable aleatoria y a, b constantes entonces

$$\mathbb{E}[aY + bZ|\mathcal{F}_t] = a\mathbb{E}[Y|\mathcal{F}_t] + b\mathbb{E}[Z|\mathcal{F}_t].$$

- **Propiedad torre:** Si $s < t$ ($\mathcal{F}_s \subset \mathcal{F}_t$) entonces

$$\mathbb{E}[\mathbb{E}[Y|\mathcal{F}_t]|\mathcal{F}_s] = \mathbb{E}[Y|\mathcal{F}_s] = \mathbb{E}[\mathbb{E}[Y|\mathcal{F}_s]|\mathcal{F}_t].$$

Definición 1.2.6. Un proceso $\{M_t : t \geq 0\} = \{M_t\}$ es una *martingala* con respecto a la filtración $\{\mathcal{F}_t\}$ si

(i) M_t es \mathcal{F}_t -medible.

(ii) Es integrable, *i.e.*, $\mathbb{E}|M_t| < \infty$.

(iii) Para cada $0 \leq s \leq t$ pasa casi seguramente que

$$\mathbb{E}[M_t|\mathcal{F}_s] = M_s.$$

Definición 1.2.7. Un proceso estocástico $\{B_t : t \geq 0\} = \{B_t\}$ es un *movimiento browniano* que comienza casi seguramente (c.s.) en $x \in \mathbb{R}$, con media m y varianza σ^2 si cumple con las siguientes propiedades:

¹La prueba de la proposición se puede consultar en [7] pp. 205-2013 y [12] pp. 3-7.

(i) Las trayectorias $t \mapsto B_t$ son funciones continuas de t casi seguramente.

(ii) Tiene incrementos estacionarios

$$\forall 0 \leq s \leq t, B_t - B_s \sim \mathcal{N}(m(t-s), \sigma^2(t-s)).$$

(iii) Tiene incrementos independientes

$$s < t, (B_t - B_s) \perp B_r, \forall r \leq s.$$

Si el proceso tiene incrementos estacionarios e independientes y comienza en 0 c.s. el proceso es un *movimiento browniano estándar*. A menos que se indique lo contrario, al hablar de movimiento browniano se estará hablando del estándar.

Proposición 1.2.2. ² Algunas propiedades del movimiento browniano son:

- Es una martingala continua.
- Es un proceso de Markov.
- Es un proceso gaussiano.

Definición 1.2.8. Sean $\mathbf{X} = \{X_t\}$ e $\mathbf{Y} = \{Y_t\}$ dos procesos estocásticos definidos en $(\Omega, \mathcal{F}, \mathbb{P})$. El proceso \mathbf{X} es una modificación de \mathbf{Y} si

$$\mathbb{P}(X_t = Y_t) = 1 \quad \forall t \in [0, T]. \quad (1.2.1)$$

Los procesos son *indistinguibles* si para casi todo $\omega \in \Omega$ se tiene $X_t = Y_t$, para todo $t \in [0, T]$. Es decir

$$\mathbb{P}(X_t = Y_t \quad \forall t \in [0, T]) = 1. \quad (1.2.2)$$

1.3. Estimadores máximo verosímiles

Definición 1.3.1. Una *muestra aleatoria* es un conjunto de variables aleatorias X_1, X_2, \dots, X_n independientes e idénticamente distribuidas. La *función de densidad conjunta* de la muestra está dada por

$$f_{X_1, X_2, \dots, X_n}(x_1, x_2, \dots, x_n; \theta) = \prod_{i=1}^n f_{X_i}(x_i; \theta).$$

Definición 1.3.2. El conjunto de valores que puede tomar la muestra aleatoria X_1, \dots, X_n se llama *espacio muestral* y se denota como \mathcal{X} . El conjunto de valores donde puede tomar valores θ , el vector de parámetros, se define como *espacio parametral* denotado como $\Theta \subset \mathbb{R}^p$.

Definición 1.3.3. Una *estadística* es una función de la muestra aleatoria que no depende de parámetros desconocidos. Se denota como $T(\underline{X}) = T(X_1, \dots, X_n)$. Un *estimador* es una estadística $T(\underline{X}) = \hat{\theta}$ utilizada para aproximar los valores de θ .

²La prueba de la proposición se puede consultar en [12] Capítulo 2.6 y [10] Capítulo 1.

Definición 1.3.4. Sea X_1, \dots, X_n una muestra aleatoria proveniente de una distribución con densidad $f(x; \theta)$. La *función de verosimilitud* es la función de densidad conjunta de la muestra, denotada como $L(\theta)$ o $L(\theta|\underline{x})$.

$$L(\theta) = f_{X_1, \dots, X_n}(x_1, \dots, x_n; \theta) = \prod_{i=1}^n f_{X_i}(x_i; \theta).$$

Se define a la *log-verosimilitud* de $f(x; \theta)$ como

$$l(\theta) = \ln(L(\theta)) = \sum_{i=1}^n \ln(f_{X_i}(x_i; \theta)).$$

Definición 1.3.5. Sea X_1, \dots, X_n una muestra aleatoria de una distribución con densidad $f(x; \theta)$ y verosimilitud $L(\theta)$. Se dice que $\hat{\theta} = T(\underline{X})$ es un *estimador máximo verosímil* (MLE, por sus siglas en inglés) de θ si para cualquier $\theta \in \Theta$ se cumple que $L(\hat{\theta}) \geq L(\theta)$. Es decir

$$\hat{\theta} = \arg \max_{\theta \in \Theta} L(\theta).$$

Observación 1.3.1. Dado que el logaritmo natural es una función monótona se tiene que tanto la verosimilitud como la log-verosimilitud alcanzan su máximo en el mismo valor $\hat{\theta}$. Por lo tanto

$$\hat{\theta} = \arg \max_{\theta \in \Theta} L(\theta) = \arg \max_{\theta \in \Theta} l(\theta).$$

Observación 1.3.2. Los estimadores máximo verosímiles pueden no ser únicos. Bajo la condición de un modelo identificable, es decir que para todo $\theta \in \Theta$ la función $\theta \mapsto L(\theta)$ es inyectiva, se tiene que el estimador máximo verosímil es único.

Definición 1.3.6. La *función score*, denotada $s(\theta)$ es el gradiente de la log-verosimilitud.

$$s(\theta) = \frac{\partial l(\theta)}{\partial \theta} = \left(\frac{\partial l(\theta)}{\partial \theta_1}, \frac{\partial l(\theta)}{\partial \theta_2}, \dots, \frac{\partial l(\theta)}{\partial \theta_p} \right).$$

La *matriz Hessiana*, denotada $H(\theta)$ tiene como entradas las segundas derivadas de la función $l(\theta)$. Es decir

$$H(\theta)_{i,j} = \frac{\partial^2 l(\theta)}{\partial \theta_i \partial \theta_j}.$$

Observación 1.3.3. Si $\hat{\theta}$ satisface

$$s(\hat{\theta}) = 0, \quad \frac{\partial^2 l(\hat{\theta})}{\partial \theta^2} = 0$$

entonces $\hat{\theta}$ es un estimador máximo verosímil.

Definición 1.3.7. Se define como *matriz de información de Fisher* a la matriz de covarianzas de la función score.

$$\mathcal{I}(\theta) = \text{Var}(s(\theta)) = \mathbb{E} \left[\left(\frac{\partial}{\partial \theta} l(\theta) \right)^2 \right].$$

La matriz puede ser definida también en términos de la matriz Hessiana.

$$\mathcal{I}(\theta) = -\mathbb{E}[H(\theta)] = -\mathbb{E} \left[\frac{\partial^2 l(\theta)}{\partial \theta^2} \right].$$

Proposición 1.3.1. ³ Suponiendo que el modelo es identificable y que la función de log-verosimilitud es integrable se tienen las siguientes propiedades del estimador máximo verosímil.

- i) El estimador $\hat{\theta}$ es **consistente**. Es decir que converge en probabilidad al parámetro verdadero conforme $n \rightarrow \infty$.
- ii) **Normalidad asintótica**. El estimador $\hat{\theta}$ converge en distribución a una variable normal multivariada.

$$\sqrt{n}(\hat{\theta} - \theta_0) \xrightarrow{d} \mathcal{N}(0, (\mathcal{I}(\theta_0))^{-1}),$$

donde θ es el parámetro verdadero y n el tamaño de la muestra.

Definición 1.3.8. Se denota $ECM(\hat{\theta})$ al **error cuadrático medio** del estimador $\hat{\theta}$, definido como

$$ECM(\hat{\theta}) = \mathbb{E}[(\hat{\theta} - \theta)^2].$$

Esta función permite evaluar la bondad de un estimador puntual.

³La prueba a la proposición se puede consultar en [23].

Capítulo 2

Cálculo estocástico

En la Sección 2.1 de este capítulo se construye la integral estocástica. Primero se construye para procesos simples. Después, por medio del lema de aproximación se define para procesos continuos. Además, se enuncian propiedades importantes de la integral estocástica. La Sección 2.2 aborda las fórmulas de Itô. El capítulo se basa en [12], principalmente en los capítulos 2 y 3, donde se pueden consultar las pruebas a las proposiciones y teoremas.

2.1. Construcción integral estocástica

En el cálculo diferencial e integral usual, la derivada $\frac{d}{dt}f(t)$ representa la tasa de cambio para una función $f(t)$. La función derivada puede depender tanto del tiempo como de la función original. Usando la notación de ecuación diferencial se lee cómo

$$df(t) = C(t, f(t))dt.$$

Dada una condición inicial $f(0) = t_0$ se puede recuperar la función $f(t)$

$$f(t) = t_0 + \int_0^t C(s, f(s))ds.$$

Por otro lado, en el cálculo estocástico la tasa de cambio de un proceso $\{X_t\}$ cuenta con un factor aleatorio que puede depender tanto del tiempo como del proceso. Agregando este factor, la tasa de cambio del proceso se describe con la siguiente notación

$$dX_t = A_t dB_t + C_t dt, \quad (2.1.1)$$

$$A_t = A(t, X_t), \quad C_t = C(t, X_t),$$

donde B_t es un movimiento browniano. Dada la condición inicial X_0 , la notación (2.1.1) y la siguiente son equivalentes

$$X_t = X_0 + \int_0^t A_s dB_s + \int_0^t C_s ds.$$

El objetivo de esta sección es definir al proceso $Z_t = \int_0^t A_s dB_s$; la integral estocástica o integral de Itô.

2.1.1. Integral estocástica para procesos simples

Definición 2.1.1. $\{A_t\}$ es un *proceso simple* adaptado a la filtración $\{\mathcal{F}_t\}$ si existen tiempos $0 = t_0 < t_1 < \dots < t_n = t < \infty$ y variables aleatorias Y_j , $j = 0, \dots, n$ tales que el proceso es de la forma

$$A_t = \sum_{j=0}^{n-1} Y_j \mathbb{1}_{[t_j, t_{j+1})}.$$

Las variables aleatorias Y_j son \mathcal{F}_{t_j} -medibles y $\mathbb{E}|Y_j^2| < \infty$, $j = 0, 1, \dots, n$.

Observación: Si a, b son constantes y $\{A_t\}, \{C_t\}$ son procesos simples, entonces el proceso $\{aA_t + bC_t\}$ también es un proceso simple.

Para definir la integral estocástica de un proceso simple se recuerda la integral de Riemann de una función escalonada. Dado el intervalo $[0, 1]$, $0 = t_0 < t_1 < \dots < t_n = 1$ una partición del intervalo, y la función escalonada

$$f(t) = a_j, \quad t_j \leq t < t_{j+1},$$

con a_j constante para toda $j = 0, \dots, n-1$, se define a la integral $\int_0^1 f(t)dt$ como

$$\int_0^1 f(t)dt = \sum_{j=0}^{n-1} a_j [t_{j+1} - t_j].$$

Para integrar $\{X_t\}$ con respecto al movimiento browniano $\{B_t\}$ se hace la suma de las variables aleatorias Y_j y en vez de multiplicar por el tamaño del intervalo, se multiplica por el incremento $B_{t_{j+1}} - B_{t_j}$. Este incremento es una variable aleatoria con distribución $N(0, t_{j+1} - t_j)$.

Definición 2.1.2. La *integral estocástica del proceso simple* $\{A_t\}$, se define como

- Si $t_n = t$

$$Z_t = \sum_{j=0}^{n-1} Y_j [B_{j+1} - B_j].$$

- Si $t_n < t$

$$Z_t = \sum_{j=0}^{n-1} Y_j [B_{j+1} - B_j] + Y_n [B_t - B_{t_n}].$$

Se denota como $Z_t = \int_0^t A_s dB_s$.

El proceso $\{A_t\}$ y el movimiento browniano $\{B_t\}$ son ambos adaptados a la filtración $\{\mathcal{F}_t\}$. La siguiente proposición describe cuatro propiedades de la integral estocástica para un proceso simple.

Proposición 2.1.1. Sea $\{B_t\}$ un movimiento browniano y $\{A_t\}, \{C_t\}$ dos procesos simples, Los tres procesos son adaptados a la filtración $\{\mathcal{F}_t\}$. Entonces se cumple que

(i) **Linealidad:** para $a, b \in \mathbb{R}$, la integral del proceso simple $\{aA_t + bC_t\}$ es

$$\int_0^t (aA_s + bC_s)dB_s = a \int_0^t A_s dB_s + b \int_0^t C_s dB_s.$$

La linealidad también se cumple respecto a los límites de integración. Si $0 < r < t$ entonces

$$\int_0^t A_s dB_s = \int_0^r A_s dB_s + \int_r^t A_s dB_s.$$

(ii) **Martingala:** La integral $Z_t = \int_0^t A_s dB_s$ como proceso estocástico es una martingala con respecto a $\{\mathcal{F}_t\}$.

(iii) **Isometría:** El proceso $\{Z_t\}$ es cuadrado integrable y su varianza es

$$\text{Var}[Z_t] = \mathbb{E}[Z_t^2] = \int_0^t \mathbb{E}[A_s^2] ds.$$

(iv) **Continuidad:** La función $t \mapsto Z_t$ es continua con probabilidad 1.

2.1.2. Integral estocástica para procesos continuos

A partir de integrales de procesos simples se define la integral para un proceso continuo y acotado. Esta integral se define como un límite. El siguiente lema asegura que existe una sucesión de procesos simples $\{A_t^{(n)}\}_{n \in \mathbb{N}}$ que permite definir la integral estocástica del proceso continuo $\{A_t\}$.

Lema 2.1.1. Sea $\{A_t\}$ un proceso continuo adaptado a la filtración $\{\mathcal{F}_t\}$. Sea $C < \infty$ tal que $\mathbb{P}(|A_t| \leq C) = 1$ para toda $t \geq 0$. Entonces existe una sucesión de procesos simples $\{A_t^{(n)}\}_{n \in \mathbb{N}}$ tales que para toda $t \geq 0$

$$\lim_{n \rightarrow \infty} \int_0^t \mathbb{E}[|A_s - A_s^{(n)}|^2] ds = 0.$$

Además, $|A_t^{(n)}| \leq C$ para todo $n \in \mathbb{N}, t \geq 0$.

Definición 2.1.3. La *integral estocástica del proceso continuo y acotado* $\{A_t\}$, denotada $Z_t = \int_0^t A_s dB_s$ se define como

$$Z_t = \lim_{n \rightarrow \infty} \int_0^t A_s^{(n)} dB_s,$$

donde $\{A_t^{(n)}\}_{n \in \mathbb{N}}$ es una sucesión de procesos simples que cumple (2.1.1).

Se nota que la sucesión de procesos $\{A_t^{(n)}\}_{n \in \mathbb{N}}$ no necesariamente converge al proceso $\{A_t\}$ para un valor fijo de t . Sin embargo eso no es necesario para definir la integral de $\{A_t\}$. La propiedades de la Proposición 2.1.1 también se cumple cuando $\{A_t\}$ y $\{C_t\}$ son procesos continuos y acotados.

La integral estocástica se puede definir para procesos $\{A_t\}$ que sean continuos pero no acotados. Para ello se define al tiempo de paro $T_m = \min\{t : |A_t| = m\}$ y al proceso $\{A_s^m\} = \{A_{s \wedge T_m}\}$, el cual es continuo y acotado. Por lo tanto se puede definir la integral de $\{A_s^m\}$ como

$$Z_t^m = \int_0^t A_s^m dB_s. \quad (2.1.2)$$

Definición 2.1.4. La *integral estocástica del proceso continuo* $\{A_t\}$, denotada $Z_t = \int_0^t A_s dB_s$ es

$$Z_t = \lim_{m \rightarrow \infty} Z_t^m.$$

Z_t^m se define como en (2.1.2).

Para $\{A_t\}$ y $\{C_t\}$ procesos continuos se cumplen (i) y (iv) de la Proposición 2.1.1.

Proposición 2.1.2. Sean $\{B_t\}$ es un movimiento browniano y $\{A_t\}$ un proceso continuo, adaptados a la filtración $\{\mathcal{F}_t\}$.

(ii) **Isometría de Itô:** La varianza del proceso $\{Z_t\}$ es

$$\text{Var}[Z_t] = \mathbb{E}[Z_t^2] = \mathbb{E} \left[\int_0^t A_s^2 ds \right] = \int_0^t \mathbb{E}[A_s^2] ds.$$

La varianza puede ser infinita. Si $\text{Var}[Z_t] < \infty$ entonces Z_t es cuadrado integrable.

(iii) **Martingala:** Si $\text{Var}[Z_t] < \infty$ entonces $\{Z_t\}$ como proceso es una martingala.

2.1.3. Variación y covariación cuadrática

A continuación se definen la variación total, variación cuadrática y covariación cuadrática de un proceso $\{X_t\}$. Son necesarias para definir las fórmulas de Itô. La variación y covariación cuadrática de un proceso son a su vez procesos estocásticos.

Definición 2.1.5. Sea $\Pi = \{0 = t_0 < t_1 < \dots < t_n = t\}$ una partición del intervalo $[0, t]$. La *norma* de Π es la longitud del intervalo con mayor longitud dentro de la partición

$$\|\Pi\| = \max_{i=1, \dots, n} |t_i - t_{i-1}|.$$

Definición 2.1.6. La *variación total* del proceso $\{X_t\}$ se define como

$$V_t(X) = \lim_{\|\Pi\| \rightarrow 0} \sum_{i=1}^n |X_{t_i} - X_{t_{i-1}}|.$$

La variación total se interpreta como la longitud de la trayectoria del proceso hasta el tiempo t . Un proceso tiene *variación acotada* si la variación total es finita; $V_t(X) < \infty$.

Definición 2.1.7. La *variación cuadrática* de un proceso es

$$\langle X_t \rangle = \lim_{\|\Pi\| \rightarrow 0} \sum_{i=1}^n (X_{t_i} - X_{t_{i-1}})^2.$$

Se utilizará a Π como la partición uniforme de $[0, t]$. Así, la variación cuadrática del proceso $\{X_t\}$ se puede definir también como

$$\langle X_t \rangle = \lim_{n \rightarrow \infty} \sum_{i=1}^{tn} (X_{i/n} - X_{(i-1)/n})^2.$$

Proposición 2.1.3. El movimiento browniano $\{B_t\}$ tiene variación total infinita, pero variación cuadrática finita. Además:

$$\langle B_t \rangle = t.$$

Definición 2.1.8. Sean $\{X_t\}$ y $\{Y_t\}$ dos procesos definidos en el mismo espacio de probabilidad. La *covariación cuadrática* de $\{X_t\}$ y $\{Y_t\}$ es

$$\langle X_t, Y_t \rangle = \lim_{\|\Pi\| \rightarrow 0} \sum_{i=1}^n (X_{t_i} - X_{t_{i-1}})(Y_{t_i} - Y_{t_{i-1}}).$$

También se puede definir usando la partición uniforme de $[0, t]$.

$$\langle X_t, Y_t \rangle = \lim_{n \rightarrow \infty} \sum_{i=1}^{tn} (X_{i/n} - X_{(i-1)/n})(Y_{i/n} - Y_{(i-1)/n}).$$

Proposición 2.1.4. Si X_t es continuo y Y_t de variación acotada entonces:

$$\langle X_t, Y_t \rangle = 0, \quad \forall t > 0.$$

2.2. Fórmulas de Itô

Hasta ahora se ha definido al proceso $Z_t = \int_0^t A_s dB_s$; la integral estocástica de un proceso estocástico $\{A_t\}$. Por el teorema fundamental del cálculo la función $f \in \mathcal{C}^1$ en $[a, b]$ se puede reescribir de la siguiente manera

$$f(t) = f(0) + \int_0^t f'(s) ds.$$

Es decir, se puede ver en términos de la integral de su derivada. El equivalente en el cálculo estocástico al teorema fundamental del cálculo es la fórmula de Itô. En esta sección se define la fórmula de Itô en sus diferentes versiones. Primero para procesos estocásticos que dependen de $\{B_t\}$.

Teorema 2.2.1 (Fórmula I de Itô). Sea $f \in \mathcal{C}^2$ y $\{B_t\}$ un movimiento browniano. Entonces para toda $t \geq 0$

$$f(B_t) = f(B_0) + \int_0^t f'(B_s) dB_s + \frac{1}{2} \int_0^t f''(B_s) ds.$$

En su forma diferencial

$$df(B_t) = f'(B_t) dB_t + \frac{1}{2} f''(B_t) dt.$$

Ejemplo 2.2.1. Se resuelve la integral estocástica $\int_0^t \cos(B_s)dB_s$ usando el teorema anterior. Sea $f(t) = \sin(t)$, entonces

$$\begin{aligned} f'(t) &= \cos(t), & f''(t) &= -\sin(t). \\ f'(B_t) &= \cos(B_t), & f''(B_t) &= -\sin(B_t). \end{aligned}$$

Remplazando en la fórmula de Itô

$$\begin{aligned} \sin(B_t) &= \sin(B_0) + \int_0^t \cos(B_s)dB_s - \frac{1}{2} \int_0^t \sin(B_s)ds. \\ \int_0^t \cos(B_s)dB_s &= \frac{1}{2} \int_0^t \sin(B_s)ds - \sin(B_t). \end{aligned}$$

La siguiente versión de la fórmula de Itô se utiliza para procesos que dependen de 2 variables (por ejemplo tiempo y espacio).

Teorema 2.2.2 (Fórmula II de Itô). Sea $f(t, x)$ una función tal que $f \in \mathcal{C}^1$ en t y \mathcal{C}^2 en x , $\{B_t\}$ un movimiento browniano. Para toda $t \geq 0$

$$f(t, B_t) = f(0, B_0) + \int_0^t \partial_x f(s, B_s)dB_s + \int_0^t \left[\partial_s f(s, B_s) + \frac{1}{2} \partial_{xx} f(s, B_s) \right] ds.$$

En su forma diferencial

$$df(t, B_t) = \partial_x f(t, B_t)dB_t + \left[\partial_t f(t, B_t) + \frac{1}{2} \partial_{xx} f(t, B_t) \right] dt.$$

Ejemplo 2.2.2. Se define al proceso estocástico $X_t = e^{-t} \left(\frac{B_t}{t}\right)^2$ y se utiliza el teorema anterior para encontrar dX_t . Sea $f(t, x) = e^{-t} \left(\frac{x}{t}\right)^2$, la función cumple las hipótesis del teorema y $X_t = f(t, B_t)$.

$$\partial_x f(t, x) = \frac{2xe^{-t}}{t^2}, \quad \partial_{xx} f(t, x) = \frac{2e^{-t}}{t^2}, \quad \partial_t f(t, x) = -\left(\frac{e^{-t}x^2}{t^2}\right) \left(\frac{2+t}{t}\right).$$

$$\begin{aligned} df(t, B_t) &= \frac{2B_t e^{-t}}{t^2} dB_t + \left[-\left(\frac{e^{-t}B_t^2}{t^2}\right) \left(\frac{2+t}{t}\right) + \frac{e^{-t}}{t^2} \right] dt, \\ df(t, B_t) &= \left(\frac{2}{B_t}\right) \left(\frac{e^{-t}B_t^2}{t^2}\right) dB_t + \left[\left(\frac{e^{-t}B_t^2}{t^2}\right) \left(\frac{1}{B_t^2} - \frac{2+t}{t}\right) \right] dt, \\ df(t, B_t) &= \left(\frac{2}{B_t}\right) f(t, B_t) dB_t + \left[f(t, B_t) \left(\frac{1}{B_t^2} - \frac{2+t}{t}\right) \right] dt. \end{aligned}$$

Por lo tanto el diferencial del proceso $X_t = e^{-t} \left(\frac{B_t}{t}\right)^2$ es

$$dX_t = \frac{2X_t}{B_t} dB_t + X_t \left(\frac{1}{B_t^2} - \frac{2+t}{t}\right) dt.$$

La tercera versión de la fórmula de Itô es el equivalente a la regla de la cadena en el cálculo estocástico. Describen a $f(X_t)$ y $f(t, X_t)$; funciones que dependen de un proceso $\{X_t\}$. Por último, se define la regla del producto para el cálculo estocástico.

Definición 2.2.1. Un *Proceso de Itô* $\{X_t : t \in [0, T]\}$ es de la forma

$$X_t = X_0 + \int_0^t H_s ds + \int_0^t A_s dB_s, \quad (2.2.1)$$

donde $\{H_t\}$, $\{A_t\}$ son dos procesos estocásticos adaptados a la filtración browniana $\{\mathcal{F}_t\}$, y que cumplen casi seguramente

$$H \text{ es integrable:} \quad \int_0^T |H_t| dt < \infty. \quad (2.2.2)$$

$$A \text{ es cuadrado integrable:} \quad \int_0^T A_t^2 dt < \infty. \quad (2.2.3)$$

En su notación diferencial

$$dX_t = H_t dt + A_t dB_t. \quad (2.2.4)$$

La ecuación (2.2.4) es una *ecuación diferencial estocástica*. El proceso X_t es *solución a la ecuación* (2.2.4) si se escribe como en (2.2.1).

Proposición 2.2.1. Si $\{X_t\}$ es un proceso de Itô, definido como en 2.2.1 entonces

$$\langle X_t \rangle = \int_0^t A_s^2 ds.$$

Al ser una integral estocástica, la variación cuadrática se puede escribir en su forma diferencial

$$d\langle X_t \rangle = A_t^2 dt.$$

Teorema 2.2.3 (Fórmula de Itô III). Sea $\{X_t\}$ un proceso de Itô que satisface (2.2.4), $f \in \mathcal{C}^2$, $f : \mathbb{R} \rightarrow \mathbb{R}$ una función. Entonces

$$\begin{aligned} f(X_t) &= f(X_0) + \int_0^t f'(X_s) dX_s + \frac{1}{2} \int_0^t f''(X_s) d\langle X_s \rangle, \\ &= f(X_0) + \int_0^t f'(X_s) [H_s ds + A_s dB_s] + \frac{1}{2} \int_0^t f''(X_s) A_s^2 ds, \\ &= f(X_0) + \int_0^t \left[f'(X_s) H_s + \frac{1}{2} f''(X_s) A_s^2 \right] ds + \int_0^t f'(X_s) A_s dB_s. \end{aligned}$$

En su forma diferencial

$$d(f(X_t)) = \left[f'(X_t) H_t + \frac{1}{2} f''(X_t) A_t^2 \right] dt + f'(X_t) A_t dB_t.$$

En cambio si $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(t, x) \in \mathcal{C}^1$ para t y $f(t, x) \in \mathcal{C}^2$ para x entonces

$$\begin{aligned} f(t, X_t) &= f(0, X_0) + \int_0^t \partial_s f(s, X_s) ds + \int_0^t \partial_x f(s, X_s) dX_s + \frac{1}{2} \int_0^t \partial_{xx} f(s, X_s) d\langle X_s \rangle \\ &= f(0, X_0) + \int_0^t \left[\partial_s f(s, X_s) + H_s \partial_x f(s, X_s) \right. \\ &\quad \left. + \frac{A_s^2}{2} \partial_{xx} f(s, X_s) \right] ds + \int_0^t A_s \partial_x f(s, X_s) dB_s. \end{aligned}$$

En su forma diferencial

$$d(f(t, X_t)) = \left[\partial_t f(t, X_t) + H_t \partial_x f(t, X_t) + \frac{A^2(t)}{2} \partial_{xx} f(t, X_t) \right] dt + A_t \partial_x f(t, X_t) dB_t.$$

Teorema 2.2.4. Sean $\{X_t\}$ y $\{Y_t\}$ dos procesos de Itô

$$dX_t = H_t dt + A_t dB_t, \quad dY_t = K_t dt + C_t dB_t.$$

Entonces

$$d(X_t Y_t) = X_t dY_t + Y_t dX_t + d\langle X_t, Y_t \rangle. \quad (2.2.5)$$

$\langle X_t, Y_t \rangle$ es la covariación cuadrática de los procesos $\{X_t\}$ y $\{Y_t\}$.

Capítulo 3

Ecuaciones diferenciales estocásticas y procesos de difusión

Aunque existe una fuerte relación las ecuaciones diferenciales estocásticas (EDE), sus soluciones y los procesos de difusión, no son conceptos equivalentes. De forma sencilla se puede definir un proceso de difusión como un proceso de Markov a tiempo continuo cuyas trayectorias son continuas casi seguramente.

Como se definió en la sección anterior, una EDE es un sistema dinámico estocástico de la forma

$$dX_t = H_t dt + A_t dB_t,$$

donde $\{H_t\}$ y $\{A_t\}$ son dos procesos estocásticos. Es importante notar que no siempre se tiene una solución explícita para la EDE, o esta puede no ser única. En la Sección 3.1 de este capítulo se presentan condiciones de existencia y unicidad para verificar si una EDE tiene o no soluciones, y si estas son fuertes o débiles.

Es posible definir a un proceso de difusión por medio del cálculo estocástico, ya que una representación de una difusión es como solución a una EDE. En la Sección 3.2 se presentan algunas definiciones relativas a este tipo de procesos.

Por último, en la Sección 3.3 se mencionan algunos ejemplos de procesos de difusión y las EDEs que satisfacen. Se obtiene el proceso solucionando la ecuación y se presentan algunas propiedades cada proceso.

El capítulo se basa en [17] donde se pueden consultar las pruebas e información complementaria.

3.1. Existencia y unicidad de soluciones fuertes y débiles

Definición 3.1.1. Consideremos la siguiente ecuación diferencial estocástica:

$$\begin{cases} dX_t &= m(t, X_t)dt + \sigma(t, X_t)dB_t \\ X_0 &= x_0 \end{cases} \quad (3.1.1)$$

donde dado $T > 0$:

$$\begin{aligned} &\text{Coeficiente de deriva. } m : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}, \\ &\text{Coeficiente de difusión. } \sigma : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}, \end{aligned}$$

son dos funciones medibles que satisfacen las condiciones (2.2.2) y (2.2.3), respectivamente. Un proceso \mathbf{X} que satisface (3.1.1) es un *proceso de difusión*.

Si $m(t, X_t) = m(X_t)$ y $\sigma(t, X_t) = \sigma(X_t)$ *proceso de difusión homogéneo*. Los coeficientes de deriva y difusión no dependen del tiempo.

Definición 3.1.2. El proceso $\mathbf{X} = \{X_t\}$ es una *solución fuerte* de (3.1.1) si dado cualquier espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$ con la filtración $\{\mathcal{F}_t\}$ y $\{B_t\}$ un movimiento browniano sucede que

- (a) \mathbf{X} es continuo y adaptado a $\{\mathcal{F}_t\}$,
- (b) Para todo $t \in [0, T]$

$$X_t = X_0 + \int_0^t m(s, X_s) ds + \int_0^t \sigma(s, X_s) dB_s \quad \text{c.s.} \quad (3.1.2)$$

$$\text{i.e., } \mathbb{P} \left(X_t = X_0 + \int_0^t m(s, X_s) ds + \int_0^t \sigma(s, X_s) dB_s, \forall t \in [0, T] \right) = 1.$$

X_0 es la condición inicial del proceso, la cual puede o no ser aleatoria. En caso de ser aleatoria debe ser independiente de $\{\mathcal{F}_t\}$ y cuadrado integrable. En este caso se utilizará una condición inicial determinista $X_0 = x \in \mathbb{R}$.

Definición 3.1.3. Una *solución débil* con condición inicial x para (3.1.1) es un proceso $\{X_t\}$ definido en algún espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$ tal que para algún movimiento browniano $\{B_t\}$ y filtración $\{\mathcal{F}_t\}$ el proceso es adaptado y satisface (3.1.2).

Observación 3.1.1. En la solución fuerte, a diferencia de la débil, se conoce a priori el espacio de probabilidad, la filtración y el movimiento browniano. El proceso $\{X_t\}$ se construye a partir de la condición inicial y la información generada por el movimiento browniano hasta el tiempo t .

En cambio en la solución débil el proceso puede no depender del browniano, pues no se necesita conocer la filtración previamente. Si una ecuación diferencial tiene una solución fuerte, entonces es también débil. El regreso de esta implicación no se cumple.

A continuación se enuncia el Teorema de existencia y unicidad para soluciones fuertes de (3.1.1), donde tanto el proceso \mathbf{X} como el movimiento browniano son unidimensionales. Para ello se definen las siguientes dos condiciones.

Condición 1. (*Globalmente Lipschitz*). Para todo $x, y \in \mathbb{R}$ y $t \in [0, T]$, existe una constante $D < \infty$ tal que

$$|m(t, x) - m(t, y)| + |\sigma(t, x) - \sigma(t, y)| \leq D|x - y|. \quad (3.1.3)$$

Observación. De la condición anterior se sigue que

$$|m(t, x) - m(t, y)| \leq D|x - y|, \quad |\sigma(t, x) - \sigma(t, y)| \leq D|x - y|. \quad (3.1.4)$$

Condición 2. (*Crecimiento lineal*). Para todo $x \in \mathbb{R}$ y $t \in [0, T]$, existe una constante $C < \infty$ tal que

$$|m(t, x)| + |\sigma(t, x)| \leq C(1 + |x|). \quad (3.1.5)$$

Esta condición garantiza que el proceso \mathbf{X} no explote. Es decir, que no tome valores infinitos en tiempos finitos.

Teorema 3.1.1. (Existencia y unicidad)¹ Sea $T > 0$ y

$$m : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}, \quad \sigma : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}$$

dos funciones medibles que cumplen las condiciones (2.2.2) y (2.2.3), respectivamente. Si se satisfacen también las condiciones (3.1.3) y (3.1.5) entonces existe $\mathbf{X} = \{X_t\}$, una única solución fuerte, continua y adaptada para la ecuación diferencial estocástica (3.1.1). Además

$$\mathbb{E} \left[\int_0^T |X_t|^2 dt \right] < \infty. \quad (3.1.6)$$

El teorema habla de unicidad en el sentido de trayectoria. Esto significa que cualesquiera dos soluciones que satisfagan (3.1.2) deben ser indistinguibles.

Ejemplo 3.1.1. (*Ecuación de Tanaka*) Se considera la función $g(x) = |x|$. Al aplicar la fórmula de Itô a $g(B_t)$ se obtendría

$$g(B_t) = g(B_0) + \int_0^t g'(B_s) dB_s + \frac{1}{2} \int_0^t g''(B_s) ds.$$

Sin embargo esto no es posible, pues g no es una función de clase C^2 . Para solucionar el problema se definen para $\varepsilon > 0$ las funciones g_ε .

$$g_\varepsilon(x) = \begin{cases} |x|, & |x| \geq \varepsilon \\ \frac{1}{2} \left(\varepsilon + \frac{x^2}{\varepsilon} \right), & |x| < \varepsilon. \end{cases} \quad (3.1.7)$$

Si $x \in (-\varepsilon, \varepsilon)$

$$g'_\varepsilon(x) = \frac{x}{\varepsilon}.$$

Por lo tanto

$$\int_0^t g'_\varepsilon(B_s) \mathbb{1}_{B_s \in (-\varepsilon, \varepsilon)} dB_s = \int_0^t \frac{B_s}{\varepsilon} \mathbb{1}_{B_s \in (-\varepsilon, \varepsilon)} dB_s.$$

La integral anterior converge en $L^2(\mathbb{P})$ a 0 si $\varepsilon \rightarrow 0$. Por isometría de la integral estocástica, y dado que $B_s < \varepsilon$

$$\begin{aligned} \mathbb{E} \left[\left(\int_0^t \frac{B_s}{\varepsilon} \mathbb{1}_{B_s \in (-\varepsilon, \varepsilon)} dB_s \right)^2 \right] &\leq \mathbb{E} \left[\int_0^t \frac{\varepsilon^2}{\varepsilon^2} \mathbb{1}_{B_s \in (-\varepsilon, \varepsilon)} ds \right] \\ &= \int_0^t \mathbb{E} [\mathbb{1}_{B_s \in (-\varepsilon, \varepsilon)}] ds \\ &= \int_0^t \mathbb{P}(B_s \in (-\varepsilon, \varepsilon)) ds. \end{aligned}$$

¹La prueba al teorema puede ser consultada en [21].

Notemos que $\lim_{\varepsilon \rightarrow 0} \mathbb{P}(B_s \in (-\varepsilon, \varepsilon)) = \mathbb{P}(B_s = 0) = 0$. Por lo tanto si $\varepsilon \rightarrow 0$ ocurre

$$\int_0^t g'_\varepsilon(B_s) \mathbb{1}_{B_s \in (-\varepsilon, \varepsilon)} dB_s \xrightarrow{L^2} 0.$$

Si $|B_s| > \varepsilon$ entonces

$$\int_0^t g'_\varepsilon(B_s) \mathbb{1}_{B_s \notin (-\varepsilon, \varepsilon)} dB_s = \int_0^t \text{sign}(B_s) \mathbb{1}_{B_s \notin (-\varepsilon, \varepsilon)} dB_s,$$

$$\text{sign}(x) = \begin{cases} +1 & x \geq 0 \\ -1 & x < 0 \end{cases} . \quad (3.1.8)$$

Por otro lado la segunda derivada de g_ε es

$$g''_\varepsilon(x) = \begin{cases} 0, & |x| \geq \varepsilon \\ \frac{1}{\varepsilon}, & |x| < \varepsilon \end{cases} . \quad (3.1.9)$$

Entonces

$$\frac{1}{2} \int_0^t g''_\varepsilon(B_s) ds = \frac{1}{2\varepsilon} \int_0^t \mathbb{1}_{B_s \in (-\varepsilon, \varepsilon)} ds.$$

Aplicando la fórmula de Itô a $g_\varepsilon(B_t)$

$$g_\varepsilon(B_t) = g_\varepsilon(B_0) + \int_0^t \left(g'_\varepsilon(B_s) \mathbb{1}_{B_s \notin (-\varepsilon, \varepsilon)} + g'_\varepsilon(B_s) \mathbb{1}_{B_s \in (-\varepsilon, \varepsilon)} \right) dB_s + \frac{1}{2\varepsilon} \int_0^t \mathbb{1}_{B_s \in (-\varepsilon, \varepsilon)} ds.$$

Tomando el límite cuando $\varepsilon \rightarrow 0$ se obtiene

$$|B_t| = |B_0| + \int_0^t \text{sign}(B_s) dB_s + L^t, \quad (3.1.10)$$

donde L^t es el tiempo local del movimiento browniano en 0 definido por

$$L^t = \lim_{\varepsilon \rightarrow 0} \frac{1}{2\varepsilon} \int_0^t \mathbb{1}_{B_s \in (-\varepsilon, \varepsilon)} ds.$$

Ahora, se define la siguiente ecuación diferencial estocástica

$$\begin{cases} dX_t = \text{sign}(X_t) dB_t \\ X_0 = 0 \end{cases} . \quad (3.1.11)$$

La función $\sigma(t, x) = \text{sign}(x)$ no es continua. Por lo tanto no cumple la condición de Lipschitz y la ecuación diferencial no tiene una solución fuerte, pero sí tiene una solución débil. Sea $\{\widehat{B}_s\}$ un movimiento browniano definido en un espacio de probabilidad filtrado $(\Omega, \mathcal{F}, \mathbb{P})$ con filtración $\{\mathcal{F}_t\}$. Sea $\{Y_t\}$ el proceso definido como

$$Y_t = \int_0^t \text{sign}(\widehat{B}_s) d\widehat{B}_s.$$

Teorema 3.1.2. ² Sea $dY_t = v dB_t$ un proceso de Itô con condición inicial $Y_0 = 0$. El proceso coincide en ley con un movimiento browniano si y sólo si $v^2 = 1$ casi seguramente.

²Una redacción más formal y prueba de este teorema se puede encontrar en [17] p. 143-146.

Por el teorema anterior $\{Y_t\}$ coincide en ley con un movimiento browniano. Entonces

$$dY_t = \text{sign}(\widehat{B}_t)d\widehat{B}_t \Rightarrow \frac{1}{\text{sign}(\widehat{B}_t)}dY_t = d\widehat{B}_t$$

$$d\widehat{B}_t = \text{sign}(\widehat{B}_t)dY_t$$

Por lo tanto $\{\widehat{B}_t\}$ es una solución débil para la ecuación diferencial estocástica (3.1.11). Por la igualdad (3.1.10) se tiene la siguiente contención de filtraciones

$$Y_t = |\widehat{B}_t| - |\widehat{B}_0| - L^t.$$

$$\sigma(Y_s : 0 \leq s \leq t) \subset \sigma(|\widehat{B}_s| : 0 \leq s \leq t) \subset \sigma(\widehat{B}_s : 0 \leq s \leq t). \quad (3.1.12)$$

Supongamos que $\{X_t\}$ es una solución fuerte para (3.1.11). En ese caso $\{B_t\}$ es un movimiento browniano y $\{X_t\}$ es adaptado a la filtración generada por $\{B_t\}$.

$$dX_t = \text{sign}(X_t)dB_t$$

$$dB_t = \text{sign}(X_t)dX_t$$

$\{X_t\}$ coincide en ley con un browniano. Usando la misma idea que en (3.1.12) se tiene

$$\sigma(B_s : 0 \leq s \leq t) \subset \sigma(X_s : 0 \leq s \leq t)!!!$$

Se tiene que $\{X_t\}$ no es adaptado a la filtración browniana, lo cual es una contradicción. Por lo tanto la ecuación de Tanaka no tiene soluciones fuertes.

3.2. Propiedades de los procesos de difusión

Definición 3.2.1. Sea $\{X_t\}$ un proceso de difusión definido como en (3.1.1). La *densidad de transición* del estado y al estado x , en un lapso de tiempo $t - s$ es un kernel de probabilidad que satisface

$$p(t, x|s, y) = \int_{-\infty}^{\infty} p(t, x|r, z)p(r, z|s, y)dz \quad \text{para cualquier } s < r < t.$$

Definición 3.2.2. El proceso $\{X_t\}$ es *homogéneo* (con respecto al tiempo) si

$$p(t, x|s, y) = p(t - s, x|0, y) = p(t - s, x|y).$$

Este tipo de difusiones son invariantes bajo traslaciones en el tiempo, pues los coeficientes no dependen explícitamente de t . En este caso la densidad de transición puede ser escrita como

$$p(t, x|y) = \int_{-\infty}^{\infty} p(t - s, x|z)p(s, z|y)dz \quad \text{para cualquier } 0 < s < t.$$

3.3. Ejemplos ecuaciones diferenciales estocásticas

3.3.1. Ornstein-Uhlenbeck y Vasicek

Definición 3.3.1. El proceso $\{X_t\}$ es un *proceso de Ornstein-Uhlenbeck (OU)* si satisface la siguiente ecuación diferencial estocástica

$$dX_t = \left(\alpha(t)X_t + \beta(t) \right) dt + \sigma(t, B_t) dB_t, \quad t \geq 0, \quad (3.3.1)$$

donde $\alpha(t)$ y $\beta(t)$ son funciones continuas e integrables y $\sigma(t, B_t)$ es continua y cuadrado integrable. Bajo estas condiciones el proceso tiene una única solución fuerte.

Solución: Se define $A(t) = \int_0^t \alpha(s) ds$.

$$\begin{aligned} dX_t &= \left(\alpha(t)X_t + \beta(t) \right) dt + \sigma(t, B_t) dB_t \\ e^{-A(t)} dX_t - e^{-A(t)} \alpha(t) X_t dt &= e^{-A(t)} \beta(t) dt + e^{-A(t)} \sigma(t, B_t) dB_t. \end{aligned}$$

Se usa la igualdad (2.2.5) y la Propiedad 2.1.4.

$$\begin{aligned} d(e^{-A(t)} X_t) &= e^{-A(t)} dX_t + X_t d(e^{-A(t)}) + d\langle e^{-A(t)}, X_t \rangle \\ &= e^{-A(t)} dX_t - X_t e^{-A(t)} \alpha(t) dt + \cancel{d\langle e^{-A(t)}, X_t \rangle}. \end{aligned}$$

$$\begin{aligned} d(e^{-A(t)} X_t) &= e^{-A(t)} \beta(t) dt + e^{-A(t)} \sigma(t, B_t) dB_t \\ e^{-A(t)} X_t &= e^{-A(0)} X_0 + \int_0^t e^{-A(s)} \beta(s) ds + \int_0^t e^{-A(s)} \sigma(s, B_s) dB_s \end{aligned}$$

Por lo tanto la solución fuerte a la ecuación diferencial (3.3.1) es

$$X_t = e^{A(t)} \left[X_0 + \int_0^t e^{-A(s)} \beta(s) ds + \int_0^t e^{-A(s)} \sigma(s, B_s) dB_s \right]. \quad (3.3.2)$$

Definición 3.3.2. El *proceso de Vasicek* es un caso particular del proceso Ornstein-Uhlenbeck, en el cual se satisface la siguiente ecuación diferencial estocástica

$$dX_t = a(b - X_t) dt + \sigma dB_t, \quad (3.3.3)$$

$$a, b, \sigma \in \mathbb{R}, \quad a, \sigma > 0, \quad X_0 = x_0 \geq 0.$$

A partir de la solución (3.3.2) para el proceso OU se obtiene la solución para (3.3.3)

$$\begin{aligned} X_t &= e^{-at} \left[x_0 - \int_0^t a b e^{as} ds + \int_0^t \sigma e^{as} dB_s \right], \\ X_t &= e^{-at} \left[x_0 - b(e^{at} - 1) + \sigma \int_0^t e^{as} dB_s \right]. \end{aligned}$$

Así, la solución a la ecuación (3.3.3) es

$$X_t = x_0 e^{-at} - b(e^{-at} - 1) + \sigma \int_0^t e^{-a(t-s)} dB_s. \quad (3.3.4)$$

Observación 3.3.1. Notemos que $\sigma \int_0^t e^{-a(t-s)} dB_s$ es la integral de una función determinista con respecto a B_t , que es un proceso gaussiano. Por lo tanto X_t es la suma de funciones deterministas más un término cuyos incrementos distribuyen normal.

Los coeficientes de deriva y difusión son

$$m(t, x) = a(b - x), \quad \sigma(t, X_t) = \sigma.$$

Ambas son funciones lineales respecto a X_t , así que cumplen las condiciones de Lipschitz y crecimiento lineal. Por lo tanto (3.3.4) es una solución fuerte y única para (3.3.3). Además el proceso es homogéneo.

Se calculan la esperanza y varianza del proceso. Para el cálculo de la esperanza se utiliza la propiedad de martingala de la integral estocástica, y la propiedad de isometría para el cálculo de la varianza.

$$\begin{aligned} \mathbb{E}[X_t] &= \mathbb{E}[x_0 e^{-at}] - \mathbb{E}[b(e^{-at} - 1)] + \sigma \mathbb{E} \left[\int_0^t e^{-a(t-s)} dB_s \right] \\ &= x_0 e^{-at} - b(e^{-at} - 1). \end{aligned} \quad (3.3.5)$$

$$\begin{aligned} \text{Var}(X_t) &= \text{Var} \left(x_0 e^{-at} - b(e^{-at} - 1) + \sigma \int_0^t e^{-a(t-s)} dB_s \right) \\ &= \sigma^2 e^{-2at} \text{Var} \left(\int_0^t e^{as} dB_s \right) \\ &= \sigma^2 e^{-2at} \int_0^t \mathbb{E} [e^{2as}] ds \\ &= \frac{\sigma^2 e^{-2at}}{2a} (e^{2at} - 1) \\ &= \frac{\sigma^2}{2a} (1 - e^{-2at}). \end{aligned} \quad (3.3.6)$$

Notemos que $\lim_{t \rightarrow \infty} \mathbb{E}[X_t] = b$. Por lo tanto al coeficiente b se le llama *media a largo plazo*. El coeficiente σ representa la *volatilidad*, y a la *velocidad de reversión a la media*. A partir de la Observación 3.3.1 y los resultados (3.3.5) y (3.3.6) se tiene la siguiente Proposición.

Proposición 3.3.1. Los incrementos de $\{X_t\}$ un proceso de Vasicek distribuyen como

$$X_t | X_s \sim \mathcal{N} \left(X_s e^{-a(t-s)} - b(e^{-a(t-s)} - 1), \frac{\sigma^2}{2a} (1 - e^{-2a(t-s)}) \right), \quad 0 \leq s < t. \quad (3.3.7)$$

Por lo tanto, la densidad de transición de un estado y al estado x en un tiempo t está dada por

$$p(t, x|y) = \frac{1}{\sqrt{\pi \sigma^2 (1 - e^{-2at})/a}} \exp \left(-\frac{(x - ye^{-at} + b(e^{-at} - 1))^2}{\sigma^2 (1 - e^{-2at})/a} \right). \quad (3.3.8)$$

El modelo de Vasicek es utilizado en finanzas para modelar tasas de interés. Esta aplicación se verá a mayor profundidad en el Capítulo 6.

3.3.2. Movimiento browniano geométrico

Definición 3.3.3. El proceso $\{X_t\}$ es un movimiento browniano geométrico si satisface la siguiente ecuación diferencial estocástica:

$$dX_t = \mu X_t dt + \sigma X_t dB_t, \quad \mu \in \mathbb{R}, \sigma > 0, \quad i = 1, \dots, N. \quad (3.3.9)$$

Solución. Sea $x_0 \in \mathbb{R}$ una condición inicial. El siguiente proceso es un movimiento browniano geométrico

$$X_t = x_0 \exp \left(\left(\mu - \frac{\sigma^2}{2} \right) t + \sigma B_t \right). \quad (3.3.10)$$

Se verifica utilizando la fórmula II de Itô:

$$f(t, x) = x_0 \exp \left(\left(\mu - \frac{\sigma^2}{2} \right) t + \sigma x \right), \quad f(t, B_t) = X_t$$

$$\partial_t f(t, x) = \left(\mu - \frac{\sigma^2}{2} \right) f(t, x), \quad \partial_x f(t, x) = \sigma f(t, x), \quad \partial_{xx} f(t, x) = \sigma^2 f(t, x).$$

Sustituyendo en la fórmula de Itô

$$\begin{aligned} dX_t &= \sigma f(t, B_t) dB_t + \left[\left(\mu - \frac{\sigma^2}{2} \right) f(t, B_t) + \frac{1}{2} \sigma^2 f(t, B_t) \right] dt \\ &= \sigma X_t dB_t + \mu X_t dt. \end{aligned}$$

Por lo tanto, X_t satisface (3.3.9). Los coeficientes de deriva y difusión son

$$m(t, X_t) = \mu X_t, \quad \sigma(t, X_t) = \sigma X_t.$$

Ambas funciones son lineales respecto a X_t . Por lo tanto la solución a la ecuación es fuerte y única, y el proceso es homogéneo,

Sea

$$Y_t = \left(\mu - \frac{\sigma^2}{2} \right) t + \sigma B_t.$$

Dado $t > 0$, la variable Y_t tiene distribución normal,

$$Y_t \sim \mathcal{N} \left(\left(\mu - \frac{\sigma^2}{2} \right) t, \sigma^2 t \right).$$

Por lo tanto $X_t = \exp(Y_t)$ es una variable con distribución log-normal.

Proposición 3.3.2. Si $\{X_t\}$ es un movimiento browniano geométrico entonces

$$X_t | X_0 = x_0 \sim \text{Lognormal} \left(\ln(x_0) + \left(\mu - \frac{\sigma^2}{2} \right) t, \sigma^2 t \right). \quad (3.3.11)$$

La densidad de transición del proceso está dada por

$$p(t, x|y) = \frac{1}{\sigma x \sqrt{2\pi t}} \exp \left(-\frac{\left(\ln(x) - \ln(y) - \left(\mu - \frac{\sigma^2}{2} \right) t \right)^2}{2\sigma^2 t} \right).$$

Como consecuencia de (3.3.10) y (3.3.11) se tiene que, dado $X_0 = x_0$

$$\ln\left(\frac{X_t}{x_0}\right) = \left(\mu - \frac{\sigma^2}{2}\right)t + \sigma B_t \sim \mathcal{N}\left(\left(\mu - \frac{\sigma^2}{2}\right)t, \sigma^2 t\right). \quad (3.3.12)$$

Se calcula la esperanza de X_t utilizando la función generadora de momentos de Y_t .

$$\begin{aligned} \mathbb{E}[X_t] &= \mathbb{E}[x_0 e^{Y_t}] \\ &= x_0 e^{(\mu - \frac{\sigma^2}{2})t} e^{t \frac{\sigma^2}{2}} \\ &= x_0 e^{t\mu}. \end{aligned}$$

Se calcula la varianza de X_t .

$$\mathbb{E}[X_t^2] = x_0^2 \mathbb{E}[e^{2Y_t}] = x_0^2 e^{2(\mu - \frac{\sigma^2}{2})t} e^{4t \frac{\sigma^2}{2}} = x_0^2 e^{2\mu t + \sigma^2 t}.$$

$$\begin{aligned} \text{Var}(X_t) &= \mathbb{E}[X_t^2] - (\mathbb{E}[X_t])^2 \\ &= x_0^2 e^{2\mu t + \sigma^2 t} - x_0^2 e^{2t\mu} \\ &= x_0^2 e^{2\mu t} (e^{\sigma^2 t} - 1). \end{aligned}$$

Una de las aplicaciones más importantes del proceso es modelar precios tal como el de acciones financieras. Gracias a la función exponencial, $X_t \geq 0$ para toda $t > 0$; como el precio de las acciones, el proceso sólo toma valores positivos. Esta aplicación se retomará en el Capítulo 6.

Así como el movimiento browniano tiene incrementos independientes, el browniano geométrico tiene razones de cambio independientes. Por ejemplo, dados los tiempos $0 = t_{i-1} < t_i < t_{i+1}$, Y_{t_i} es independiente a $Y_{t_{i+1}} - Y_{t_i}$.

$$\frac{X_{t_i}}{X_{t_{i-1}}} = e^{Y_{t_i}}, \quad \frac{X_{t_{i+1}}}{X_{t_i}} = e^{Y_{t_{i+1}} - Y_{t_i}}.$$

Por lo tanto la razón de cambio $\frac{X_{t_i}}{X_{t_{i-1}}}$ es independiente a $\frac{X_{t_{i+1}}}{X_{t_i}}$.

3.3.3. Puentes brownianos

Definición 3.3.4. Sea $\{B_t\}$ un movimiento browniano. El proceso $\{X_t : t \in [0, T]\}$ es un *puente browniano estándar* si

$$\{X_t : t \in [0, T]\} \stackrel{d}{=} \{B_t : t \in [0, T] | B_T = 0\}.$$

Es decir que tiene la misma distribución que un movimiento browniano condicional a tomar el valor 0 al tiempo $t = T$.

Una segunda definición para el puente browniano estándar es la de representación adelantada, en la cual se construye de forma explícita el puente con un movimiento browniano.

Definición 3.3.5. Sea $\{B_t\}$ un movimiento browniano. El proceso $\{X_t : t \in [0, T]\}$ es un *punte browniano estándar* si

$$X_t = B_t - \frac{t}{T}B_T, \quad t \in [0, T]. \quad (3.3.13)$$

Una tercera forma de definir al puente browniano es como solución a una ecuación diferencial estocástica. Esta definición generaliza al puente browniano estándar. En este caso se tiene $X_0 = \alpha$ y $X_T = \beta$.

Proposición 3.3.3. El puente browniano $\{X_t : t \in [0, T]\}$ satisface la siguiente ecuación diferencial estocástica

$$\begin{cases} dX_t = \frac{\beta - X_t}{T-t} dt + dB_t, & t \in [0, T] \\ X_0 = \alpha \end{cases}. \quad (3.3.14)$$

Solución. Sea $Y_t = \beta - X_t$. Se aplica la fórmula III de Itô.

$$f(x) = \beta - x, \quad f'(x) = -1, \quad f''(x) = 0.$$

$$\begin{aligned} d(Y_t) &= f'(X_t)dX_t + \frac{1}{2}f''(X_t)d\langle X_t \rangle \\ &= -dX_t. \end{aligned}$$

Se hace el cambio de variable a Y_t .

$$\begin{aligned} -dY_t &= \frac{Y_t}{T-t} dt + dB_t \\ dY_t + \frac{Y_t}{T-t} &= -dB_t \\ \frac{dY_t}{T-t} + \left(\frac{1}{T-t}\right)^2 Y_t dt &= -\frac{dB_t}{T-t} \\ d\left(\frac{Y_t}{T-t}\right) &= -\frac{dB_t}{T-t} \\ \frac{Y_t}{T-t} &= \frac{Y_0}{T} - \int_0^t \frac{1}{T-s} dB_s \\ (\beta - X_t) &= \frac{\beta - \alpha}{T}(T-t) - (T-t) \int_0^t \frac{1}{T-s} dB_s. \end{aligned}$$

Por lo tanto la solución a la ecuación (3.3.14) es

$$X_t = a + \frac{t}{T}(b-a) + (T-t) \int_0^t \frac{1}{T-s} dB_s.$$

3.3.4. Puentes de difusión

Todas las EDEs presentadas anteriormente se pueden asociar a un proceso de difusión que las soluciona. Para la ecuación que se presenta a continuación no siempre se conoce de forma explícita un proceso de difusión que la solucione.

Definición 3.3.6. Sea $\{X_t\}$ un proceso de difusión como en la Definición 3.1.1

$$\begin{cases} dX_t = m(t, X_t)dt + \sigma(t, X_t)dB_t, \\ X_0 = \alpha, \quad \alpha \in \mathbb{R}. \end{cases} \quad (3.3.15)$$

Se dice que $\{X^{T,\beta}\} = \{X_t^{T,\beta} : t \in [0, T]\}$ es un *punte de difusión* si tiene la misma distribución de probabilidad que $\{X_t\}$ condicionado a $X_T = \beta$, y $X_0^{T,\beta} = \alpha$.

Notación. Se denota como $(0, \alpha, T, \beta)$ -*punte* al proceso $\{X^{T,\beta}\}$. Si el proceso es solución a (3.3.15) y $X_{t_1} = \alpha$ y $X_{t_2} = \beta$, con $t_1 < t_2 \in [0, T]$ se denota como $(t_1, \alpha, t_2, \beta)$ -*punte*.

Capítulo 4

Simulación de difusiones mediante ecuaciones diferenciales estocásticas

Cuando se conoce, es posible simular un proceso de difusión mediante su densidad de transición. También es posible simular el proceso como solución a una EDE. Este método es útil cuando no se conoce de forma explícita la transición.

En este capítulo se aborda la simulación de los procesos descritas en el capítulo anterior. Aunque para la mayoría de los procesos sí se conoce la densidad de transición, la simulación se hará por medio de la ecuación diferencial que satisfacen. El capítulo está basado en [10].

En las Secciones 4.2 y 4.3 se describen los métodos de Euler y Milstein, con los cuales se discretiza y aproxima un proceso continuo. La transformada de Lamperti permite relacionar ambos métodos y tener una mejor convergencia del proceso aproximado al continuo, temas tratados en las Secciones 4.4 y 4.5. Por último, en la Sección 4.6 se muestran diferentes ejemplos de simulación de difusiones y la ecuación diferencial a la cual son solución.

4.1. Orden de convergencia

Sea $\{X_t\}$ un proceso de difusión, solución a la siguiente ecuación diferencial

$$\begin{cases} dX_t = m(t, X_t)dt + \sigma(t, X_t), & 0 \leq t \leq T, \\ X_0 = x_0. \end{cases}$$

Definición 4.1.1. Consideremos el intervalo de tiempo $[0, T]$ y la partición $\Pi_N([0, T])$

$$\Pi_N([0, T]) = \{0 = t_0 < t_1 < \dots < t_N = T\}. \quad (4.1.1)$$

Se define \mathbf{Y}_N , la *discretización o aproximación* del proceso $\mathbf{X} = \{X_t\}$ correspondiente a $\Pi_N([0, T])$ como el conjunto de observaciones

$$\mathbf{Y}_N = \{X_{t_i} : i = 0, 1, \dots, N\}.$$

Al simular ecuaciones diferenciales estocásticas se busca que las trayectorias del proceso aproximado sean similares a las del proceso de Itô original. Para verificar qué tan

óptima es la aproximación se pueden usar criterios de convergencia, ya sea fuerte o débil.

Se define $\delta = \|\Pi\|$ como la norma de la partición, o el tamaño de paso más grande. Si la partición es uniforme entonces $\delta = \frac{T}{N}$ y la partición es

$$\Pi_N([0, T]) = \left\{ t_i = \frac{iT}{N} : i = 0, 1, \dots, N \right\}. \quad (4.1.2)$$

Sea \mathbf{X} un proceso continuo que se aproxima utilizando la partición Π por medio del proceso \mathbf{Y}_N .

4.1.1. Orden fuerte de convergencia

Este criterio se utiliza para verificar la convergencia de una aproximación al proceso de forma trayectorial.

Definición 4.1.2. Se dice que \mathbf{Y} converge fuertemente con orden $\gamma \in (0, \infty]$ a \mathbf{X} si para cada tiempo $T > 0$

$$\mathbb{E}|Y_T - X_T| \leq C\delta^\gamma, \quad \forall \delta < \delta_0,$$

con $\delta_0 > 0$ y C una constante independiente a δ .

4.1.2. Orden débil de convergencia

Cuando no se necesita una aproximación en el sentido de trayectoria del proceso, sino una aproximación de la distribución de probabilidad del proceso, se puede usar el criterio de convergencia débil.

Definición 4.1.3. La aproximación \mathbf{Y} converge débilmente con orden β a \mathbf{X} si para cada $T > 0$

$$|\mathbb{E}g(Y_T) - \mathbb{E}g(X_T)| \leq C\delta^\beta, \quad \forall \delta < \delta_0,$$

donde $g \in C^{2(\beta+1)}$ es una función con crecimiento polinomial, $\delta_0 > 0$ y C una constante independiente a δ .

4.2. Método de Euler

Definición 4.2.1. El proceso estocástico continuo \mathbf{Y} es la *aproximación de Euler* de \mathbf{X} si satisface el siguiente esquema iterativo:

$$Y_{t_{i+1}} = Y_{t_i} + m(t_i, Y_{t_i})\Delta t_i + \sigma(t_i, Y_{t_i})\Delta B_{t_i}. \quad (4.2.1)$$

Donde

$$\begin{aligned} \Delta t_i &= t_{i+1} - t_i, \\ \Delta B_{t_i} &= B_{t_{i+1}} - B_{t_i}, \quad \Delta B_{t_i} \sim \mathcal{N}(0, \Delta t_i), \end{aligned}$$

para $i = 0, 1, \dots, N-1$ y $Y_0 = X_0$. Cuando el incremento en el tiempo es constante se define

$$\Delta t = t_{i+1} - t_i = T/N,$$

$$\Delta B_t = B_{t_{i+1}} - B_{t_i}, \quad \Delta B_t \sim \mathcal{N}(0, \Delta t).$$

Para que el proceso sea continuo se define Y_t haciendo interpolación lineal.

$$Y_t = Y_{t_i} + \frac{t - t_i}{t_{i+1} - t_i}(Y_{t_{i+1}} - Y_{t_i}), \quad t \in [t_i, t_{i+1}).$$

El método de Euler tiene un orden de convergencia fuerte $\gamma = 0.5$ y orden de convergencia débil $\beta = 1$.

4.3. Método de Milstein

El método de Milstein añade un término de segundo orden al esquema de Euler. Se utiliza la fórmula de Itô para mejorar la aproximación del proceso.

Definición 4.3.1. Se utiliza el mismo proceso y discretización que en el método de Euler. El proceso continuo \mathbf{Y} es la *aproximación de Milstein* de \mathbf{X} si satisface el siguiente sistema iterativo:

$$Y_{t_{i+1}} = Y_{t_i} + m(t_i, Y_{t_i})\Delta t_i + \sigma(t_i, Y_{t_i})\Delta B_{t_i} + \frac{1}{2}\sigma(t_i, Y_{t_i})\sigma'(t_i, Y_{t_i})((\Delta B_{t_i})^2 - \Delta t_i). \quad (4.3.1)$$

Como $\Delta B_{t_i} \sim \mathcal{N}(0, \Delta t_i)$ entonces $\Delta B_{t_i} = \sqrt{\Delta t_i}Z$ donde $Z \sim \mathcal{N}(0, 1)$. Por lo tanto

$$(\Delta B_{t_i})^2 - \Delta t_i = \Delta t_i Z^2 - \Delta t_i.$$

El esquema de Milstein se puede escribir también de la siguiente forma

$$Y_{t_{i+1}} = Y_{t_i} + m(t_i, Y_{t_i})\Delta t_i + \sigma(t_i, Y_{t_i})\sqrt{\Delta t_i}Z + \frac{1}{2}\sigma(t_i, Y_{t_i})\sigma'(t_i, Y_{t_i})\Delta t_i(Z^2 - 1). \quad (4.3.2)$$

Los órdenes de convergencia débil y fuerte coinciden y son $\gamma = 1 = \beta$.

4.4. Transformación de Lamperti

A continuación se define la transformación de Lamperti, la cual se usará para dar una relación entre los esquemas de Milstein y Euler.

Sea $\{X_t\}$ un proceso que satisface una ecuación diferencial estocástica, cuyo coeficiente de difusión depende únicamente del proceso

$$dX_t = m(t, X_t)dt + \sigma(X_t)dB_t. \quad (4.4.1)$$

Definición 4.4.1. La *transformación de Lamperti* de $\{X_t\}$ (definido como en (4.4.1)) es

$$Y_t = F(X_t) = \int_z^{X_t} \frac{1}{\sigma(u)} du, \quad (4.4.2)$$

donde z es un elemento arbitrario del espacio de estados de $\{X_t\}$.

La transformación de Lamperti permite transformar la ecuación (4.4.1) a una ecuación con coeficiente de difusión unitario. Esto se verifica aplicando la fórmula III de Itô a $Y_t = F(X_t)$.

$$\begin{aligned} dF(X_t) &= \left(m(t, X_t)F'(X_t) + \frac{1}{2}\sigma^2(X_t)F''(X_t) \right) dt + \sigma(X_t)F'(X_t)dB_t \\ &= \left(\frac{m(t, X_t)}{\sigma(X_t)} - \frac{1}{2}\sigma^2(X_t)\frac{\sigma'(X_t)}{\sigma^2(X_t)} \right) dt + \frac{\sigma(X_t)}{\sigma(X_t)}dB_t. \\ dY_t &= \left(\frac{m(t, X_t)}{\sigma(X_t)} - \frac{1}{2}\sigma'(X_t) \right) dt + dB_t \end{aligned}$$

4.5. Relación Euler y Milstein

A diferencia del método de Euler, el método de Milstein añade un término que involucra la primera derivada del coeficiente de difusión. Este término permite refinar el esquema de Euler. Ahora bien, si el proceso tiene un coeficiente de difusión constante se anulará el término adicional. Por lo tanto no habría diferencia entre aproximar dicho proceso con uno u otro método.

Consideremos $\{X_t\}$ un proceso tal que

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

Sea $Y_t = F(X_t)$, donde $F(x) = y$ es la transformación de Lamperti. Utilizando la fórmula III de Itô

$$dY_t = \left(F'(X_t)m(t, X_t) + \frac{1}{2}F''(X_t)\sigma^2(t, X_t) \right) dt + F'(X_t)\sigma(t, X_t)dB_t.$$

Al aplicar la transformación se obtiene

$$dY_t = \left(\frac{m(t, X_t)}{\sigma(t, X_t)} - \frac{1}{2}\sigma'(t, X_t) \right) dt + dB_t.$$

La aproximación de Euler para Y_t es

$$\begin{aligned} Y_{t_{i+1}} &= Y_{t_i} + \left(\frac{m(t_i, X_{t_i})}{\sigma(t_i, X_{t_i})} - \frac{1}{2}\sigma'(t_i, X_{t_i}) \right) \Delta t_i + \Delta B_{t_i}. \\ \Delta Y_{t_i} &= \left(\frac{m(t_i, X_{t_i})}{\sigma(t_i, X_{t_i})} - \frac{1}{2}\sigma'(t_i, X_{t_i}) \right) \Delta t_i + \sqrt{\Delta t_i}Z. \end{aligned}$$

Sea $G(y) = F^{-1}(y)$ la función inversa de F .

$$\begin{aligned} G'(y) &= (F^{-1}(y))' = \frac{1}{F'(F^{-1}(y))} = \sigma(t, G(y)), \\ G''(y) &= (\sigma(t, G(y)))' = \sigma'(t, G(y))G'(y) = \sigma'(t, G(y))\sigma(t, G(y)). \end{aligned}$$

La serie de Taylor de G es

$$G(y) = G(h) + G'(h)(y - h) + \frac{1}{2}G''(h)(y - h)^2 + O((y - h)^3)$$

Si $y = Y_{t_{i+1}} = Y_{t_i} + \Delta Y_{t_i}$ y $h = Y_{t_i}$, entonces $y - h = \Delta Y_{t_i}$.

$$\begin{aligned}
G(Y_{t_{i+1}}) &= G(Y_{t_i}) + G'(Y_{t_i})\Delta Y_{t_i} + \frac{1}{2}G''(Y_{t_i})(\Delta Y_{t_i})^2 + O((\Delta Y_{t_i})^3) \\
&= G(Y_{t_i}) + \sigma(t_i, G(Y_{t_i})) \left[\left(\frac{m(t_i, X_{t_i})}{\sigma(t_i, X_{t_i})} - \frac{1}{2}\sigma'(t_i, X_{t_i}) \right) \Delta t_i + \sqrt{\Delta t_i}Z \right] \\
&\quad + \frac{1}{2}\sigma'(t_i, G(Y_{t_i}))\sigma(t_i, G(Y_{t_i})) \left[\left(\frac{m(t_i, X_{t_i})}{\sigma(t_i, X_{t_i})} - \frac{1}{2}\sigma'(t_i, X_{t_i}) \right) \Delta t_i + \sqrt{\Delta t_i}Z \right]^2 + O((\Delta Y_{t_i})^3) \\
&= G(Y_{t_i}) + \left(m(t_i, X_{t_i}) - \frac{1}{2}\sigma(t_i, X_{t_i})\sigma'(t_i, X_{t_i}) \right) \Delta t_i + \sigma(t_i, X_{t_i})\sqrt{\Delta t_i}Z \\
&\quad + \frac{1}{2}\sigma(t_i, X_{t_i})\sigma'(t_i, X_{t_i})\Delta t_i Z^2 + O((\Delta t_i)^{3/2}).
\end{aligned}$$

Por lo tanto

$$\begin{aligned}
\Delta G(Y_{t_i}) &= \left(m(t_i, X_{t_i}) - \frac{1}{2}\sigma(t_i, X_{t_i})\sigma'(t_i, X_{t_i}) \right) \Delta t_i + \sigma(t_i, X_{t_i})\sqrt{\Delta t_i}Z \\
&\quad + \frac{1}{2}\sigma(t_i, X_{t_i})\sigma'(t_i, X_{t_i})\Delta t_i Z^2 + O((\Delta t_i)^{3/2}).
\end{aligned}$$

Utilizando (4.3.2) se tiene que la aproximación de Milstein de \mathbf{X} es

$$\begin{aligned}
\Delta X_{t_i} &= \left(m(t_i, X_{t_i}) - \frac{1}{2}\sigma(t_i, X_{t_i})\sigma'(t_i, X_{t_i}) \right) \Delta t_i + \sigma(t_i, X_{t_i})\sqrt{\Delta t_i}Z \\
&\quad + \frac{1}{2}\sigma(t_i, X_{t_i})\sigma'(t_i, X_{t_i})\Delta t_i Z^2.
\end{aligned}$$

Por lo tanto, la aproximación de Milstein de \mathbf{X} y la aproximación de Euler de la transformada del proceso son equivalentes hasta el orden Δt .

4.6. Simulación de algunas difusiones

Los códigos usados para la simulación de difusiones pueden ser consultados en el repositorio de Github dentro de la carpeta Códigos.

4.6.1. Simulación movimiento browniano

Existen diferentes formas de simular el movimiento browniano. En este trabajo se utilizará la propiedad de incrementos estacionarios para generar las trayectorias.

Algoritmo 1 : Simulación movimiento browniano

- 1: Dado $[0, T]$ un intervalo de tiempo, generar una partición uniforme con N subintervalos como en (4.1.2). El tamaño de paso será $\Delta t = T/N$.
- 2: Generar N variables aleatorias con distribución $X_{t_i} \sim \mathcal{N}(0, \Delta t)$ para $i = 1, \dots, N$, $X_0 = 0$.
- 3: Se hace una suma acumulada sobre las variables aleatorias, obteniendo una trayectoria simulada del proceso.

$$\left\{ B_{t_i} = \sum_{j=0}^i X_{t_j} : i = 0, 1, \dots, N \right\}.$$

En la Figura 4.1 se simula una trayectoria del movimiento browniano hasta el tiempo $T = 2$ y con condición inicial $X_0 = 1$. La aproximación se hace generando $N = 1000$ variables aleatorias.

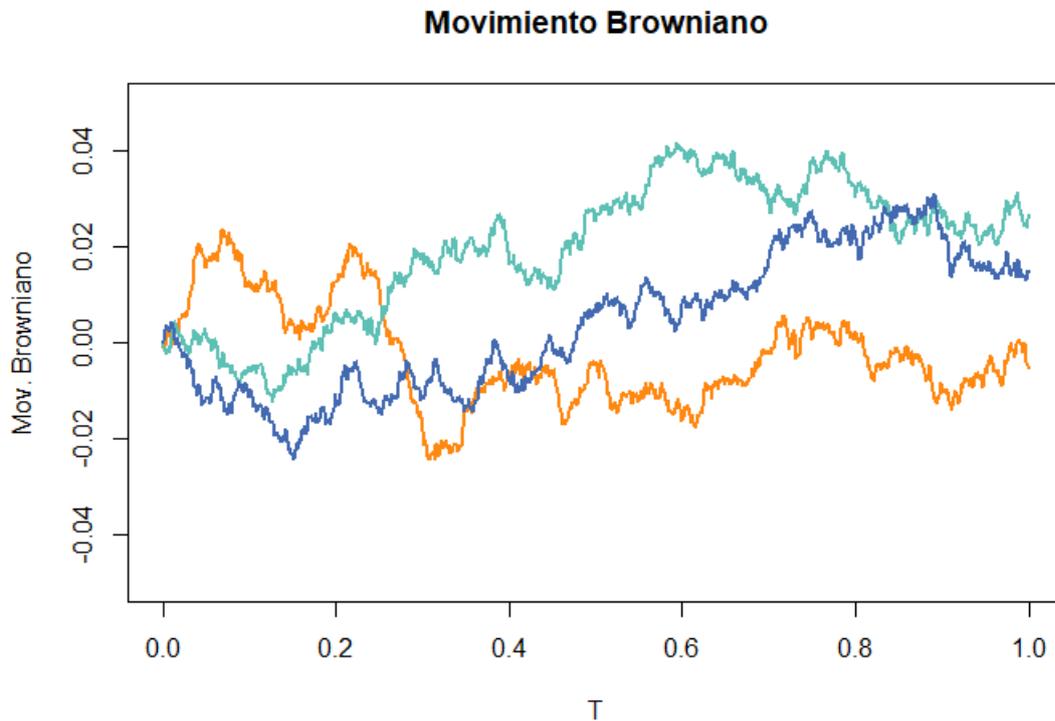


Figura 4.1: $T = 1$, $X_0 = 1$, $N = 1000$

4.6.2. Simulación Vasicek

El proceso de Vasicek tiene un coeficiente de difusión constante, por lo tanto se aproxima usando el esquema de Euler. Sean $a, b, \sigma \in \mathbb{R}$, $a, \sigma > 0$, $Z \sim \mathcal{N}(0, 1)$. La aproximación del proceso Vasicek en un intervalo de tiempo $[0, T]$, con tamaños de paso constantes $\Delta t = T/N$ sigue el siguiente esquema iterativo

$$X_{t_{i+1}} = X_{t_i} + a(b - X_{t_i})\Delta t + \sigma\sqrt{\Delta t}Z.$$

En la Figura 4.2 se simulan dos procesos con los mismos parámetros pero con diferente condición inicial y diferente parámetro a . Recordando que $b = 0.75$ representa la media a largo plazo y a representa la velocidad de reversión a la media, se puede ver cómo ambos procesos se aproximan a la media conforme avanzan en el tiempo. La trayectoria verde se aproxima más rápido que la naranja, ya que su velocidad de reversión es mayor.

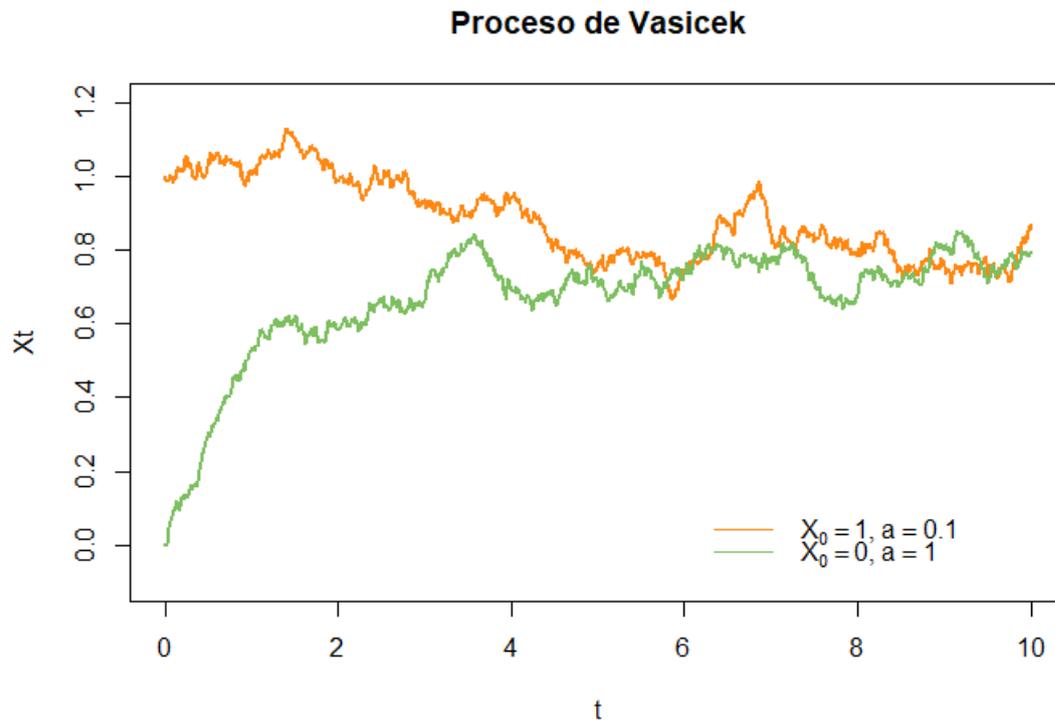


Figura 4.2: $T = 10, b = 0.75, \sigma = 0.1, N = 1000$

Las trayectorias simuladas en la Figura 4.3 tienen misma condición inicial y coeficiente a , pero diferente coeficiente b (media a largo a plazo) y σ (volatilidad). Cada trayectoria toma valores cercanos a b conforme avanza el tiempo, y la trayectoria verde tiene mayor volatilidad que la naranja, pues su coeficiente σ es mayor.

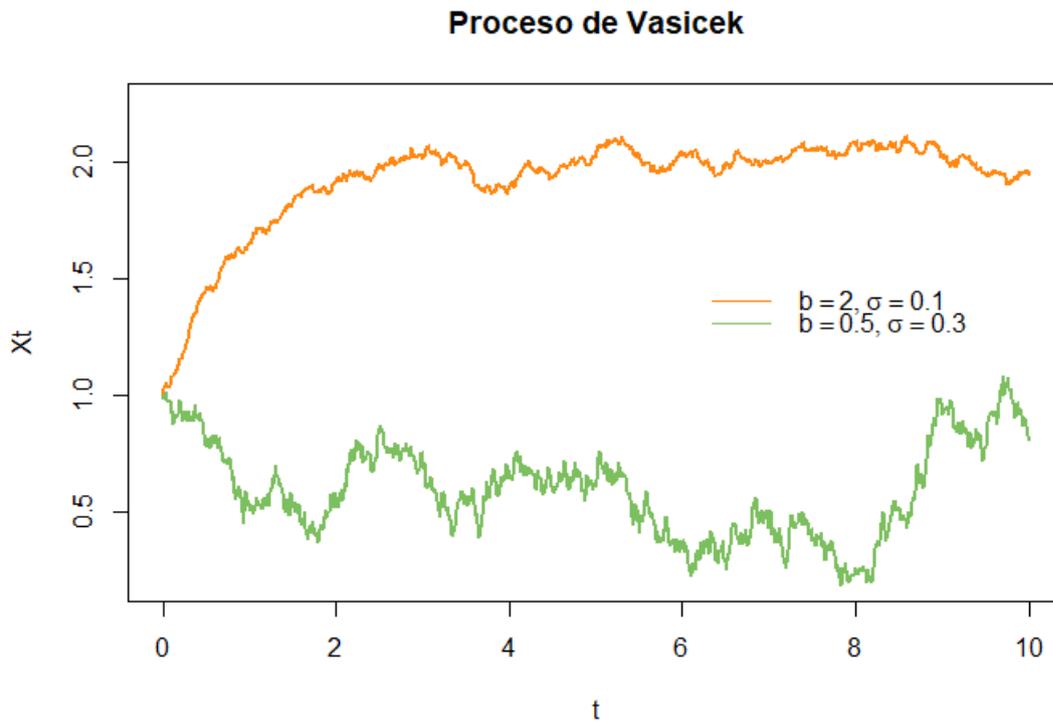


Figura 4.3: $T = 10, a = 1, N = 1000$

4.6.3. Simulación movimiento browniano geométrico

El movimiento browniano geométrico con coeficientes de deriva μ y difusión σ es solución a la ecuación diferencial (3.3.9). Se utiliza $\Delta t = T/N$ un tamaño de paso constante y la partición uniforme (4.1.2). Simulando con el método de Milstein se tiene el siguiente sistema iterativo

$$X_{t_{i+1}} = X_{t_i} + \mu X_{t_i} \Delta t + \sigma X_{t_i} \sqrt{\Delta t} Z + \frac{1}{2} \sigma^2 X_{t_i} \Delta t (Z^2 - 1).$$

En la Figura 4.4 se tienen tres trayectorias simuladas para un movimiento browniano geométrico con diferentes coeficientes de deriva. Así como en el modelo continuo, en el modelo simulado las trayectorias con un coeficiente μ mayor tienen pendientes más grandes.

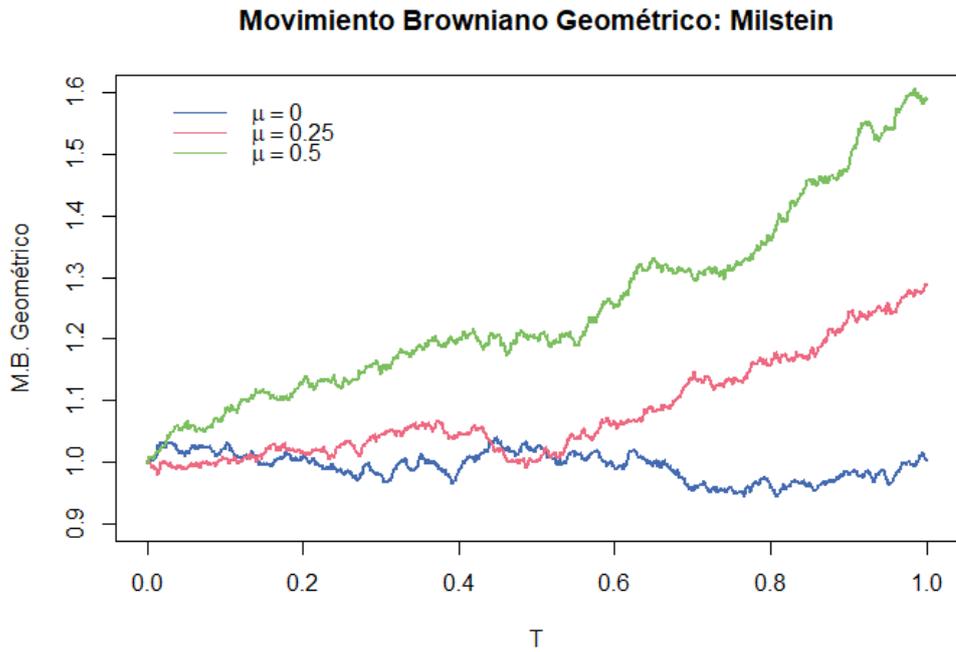


Figura 4.4: $T = 1, X_0 = 1, \sigma = 0.1, N = 1000$

En las trayectorias simuladas en la Figura 4.5 se varía el coeficiente de difusión. Como es de esperar por el modelo teórico, la simulación con $\sigma = 0.1$ tiene poca varianza comparada con la de parámetro $\sigma = 1.2$.

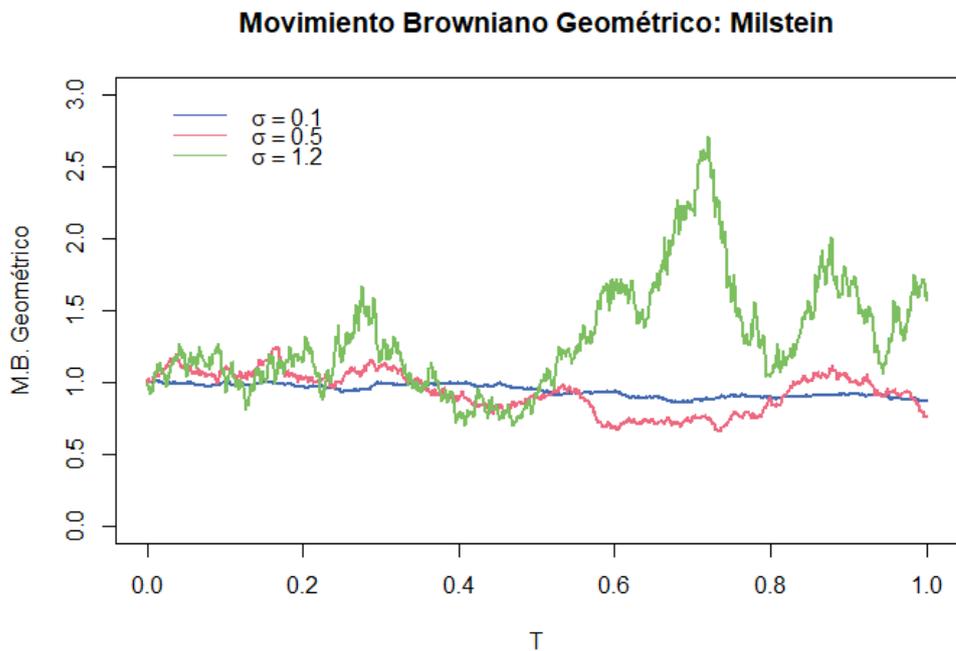


Figura 4.5: $T = 1, X_0 = 1, \mu = 0, N = 1000$

El proceso puede ser simulado también usando el método de Euler. Para ello se aplica

la transformación de Lamperti.

$$F(x) = \frac{1}{\sigma} \ln(x) \Rightarrow F'(x) = \frac{1}{\sigma x}, \quad F''(x) = -\frac{1}{\sigma x^2}.$$

Sea $F(X_t)$ la transformación de Lamperti del movimiento browniano geométrico $\{X_t\}$.

$$\begin{aligned} dF(X_t) &= \left(\frac{\mu X_t}{\sigma X_t} - \frac{1}{2}\sigma \right) dt + dB_t \\ d \ln(X_t) &= \left(\mu - \frac{1}{2}\sigma^2 \right) dt + \sigma dB_t \\ \ln(X_{t_{i+1}}) &= \ln(X_{t_i}) + \left(\mu - \frac{1}{2}\sigma^2 \right) \Delta t + \sigma \sqrt{\Delta t} Z \\ X_{t_{i+1}} &= \exp \left(\ln(X_{t_i}) + \left(\mu - \frac{1}{2}\sigma^2 \right) \Delta t + \sigma \sqrt{\Delta t} Z \right). \end{aligned}$$

En las Figuras 4.6 y 4.7 se simulan trayectorias usando el método de Euler para la transformación de Lamperti. De nuevo se varían los parámetros de deriva y difusión para comparar los comportamientos.

Movimiento Browniano Geométrico: Lamperti

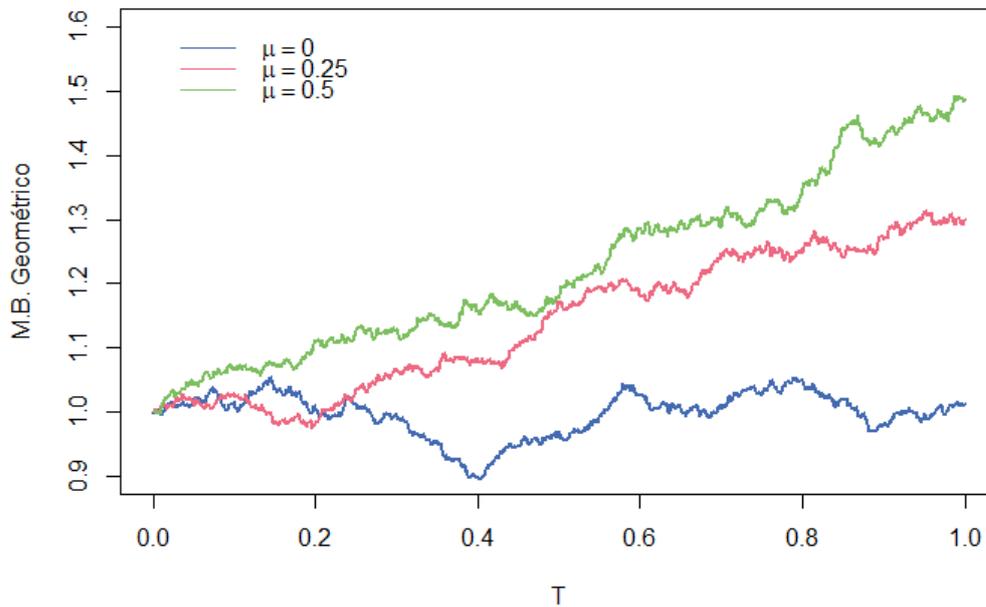


Figura 4.6: $T = 1, X_0 = 1, \sigma = 0.1, N = 1000$

Movimiento Browniano Geométrico: Lamperti

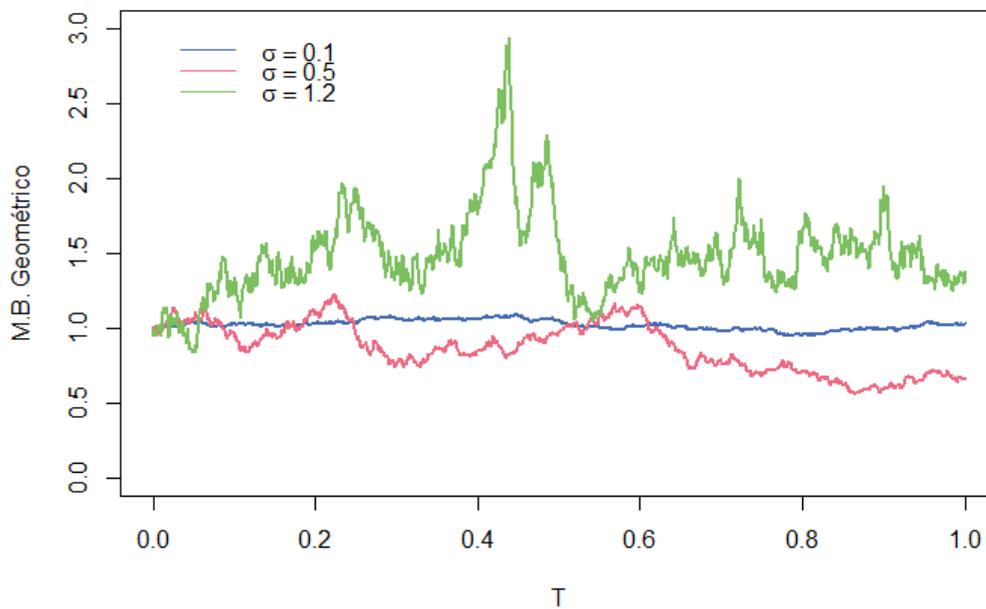


Figura 4.7: $T = 1, X_0 = 1, \mu = 0, N = 1000$

El comportamiento de las simulaciones usando el método de Euler coincide con las trayectorias simuladas usando el método de Milstein.

4.6.4. Simulación puente browniano

En el capítulo anterior se definió al puente browniano como solución a la ecuación diferencial (3.3.14) y también como una transformación del movimiento browniano (3.3.13).

La primera representación permite simular al proceso utilizando el método de Milstein. El proceso tiene un coeficiente de difusión unitario, por lo tanto en este caso el método de Euler coincide con el de Milstein. Dado el intervalo de tiempo $[0, T]$ y una partición uniforme con subintervalos de tamaño $\Delta t = T/N$, se aproxima el proceso con el siguiente esquema iterativo

$$X_{t_{i+1}} = X_{t_i} + \left(\frac{\beta - X_{t_i}}{T - t_i} \right) \Delta t + \sqrt{\Delta t} Z.$$

En la Figura 4.8 se simulan diferentes trayectorias de un puente browniano con parámetros $\alpha = 0, \beta = 2$ y $T = 1$. Todas las trayectorias comienzan en 0 y terminan en 2.

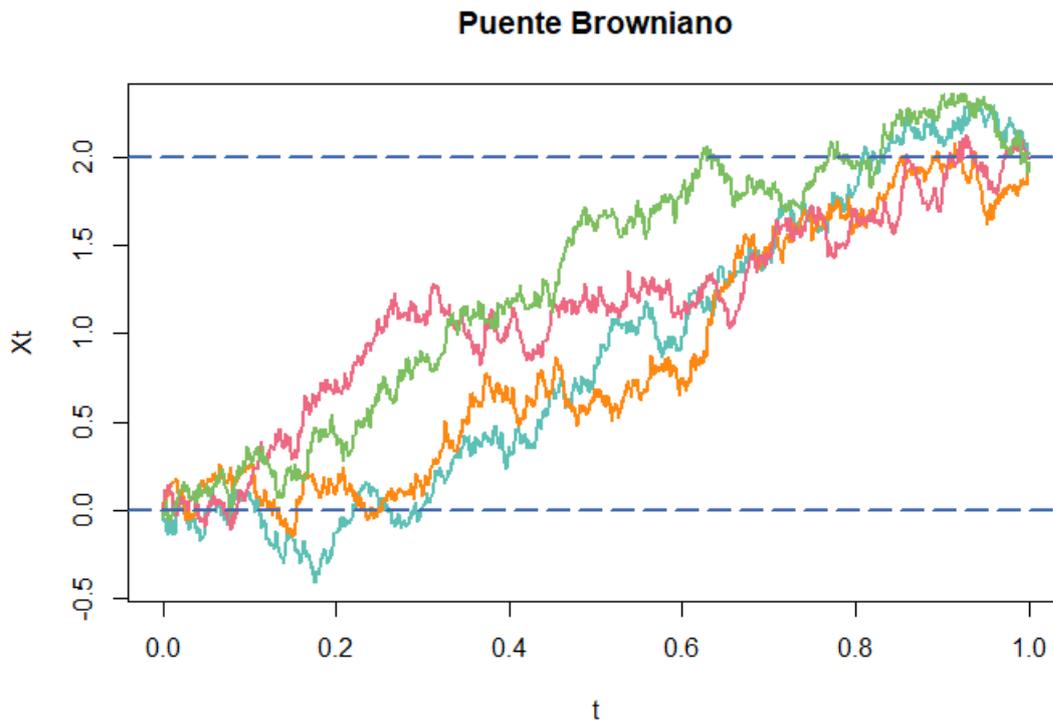


Figura 4.8: $T = 1, \alpha = 0, \beta = 2, N = 1000$

Otra forma de simular trayectorias del puente browniano es utilizando la representación adelantada.

Algoritmo 2 : Simulación puente browniano mediante representación adelantada

- 1: El intervalo de tiempo $[0, T]$ se discretiza con N valores $0 = t_0 < t_1 < \dots < t_N = T$.
 - 2: Se simula $\{B_{t_i}\}_{i=1}^N$ una trayectoria del movimiento browniano utilizando la discretización de $[0, T]$ y el Algoritmo 1.
 - 3: Nos fijamos en B_T ; el último valor que toma la trayectoria browniana simulada.
 - 4: Para cada $i = 0, \dots, N$ se define $X_i = B_i + \frac{t_i}{T}B_T$.
-

Se puede ver en la Figura 4.9 que al simular se obtienen diferentes trayectorias de un puente browniano estándar. Como se mencionó en el capítulo anterior, este es el tipo de puente generado con esta representación.

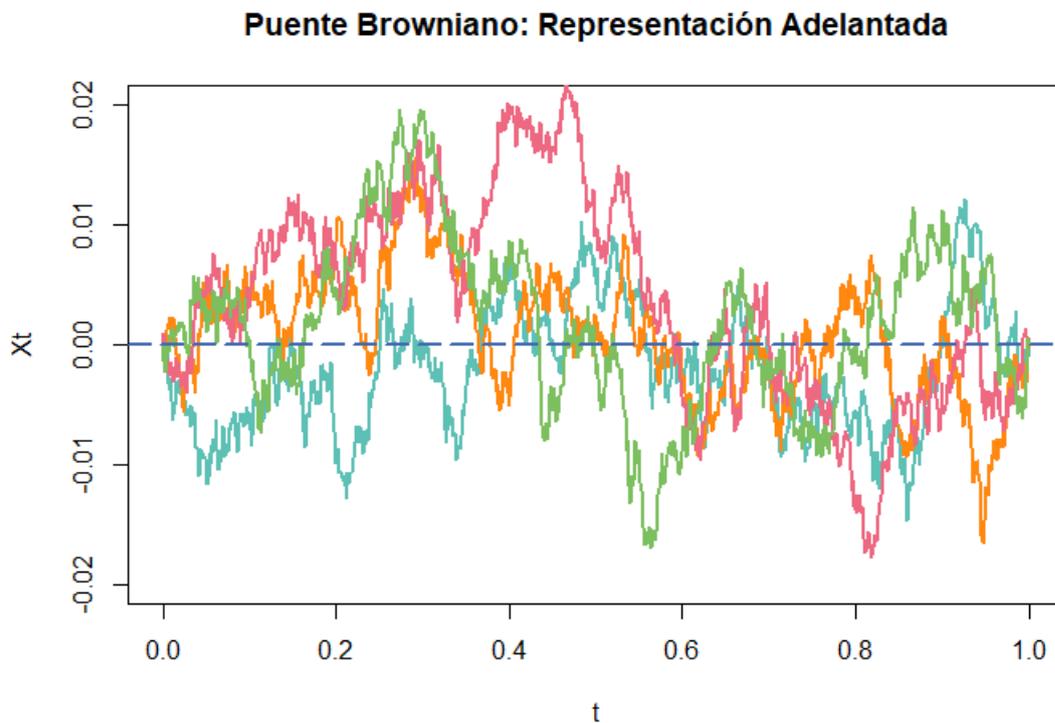


Figura 4.9: $T = 1, N = 1000$

4.6.5. Simulación puente de difusión

Los procesos hasta ahora simulados tienen una representación como solución a una ecuación diferencial estocástica. Por lo tanto, se pueden usar usando los métodos de Euler o Milstein. Sin embargo los puentes de difusión no siempre se pueden representar de esta forma, por lo que se usará otra técnica para la simulación de trayectorias. Esta subsección está basada en [3], donde se pueden consultar las pruebas a los teoremas enunciados.

Sea $\{X_t\}$ un proceso de difusión que satisface la siguiente ecuación diferencial estocástica

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

Sean B_t^1 y B_t^2 dos movimientos brownianos independientes. Se definen $\{X_t^1\}$ y $\{X_t^2\}$ como las soluciones a las siguientes ecuaciones

$$\begin{cases} dX_t^1 = m(t, X_t^1)dt + \sigma(t, X_t^1)dB_t^1 \\ X_0^1 = \alpha, \alpha \in \mathbb{R} \end{cases}, \quad \begin{cases} dX_t^2 = m(t, X_t^2)dt + \sigma(t, X_t^2)dB_t^2 \\ X_0^2 = \beta, \beta \in \mathbb{R} \end{cases}.$$

Se discretiza el intervalo de tiempo $[0, T]$ utilizando una partición uniforme $\Pi_N([0, T])$ como en (4.1.2). Usando tal partición se definen $\{Y_{t_i}^1\}$ y $\{Y_{t_i}^2\}$ aproximaciones para $\{X_t^1\}$ y $\{X_t^2\}$, respectivamente.

Consideremos $\{\tilde{Y}_{t_i}^2\} = \{Y_{t_{N-i}}^2\}$ el proceso a tiempo reverso de $\{Y_{t_i}^2\}$. Notamos que

$$\tilde{Y}_T = Y_0 = \beta.$$

Se simulan trayectorias de $\{Y_{t_i}^1\}$ y $\{\tilde{Y}_{t_i}^2\}$ hasta que haya un cruce entre ellas. Por ser discretizaciones, la única manera computacional de detectar un cruce es si hay i tal que $Y_{t_i}^1 \leq \tilde{Y}_{t_i}^2$ y $Y_{t_{i+1}}^1 \geq \tilde{Y}_{t_{i+1}}^2$ o $\tilde{Y}_{t_i}^2 \leq Y_{t_i}^1$ y $\tilde{Y}_{t_{i+1}}^2 \geq Y_{t_{i+1}}^1$. Se define ν como

$$\nu = \begin{cases} \min \{i \in \{1, \dots, N\} : Y_{t_i}^1 \leq \tilde{Y}_{t_i}^2\} & \text{si } Y_0^1 \geq \tilde{Y}_0^2 \\ \min \{i \in \{1, \dots, N\} : Y_{t_i}^1 \geq \tilde{Y}_{t_i}^2\} & \text{si } Y_0^1 \leq \tilde{Y}_0^2 \end{cases}.$$

Entonces una aproximación del $(0, \alpha, T, \beta)$ -puente es

$$Y_{t_i} = \begin{cases} Y_{t_i}^1 & i = 0, 1, \dots, \nu - 1 \\ \tilde{Y}_{t_i}^2 & i = \nu, \dots, N \end{cases}, \quad (4.6.1)$$

dado $\nu \leq N$. Es decir, antes del cruce se construye con los valores de $\{Y_{t_i}^1\}$ y después del cruce con $\{\tilde{Y}_{t_i}^2\}$. Existe la probabilidad de que las trayectorias no se crucen. Esta depende de la longitud del intervalo, los coeficientes de deriva y difusión y la distancia entre α y β .

La justificación teórica se basa en el siguiente teorema.

Teorema 4.6.1. Sea $\tau = \inf\{t \in [0, T] : X_t^1 = X_{T-t}^2\}$ ($\inf \emptyset = \infty$). Se define

$$Z_t = \begin{cases} X_t^1 & 0 \leq t \leq \tau \\ X_{T-t}^2 & \tau < t \leq T \end{cases}.$$

Entonces la distribución de $\{Z_t : 0 \leq t \leq T\}$ condicionado a $\{\tau \leq T\}$ es igual a la distribución condicional de $\{X_t : 0 \leq t \leq T\}$ dado $X_0 = \alpha$ y $X_T = \beta$. Es decir, $\{Z_t\}$ es un $(0, \alpha, T, \beta)$ -puente.

4.6.6. Simulación puente Vasicek

Se hace una implementación del algoritmo anterior para el caso particular de un proceso de Vasicek. Consideremos $\{X_t\}$ un proceso de Vasicek cuya media a largo plazo es 0

$$\begin{aligned} dX_t &= -\theta X_t dt + \sigma dB_t, \\ X_0 &= x_0. \end{aligned}$$

El siguiente lema da la distribución de un puente de Vasicek exacto.

Lema 4.6.1. Sea $\{0 = t_0 < t_1 < \dots < t_N = 1\}$ una partición del intervalo $[0, 1]$. Se define X_{t_i} como

$$\begin{aligned} X_{t_{i+1}} &= e^{-\theta\Delta t} X_{t_i} + W_i \\ X_0 &= x_0, \end{aligned}$$

para $i = 0, \dots, N - 1$, donde las variables W_i son independientes y distribuyen como

$$W_i \sim \mathcal{N}\left(0, (1 - e^{2\theta\Delta t}) \frac{\sigma^2}{2\theta}\right).$$

Se define Z_{t_i} para $i = 0, \dots, N$ como

$$Z_{t_i} = X_{t_i} + (x - X_{t_N}) \frac{e^{\theta t_i} - e^{-\theta t_i}}{e^{\theta t_N} - e^{-\theta t_N}}.$$

Entonces $\{Z_{t_i} : i = 0, 1, \dots, N\}$ distribuye como un puente de Vasicek con $Z_0 = x_0$ y $Z_{t_N} = x$.

Ejemplo 4.6.1. Consideremos $\{X_t\}$ un proceso de Vasicek que satisface la siguiente ecuación

$$dX_t = 0.1X_t dt + 0.5dB_t, t \in [0, 1].$$

Se simulará $\{X_t\}$ usando la aproximación descrita anteriormente y usando el Lema 4.6.1.

En las Figuras 4.10 y 4.11 se muestran diferentes trayectorias del $(0,0,0,1)$ -puente de difusión para $\{X_t\}$; aproximado y puente exacto. Se utilizan los parámetros $T = 1$ y $N = 100$.

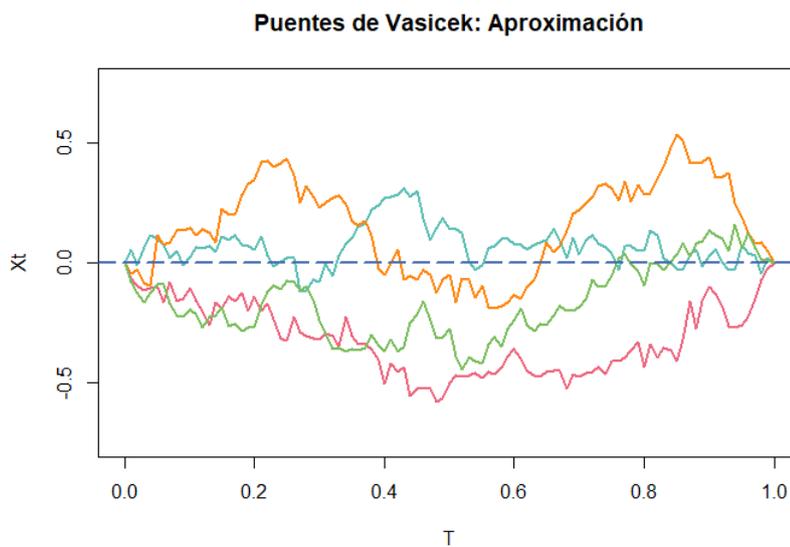


Figura 4.10: Puente Aproximado

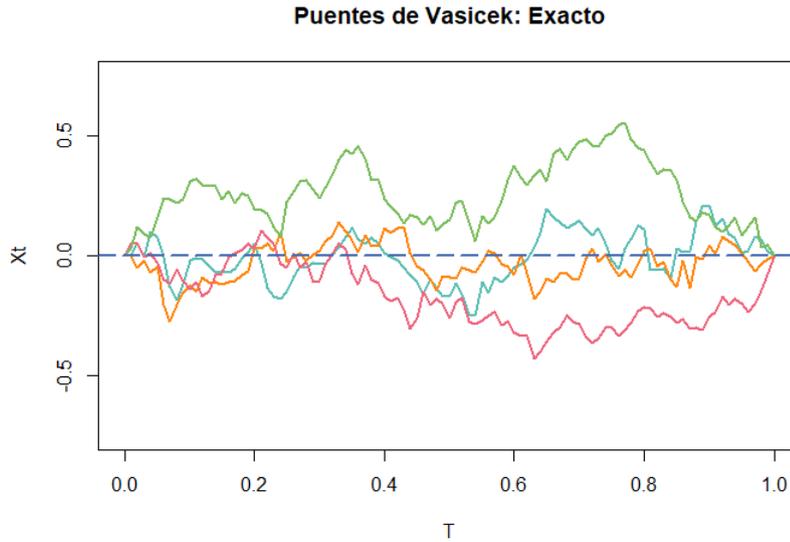


Figura 4.11: Puente Exacto

Estas dos simulaciones permiten que se haga una comparación entre las distribuciones de los puentes resultantes. Las distribuciones empíricas se comparan utilizando un gráfico Q-Q para poner a prueba el algoritmo.

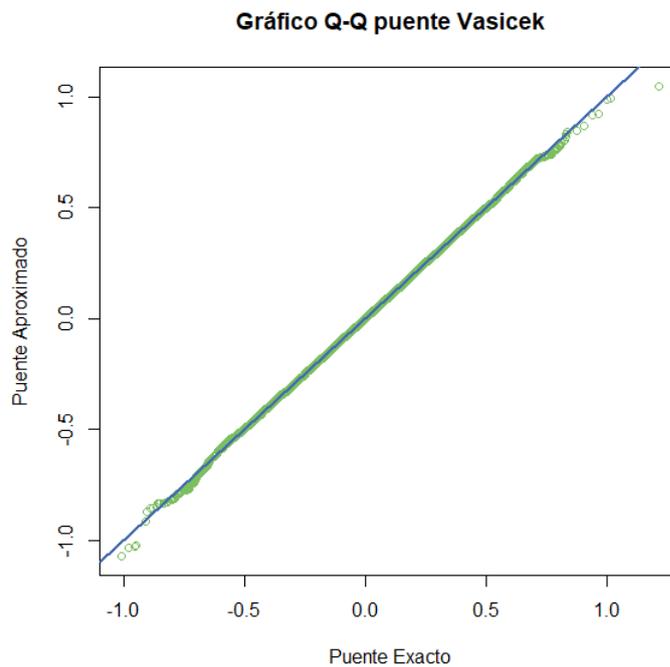


Figura 4.12: Gráfico de Q-Q

En la Figura 4.12 se aprecia cómo la distribución se aproxima a la recta identidad. Por lo tanto, hay evidencia empírica de que el puente aproximado tiene la misma distribución que el puente exacto. Entonces, el algoritmo para aproximar puentes de difusión se verifica.

4.6.7. Simulación puente browniano geométrico

Ejemplo 4.6.2. Se simulará un $(0,1,1,1)$ -puente un proceso con la siguiente difusión

$$dX_t = 0.3X_t dt + X_t dB_t.$$

Es decir, un movimiento browniano geométrico con $\mu = 0.3$ y $\sigma = 1$. Las trayectorias del proceso simulado se muestran en la Figura 4.13.

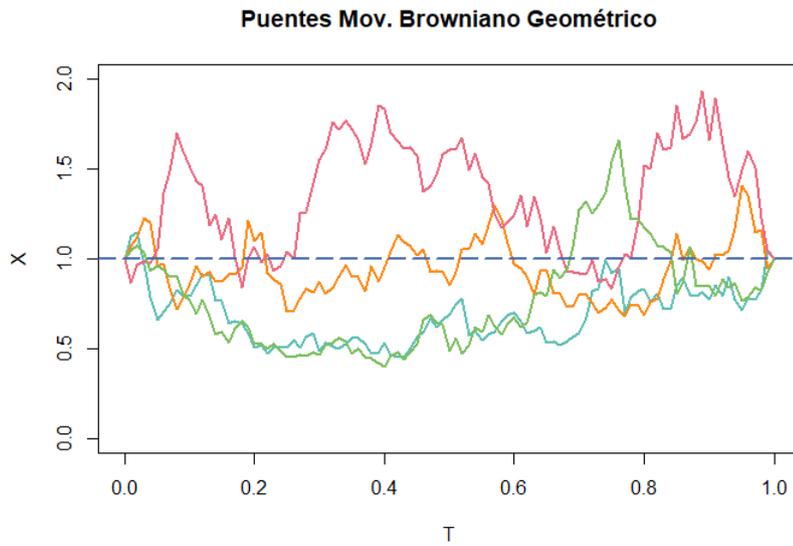


Figura 4.13: Puente Browniano Geométrico

Capítulo 5

Estimación de parámetros

Al estudiar datos que pueden ser modelados con la solución de una ecuación diferencial estocástica, es común no saber cuáles son los parámetros de dicho proceso. En este capítulo se presentan métodos para la estimación de los parámetros. En la Sección 5.1 se trata a los estimadores máximo verosímiles, en particular para procesos de difusión. Se hace un cálculo de los estimadores máximo verosímiles para los procesos de Vasicek y movimiento browniano geométrico. Para detalles sobre los conceptos presentados en esta sección ver [9], [24], [16], [4] y [5].

En la Sección 5.2 se aborda el algoritmo EM (Expectation-Maximization por sus siglas en inglés). Este será útil cuando los datos observados no son suficientes para tener convergencia con el estimador máximo verosímil. De igual manera, se pone a prueba estimando parámetros para los procesos de Vasicek y movimiento browniano geométrico. Ver [2] para más información sobre el algoritmo EM.

Los códigos con la implementación de los estimadores se encuentran dentro del repositorio de Github en la carpeta Códigos.

5.1. Estimadores basados en verosimilitud

Consideremos $\{X_t\}$ el proceso de difusión homogéneo que satisface la siguiente ecuación diferencial estocástica

$$dX_t = m(X_t; \theta)dt + \sigma(X_t; \theta)dB_t, \quad (5.1.1)$$

donde $\theta \in \Theta \subset \mathbb{R}^p$ es el vector de parámetros y X_0 es la condición inicial. Las funciones $m : \mathbb{R} \times \Theta \rightarrow \mathbb{R}$ y $\sigma : \mathbb{R} \times \Theta \rightarrow (0, \infty)$ cumplen las condiciones del Teorema 3.1.1. Por lo tanto existe una solución para (5.1.1). Adicionalmente a las 2 condiciones del Teorema 3.1.1, los coeficientes de deriva y difusión deben cumplir también con las siguientes condiciones para poder hacer inferencia basada en verosimilitud.

Condición 3. (*Positividad del coeficiente de difusión*).

$$\inf_x \sigma^2(x; \theta) > 0.$$

Condición 4. (*Momentos acotados*). Para cada $k > 0$ el k -ésimo momento del proceso $\{X_t\}$ existe y es acotado.

$$\sup_t \mathbb{E}|X_t|^k < \infty.$$

Condición 5. (*Coefficientes suaves o diferenciables*). Los coeficientes de deriva y difusión son 3 veces derivables respecto a θ , y de crecimiento polinomial respecto a x .

Definición 5.1.1. Supongamos que se tienen observaciones discretas X_{t_i} en los tiempos $0 = t_0 < t_1 < \dots < t_n = T$. La *función de verosimilitud* para la muestra de observaciones está dada por

$$L_n(\theta) = \prod_{i=1}^n p_\theta(\Delta t_i, X_{t_i} | X_{t_{i-1}}) p_\theta(X_0), \quad (5.1.2)$$

donde $\Delta t_i = t_i - t_{i-1}$ y $p_\theta(t, \cdot | x)$ es la densidad de transición de X_t dado $X_0 = x$. La *función de log-verosimilitud*, denotada $l_n(\theta)$ es

$$\begin{aligned} l_n(\theta) &= \ln \left(\prod_{i=1}^n p_\theta(\Delta t_i, X_{t_i} | X_{t_{i-1}}) p_\theta(X_0) \right) \\ &= \sum_{i=1}^n \ln (p_\theta(\Delta t_i, X_{t_i} | X_{t_{i-1}}) p_\theta(X_0)) \\ &= \sum_{i=1}^n \ln (p_\theta(\Delta t_i, X_{t_i} | X_{t_{i-1}})) + \ln (p_\theta(X_0)). \end{aligned} \quad (5.1.3)$$

Los procesos utilizados en este trabajo tienen densidades de transición conocidas explícitamente. Esto representa una gran ventaja porque se puede maximizar la verosimilitud de forma directa.

5.1.1. Verosimilitud del proceso Vasicek

Sea $\{X_t\}$ un proceso de Vasicek. El proceso satisface la siguiente ecuación diferencial estocástica

$$dX_t = a(b - X_t)dt + \sigma dB_t. \quad (5.1.4)$$

$$a, b, \sigma \in \mathbb{R}, \quad a, \sigma > 0, \quad X_0 = x_0 \geq 0.$$

La ecuación cuenta con tres parámetros $\theta = (\alpha, b, \sigma)$. Se calcula la función de log-verosimilitud utilizando (3.3.1).

Proposición 5.1.1. La log-verosimilitud de $\{X_t\}$ está dada por

$$l_n(\theta) = -\frac{n}{2} \ln(2\pi) - \frac{1}{2} \sum_{i=1}^n \ln \left(\frac{\sigma^2}{2a} (1 - e^{-2a\Delta t_i}) \right) - \sum_{i=1}^n \frac{a(X_{t_i} - X_{t_{i-1}} e^{-a\Delta t_i} - b(1 - e^{-a\Delta t_i}))^2}{\sigma^2(1 - e^{-2a\Delta t_i})}. \quad (5.1.5)$$

Demostración. Se calcula $\ln(p_\theta(\Delta t_i, X_{t_i} | X_{t_{i-1}}))$.

$$\begin{aligned} \ln(p_\theta(\Delta t_i, X_{t_i} | X_{t_{i-1}})) &= \ln \left(\frac{1}{\sqrt{\pi\sigma^2(1 - e^{-2a\Delta t_i})}/a} \exp \left(\frac{-(X_{t_i} - X_{t_{i-1}} e^{-a\Delta t_i} + b(e^{-a\Delta t_i} - 1))^2}{\sigma^2(1 - e^{-2a\Delta t_i})/a} \right) \right) \\ &= -\frac{1}{2} \ln \left(2\pi \frac{\sigma^2}{2a} (1 - e^{-2a\Delta t_i}) \right) + \left(\frac{-(X_{t_i} - X_{t_{i-1}} e^{-a\Delta t_i} + b(e^{-a\Delta t_i} - 1))^2}{\sigma^2(1 - e^{-2a\Delta t_i})/a} \right) \\ &= -\frac{1}{2} \ln(2\pi) - \frac{1}{2} \ln \left(\frac{\sigma^2}{2a} (1 - e^{-2a\Delta t_i}) \right) \\ &\quad - \frac{a}{\sigma^2(1 - e^{-2a\Delta t_i})} (X_{t_i} - X_{t_{i-1}} e^{-a\Delta t_i} + b(e^{-a\Delta t_i} - 1))^2. \end{aligned}$$

Como $X_0 = x_0$ es una condición inicial determinista entonces $p_\theta(X_0) = 1$ y $\ln(p_\theta(X_0)) = 0$. Por último se calcula $\sum_{i=1}^n \ln(p_\theta(\Delta t_i, X_{t_i}|X_{t_{i-1}}))$.

$$\begin{aligned} \sum_{i=1}^n \ln(p_\theta(\Delta t_i, X_{t_i}|X_{t_{i-1}})) &= \sum_{i=1}^n \left(-\frac{1}{2} \ln(2\pi) - \frac{1}{2} \ln \left(\frac{\sigma^2}{2a} (1 - e^{-2a\Delta t_i}) \right) \right) \\ &\quad - \frac{a}{\sigma^2(1 - e^{-2a\Delta t_i})} (X_{t_i} - X_{t_{i-1}} e^{-a\Delta t_i} + b(e^{-a\Delta t_i} - 1))^2 \\ &= -\frac{n}{2} \ln(2\pi) - \frac{1}{2} \sum_{i=1}^n \ln \left(\frac{\sigma^2}{2a} (1 - e^{-2a\Delta t_i}) \right) \\ &\quad - \frac{a}{\sigma^2(1 - e^{-2a\Delta t_i})} \sum_{i=1}^n (X_{t_i} - X_{t_{i-1}} e^{-a\Delta t_i} + b(e^{-a\Delta t_i} - 1))^2. \end{aligned}$$

Por lo tanto la log-verosimilitud del proceso Vasicek es

$$l_n(\theta) = -\frac{n}{2} \ln(2\pi) - \frac{1}{2} \sum_{i=1}^n \ln \left(\frac{\sigma^2}{2a} (1 - e^{-2a\Delta t_i}) \right) - \sum_{i=1}^n \frac{a(X_{t_i} - X_{t_{i-1}} e^{-a\Delta t_i} - b(1 - e^{-a\Delta t_i}))^2}{\sigma^2(1 - e^{-2a\Delta t_i})}.$$

Para facilitar la lectura se define

$$Y_{t_i} = X_{t_i} - X_{t_{i-1}} e^{-a\Delta t_i} - b(1 - e^{-a\Delta t_i}), \quad \gamma_i = e^{-a\Delta t_i}.$$

Se reescribe la log-verosimilitud como

$$l_n(\theta) = -\frac{n}{2} \ln(2\pi) - \frac{1}{2} \sum_{i=1}^n \ln \left(\frac{\sigma^2}{2a} (1 - \gamma_i^2) \right) - \sum_{i=1}^n \frac{aY_{t_i}^2}{\sigma^2(1 - \gamma_i^2)}.$$

Observación 5.1.1. Las derivadas parciales de $l_n(\theta)$ están dadas por

$$\frac{\partial l_n(\theta)}{\partial a} = \frac{n}{2a} - \sum_{i=1}^n \frac{\Delta t_i}{\gamma_i^{-2} - 1} + \frac{1}{\sigma^2} \sum_{i=1}^n \frac{Y_{t_i} ((Y_{t_i}(\gamma_i^2 - 1) + 2a\Delta t_i \gamma_i ((X_{t_i} - b)\gamma_i - (X_{t_{i-1}} - b)))}{(1 - \gamma_i^2)^2},$$

$$\frac{\partial l_n(\theta)}{\partial b} = \frac{2a}{\sigma^2} \sum_{i=1}^n \frac{Y_{t_i}(1 - \gamma_i)}{1 - \gamma_i^2},$$

$$\frac{\partial l_n(\theta)}{\partial \sigma} = -\frac{n}{2\sigma^2} + \frac{a}{\sigma^4} \sum_{i=1}^n \frac{Y_{t_i}^2}{1 - \gamma_i^2}.$$

Las segundas derivadas parciales están dadas por

$$\begin{aligned} \frac{\partial^2 l_n(\theta)}{\partial a^2} &= -\frac{n}{2a^2} + 2 \sum_{i=1}^n \frac{\Delta t_i^2 \gamma_i^2}{(1 - \gamma_i^2)^2} - \frac{4a}{\sigma^2} \sum_{i=1}^n \frac{\Delta t_i^2 \gamma_i^2 Y_{t_i}^2}{(1 - \gamma_i^2)^2} - \frac{8a}{\sigma^2} \sum_{i=1}^n \frac{\Delta t_i^2 \gamma_i^4 Y_{t_i}^2}{(1 - \gamma_i^2)^3} \\ &\quad + \frac{2a}{\sigma^2} \sum_{i=1}^n \frac{\Delta t_i^2 \gamma_i (X_{t_{i-1}} - b) Y_{t_i}}{1 - \gamma_i^2} + \frac{8a}{\sigma^2} \sum_{i=1}^n \frac{\Delta t_i^2 \gamma_i^3 (X_{t_{i-1}} - b) Y_{t_i}}{(1 - \gamma_i^2)^2} - \frac{2a}{\sigma^2} \sum_{i=1}^n \frac{\Delta t_i^2 \gamma_i^2 (X_{t_{i-1}} - b)^2}{1 - \gamma_i^2} \\ &\quad + \frac{4}{\sigma^2} \sum_{i=1}^n \frac{\Delta t_i \gamma_i^2 Y_{t_i}^2}{(1 - \gamma_i^2)^2} - \frac{4}{\sigma^2} \sum_{i=1}^n \frac{\Delta t_i \gamma_i (X_{t_{i-1}} - b) Y_{t_i}}{1 - \gamma_i^2}, \end{aligned}$$

$$\frac{\partial^2 l_n(\theta)}{\partial b^2} = -\frac{2a}{\sigma^2} \sum_{i=1}^n \frac{(1 - \gamma_i)^2}{1 - \gamma_i^2},$$

$$\frac{\partial^2 l_n(\theta)}{\partial (\sigma^2)^2} = \frac{n}{2\sigma^2} - \frac{2a}{\sigma^6} \sum_{i=1}^n \frac{Y_{t_i}^2}{1 - \gamma_i^2}.$$

Proposición 5.1.2. ¹ Dado Δt_i un tamaño de paso constante, los estimadores máximo verosímiles del proceso de Vasicek están dados por

$$\begin{aligned}\hat{a} &= -\frac{1}{\Delta t_i} \ln \left(\frac{n \sum_{i=1}^n X_{t_i} X_{t_{i-1}} - \sum_{i=1}^n X_{t_i} \sum_{i=1}^n X_{t_{i-1}}}{n \sum_{i=1}^n X_{t_{i-1}}^2 - \left(\sum_{i=1}^n X_{t_{i-1}} \right)^2} \right), \\ \hat{b} &= \frac{1}{n} \left(\sum_{i=1}^n \frac{X_{t_i}}{(1 - e^{-\hat{a} \Delta t_i})} - \sum_{i=1}^n \frac{X_{t_{i-1}}}{e^{\hat{a} \Delta t_i} - 1} \right), \\ \hat{\sigma}^2 &= \frac{2\hat{a}}{n} \sum_{i=1}^n \frac{\left(X_{t_i} - X_{t_{i-1}} e^{-\hat{a} \Delta t_i} - \hat{b}(1 - e^{-\hat{a} \Delta t_i}) \right)^2}{(1 - e^{-2\hat{a} \Delta t_i})}.\end{aligned}$$

Observación 5.1.2. Aunque en este trabajo se utilicen tamaños de paso constantes, es posible que los momentos donde se observan los datos no sean equidistantes. En ese caso el parámetro Δt_i de las funciones (5.1.2) y (5.1.3) serían una muestra de tiempos de observación, en lugar de un valor fijo. Los momentos en los que se observan los datos pueden afectar a los parámetros obtenidos.

Resultados y análisis. Se observó que independientemente de la combinación de parámetros b y σ , el estimador a está sesgado, en especial para valores cercanos a 0. En la Figura 5.1 se muestra el estimador a contra el valor del parámetro, para dos combinaciones diferentes de b y σ .

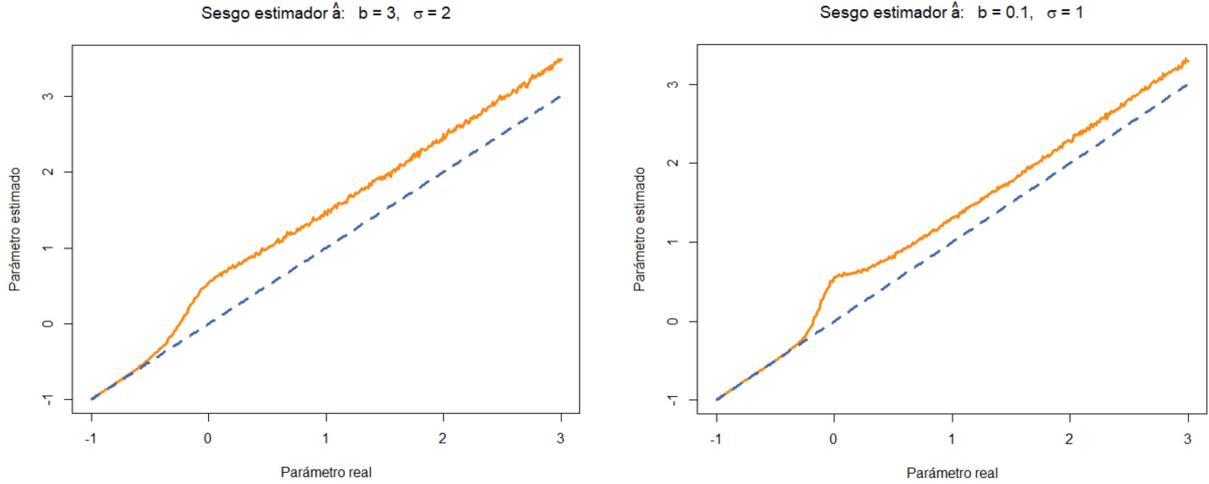


Figura 5.1: Sesgo estimador \hat{a}

En el Teorema 3.4 de [9] se afirma que para $T/n = \Delta t_i$ fijo y $n \rightarrow \infty$

$$\mathbb{E}[\hat{a}] = a + \frac{5 + 2e^{a\Delta t_i} + e^{2a\Delta t_i}}{2n\Delta t_i} + O(n^{-2}).$$

Por lo tanto, para tamaños de paso Δt_i pequeños se puede aproximar $e^{a\Delta t_i}$ con $1 + a\Delta t_i$ y aproximar el parámetro \hat{a} “sin sesgo” como

$$a \approx \frac{n(\mathbb{E}[\hat{a}] - \frac{4}{T})}{n + 2}.$$

¹La prueba puede ser consultada en [9] y [24]

Ver [9] y [24] para pruebas y más detalles sobre el sesgo del estimador.

Se prueba la implementación de los estimadores con los parámetros $a = 1, b = 3$ y $\sigma = 1$, utilizando $X_0 = 0, T = 10$ y $N = 1000$.

$$dX_t = (3 - X_t)dt + dB_t.$$

Se pone a prueba la consistencia de los estimadores. Para ello se simula una trayectoria y se grafica el número de observaciones usadas en la estimación vs. el parámetro estimado. En las Figuras 5.2, 5.3 y 5.4 se observa cómo al incrementar el número de observaciones el estimador se aproxima mejor al parámetro real. Por lo tanto son consistentes. Además se tiene evidencia empírica de que el tamaño de paso $\Delta t_i = 10/1000 = 0.01$ da estimadores consistentes.

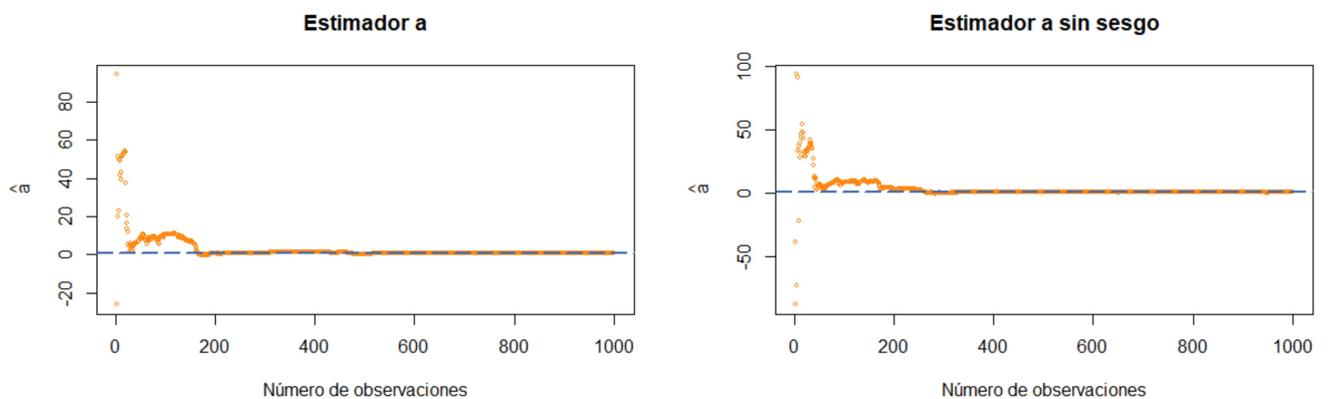


Figura 5.2: Consistencia estimador $a = 1$

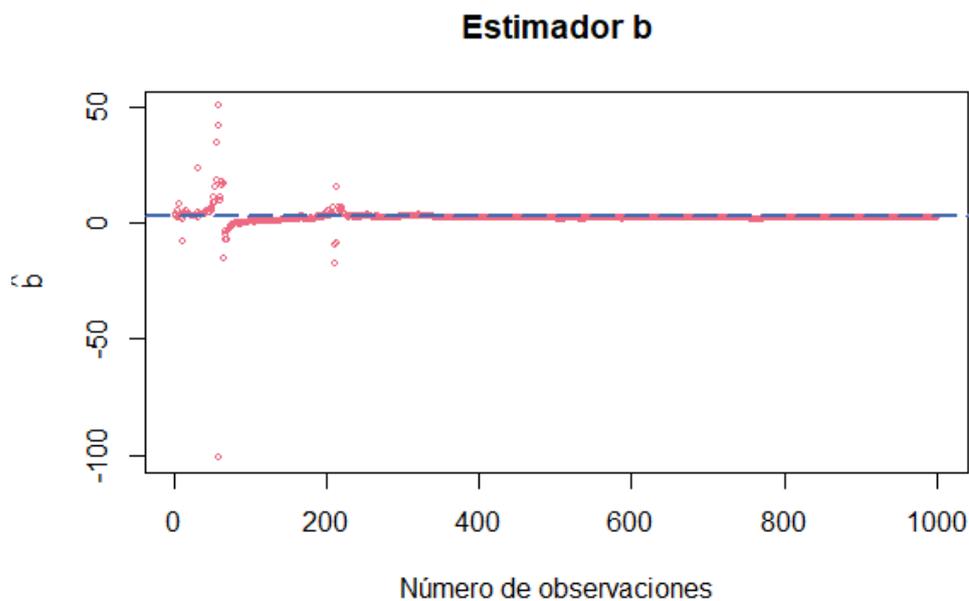


Figura 5.3: Consistencia estimador $b = 3$

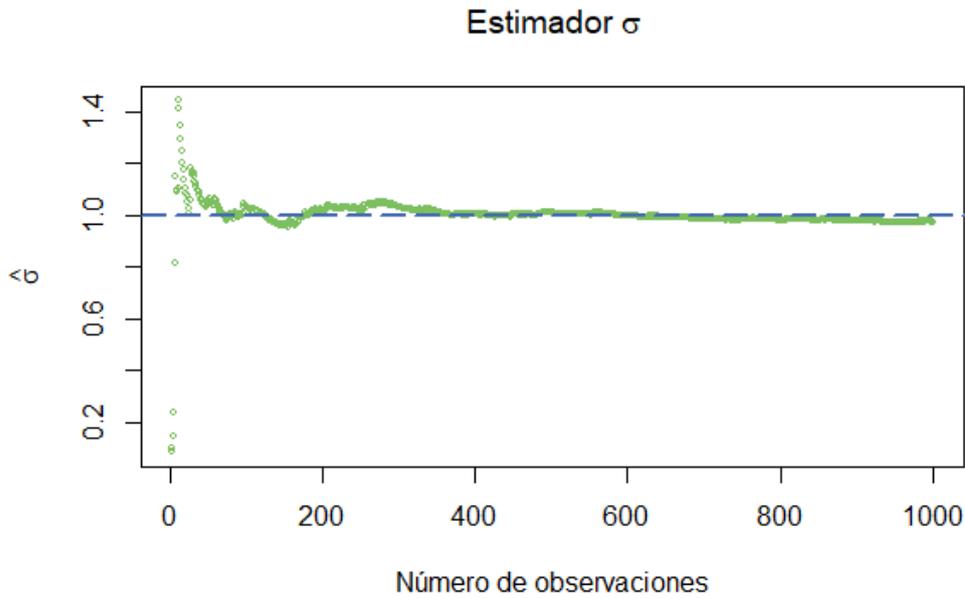


Figura 5.4: Consistencia estimador $\sigma = 1$

En la Tabla 5.1 se muestran los estimadores puntuales, estimadores por intervalos de 95% de confianza y error relativo del estimador puntual. Se usa el estimador \hat{a} con sesgo. La Tabla 5.2 exhibe la misma información pero estimando \hat{a} sin sesgo. Se utilizan 1000 trayectorias diferentes.

	Estimador puntual	Estimador por intervalos	Error relativo
\hat{a}	1.46582	(1.42696, 1.50468)	0.46582
\hat{b}	2.9938	(2.97356, 3.01405)	0.00206
$\hat{\sigma}$	1.00553	(1.00414, 1.00691)	0.00553

Tabla 5.1: Estimadores MLE y error, \hat{a} sesgado

	Estimador puntual	Estimador por intervalos	Error relativo
\hat{a}	1.042	(1.00345, 1.08056)	0.042
\hat{b}	2.9791	(2.9089, 3.0493)	0.007
$\hat{\sigma}$	1.00348	(1.00203, 1.00488)	0.00348

Tabla 5.2: Estimadores MLE y error, \hat{a} sin sesgo

Notamos que al hacer la corrección del sesgo mejora la estimación del parámetro a . Las estimaciones para los parámetros b y σ así como los intervalos de confianza son casi iguales en ambos casos y en ambos se obtienen buenos estimadores.

Al hacer la corrección del sesgo, 7 de los estimadores 1000 para a tomaron valores negativos, ya que sin la corrección eran valores cercanos a cero. Este es un problema pues por definición $a > 0$. Aún así estos estimadores sí fueron considerados para el estimador puntual e intervalos de confianza para no alterar la muestra.

Se verifica que el estimador tenga normalidad multivariada asintótica, es decir

$$\sqrt{n}(\hat{\theta} - \theta) \xrightarrow{d} \mathcal{N}\left(0, (\mathcal{I}(\hat{\theta}))^{-1}\right).$$

Se utilizará una prueba de Mardia² que prueba normalidad multivariada de una muestra por medio de la asimetría y curtosis. Se utiliza la función `mvn()` de la paquetería MVN en R, obteniendo el siguiente resultado para una muestra de 100 estimadores. La prueba en el Código 5.1 da evidencia de que el estimador máximo verosímil para el proceso de Vasicek sigue normalidad asintótica.

Código 5.1: Prueba Mardia: Normalidad estimadores Vasicek

```
> n = 100
> muestra_estimadores = matrix(NA,n,3)
> for(i in 1:n){
+   X = vasicek(a_real,b_real,sigma_real,X0,T,N)
+   muestra_estimadores[i,] = (MLE_vasicek(T_, X)-c(a_real,b_real, sigma_real))*N
+ }
> mvn(data = muestra_estimadores,mvnTest = "mardia")$multivariateNormality
```

	Test	Statistic	p value	Result
1	Mardia Skewness	14.3361165884682	0.158204902893079	YES
2	Mardia Kurtosis	-1.67262837978281	0.0944004764227302	YES
3	MVN	<NA>	<NA>	YES

²Ver [14] y [1] para más información sobre la prueba e implementación computacional.

5.1.2. Verosimilitud del movimiento browniano geométrico

Consideremos ahora $\{X_t\}$ un movimiento browniano geométrico, el cual satisface la siguiente ecuación diferencial

$$dX_t = \mu X_t dt + \sigma X_t dB_t, \quad (5.1.6)$$

$$\mu \in \mathbb{R}, \sigma > 0, \quad X_0 = x_0 > 0.$$

En este caso $\theta = (\mu, \sigma)$.

Definición 5.1.2. Se definen los *log-rendimientos* de X_t como

$$Y_{t_i} = \ln(X_{t_i}) - \ln(X_{t_{i-1}}) = \ln\left(\frac{X_{t_i}}{X_{t_{i-1}}}\right).$$

Por (3.3.12) sabemos que

$$Y_{t_i} \sim \mathcal{N}\left(\left(\mu - \frac{\sigma^2}{2}\right) \Delta t_i, \sigma^2 \Delta t_i\right)$$

Proposición 5.1.3. Las funciones de log-verosimilitud de $\{X_t\}$ y $\{Y_t\}$ estan dadas por

$$l_n^X(\theta) = - \sum_{i=1}^n \left(\frac{1}{2} \ln(2\pi \Delta t_i \sigma^2) + \ln(X_{t_i}) + \frac{(\ln(X_{t_i}) - \ln(X_{t_{i-1}}) - (\mu - \frac{\sigma^2}{2}) \Delta t_i)^2}{2\sigma^2 \Delta t_i} \right), \quad (5.1.7)$$

$$l_n^Y(\theta) = -n \ln(\sigma) - \sum_{i=1}^n \left(\frac{1}{2} \ln(2\pi \Delta t_i) + \frac{(Y_{t_i} - (\mu - \frac{\sigma^2}{2}) \Delta t_i)^2}{2\sigma^2 \Delta t_i} \right). \quad (5.1.8)$$

Demostración. Se demuestra primero (5.1.7) calculando $\ln(p_\theta(\Delta t_i, X_{t_i} | X_{t_{i-1}}))$.

$$\begin{aligned} \ln(p_\theta(\Delta t_i, X_{t_i} | X_{t_{i-1}})) &= \ln \left(\frac{1}{\sigma X_{t_i} \sqrt{2\pi \Delta t_i}} \exp \left(- \frac{(\ln(X_{t_i}) - \ln(X_{t_{i-1}}) - (\mu - \frac{\sigma^2}{2}) \Delta t_i)^2}{2\sigma^2 \Delta t_i} \right) \right) \\ &= \ln((\sqrt{2\sigma^2 \pi \Delta t_i})^{-1}) - \ln(X_{t_i}) - \frac{(\ln(X_{t_i}) - \ln(X_{t_{i-1}}) - (\mu - \frac{\sigma^2}{2}) \Delta t_i)^2}{2\sigma^2 \Delta t_i} \\ &= -\frac{1}{2} \ln(2\pi \sigma^2 \Delta t_i) - \left(\ln(X_{t_i}) + \frac{(\ln(X_{t_i}) - \ln(X_{t_{i-1}}) - (\mu - \frac{\sigma^2}{2}) \Delta t_i)^2}{2\sigma^2 \Delta t_i} \right). \end{aligned}$$

Sumando sobre $i = 1, \dots, n$.

$$\sum_{i=1}^n \ln(p_\theta(\Delta t_i, X_{t_i} | X_{t_{i-1}})) = \sum_{i=1}^n \left(-\frac{1}{2} \ln(2\pi \sigma^2 \Delta t_i) - \left(\ln(X_{t_i}) + \frac{(\ln(X_{t_i}) - \ln(X_{t_{i-1}}) - (\mu - \frac{\sigma^2}{2}) \Delta t_i)^2}{2\sigma^2 \Delta t_i} \right) \right)$$

Por lo tanto

$$l_n^X(\theta) = \sum_{i=1}^n \left(-\frac{1}{2} \ln(2\pi\sigma^2\Delta t_i) - \left(\ln(X_{t_i}) + \frac{\left(\ln(X_{t_i}) - \ln(X_{t_{i-1}}) - \left(\mu - \frac{\sigma^2}{2}\right)\Delta t_i \right)^2}{2\sigma^2\Delta t_i} \right) \right).$$

Se demuestra (5.1.8). Por 3.3.12 se tiene que

$$p_\theta(\Delta t_i, Y_{t_i}|Y_{t_{i-1}}) = \frac{1}{\sqrt{2\pi\sigma^2\Delta t_i}} \exp \left(-\frac{\left(Y_{t_i} - \left(\mu - \frac{\sigma^2}{2}\right)\Delta t_i \right)^2}{2\sigma^2\Delta t_i} \right).$$

Se calcula $\ln(p_\theta(\Delta t_i, Y_{t_i}|Y_{t_{i-1}}))$ y después se suma sobre $i = 1, \dots, n$.

$$\begin{aligned} \ln(p_\theta(\Delta t_i, Y_{t_i}|Y_{t_{i-1}})) &= \ln \left(\frac{1}{\sqrt{2\pi\sigma^2\Delta t_i}} \exp \left(-\frac{\left(Y_{t_i} - \left(\mu - \frac{\sigma^2}{2}\right)\Delta t_i \right)^2}{2\sigma^2\Delta t_i} \right) \right) \\ &= -\ln(\sqrt{2\pi\Delta t_i}) - \ln(\sigma) - \frac{\left(Y_{t_i} - \left(\mu - \frac{\sigma^2}{2}\right)\Delta t_i \right)^2}{2\sigma^2\Delta t_i}, \end{aligned}$$

$$\begin{aligned} \sum_{i=1}^n (\ln(p_\theta(\Delta t_i, Y_{t_i}|Y_{t_{i-1}}))) &= \sum_{i=1}^n \left(-\ln(\sqrt{2\pi\Delta t_i}) - \ln(\sigma) - \frac{\left(Y_{t_i} - \left(\mu - \frac{\sigma^2}{2}\right)\Delta t_i \right)^2}{2\sigma^2\Delta t_i} \right) \\ &= -n \ln(\sigma) - \sum_{i=1}^n \left(\frac{1}{2} \ln(2\pi\Delta t_i) + \frac{\left(Y_{t_i} - \left(\mu - \frac{\sigma^2}{2}\right)\Delta t_i \right)^2}{2\sigma^2\Delta t_i} \right). \end{aligned}$$

Por lo tanto

$$l_n^Y(\theta) = -n \ln(\sigma) - \sum_{i=1}^n \left(\frac{1}{2} \ln(2\pi\Delta t_i) + \frac{\left(Y_{t_i} - \left(\mu - \frac{\sigma^2}{2}\right)\Delta t_i \right)^2}{2\sigma^2\Delta t_i} \right).$$

Observación 5.1.3. Las primeras derivadas parciales de $l_n^X(\theta)$ y $l_n^Y(\theta)$ son

$$\begin{aligned}\frac{\partial l_n^X(\theta)}{\partial \mu} &= -\sum_{i=1}^n \frac{2\mu\Delta t_i^2 \ln^2(X_{t_{i-1}}) - \sigma^2\Delta t_i^2 \ln^2(X_{t_{i-1}}) - 2\Delta t_i \ln(X_{t_i}) \ln(X_{t_{i-1}})}{2\sigma^2\Delta t_i}, \\ \frac{\partial l_n^X(\theta)}{\partial \sigma} &= -\frac{n}{\sigma} - \sum_{i=1}^n \left(-\frac{\left(\ln(X_{t_i}) - \Delta t_i \left(\mu - \frac{\sigma^2}{2}\right) \ln(X_{t_{i-1}})\right)^2}{\sigma^3\Delta t_i} \right. \\ &\quad \left. + \frac{-2\mu\sigma\Delta t_i^2 \ln^2(X_{t_{i-1}}) + \sigma^3\Delta t_i^2 \ln^2(X_{t_{i-1}}) + 2\sigma\Delta t_i \ln(X_{t_i}) \ln(X_{t_{i-1}})}{2\sigma^2\Delta t_i} \right), \\ \frac{\partial l_n^Y(\theta)}{\partial \mu} &= \sum_{i=1}^n \frac{Y_{t_i} - \left(\mu - \frac{\sigma^2}{2}\right) \Delta t_i}{\sigma^2}, \\ \frac{\partial l_n^Y(\theta)}{\partial \sigma} &= -\frac{n}{\sigma} - \sum_{i=1}^n \left(\frac{Y_{t_i} - \left(\mu - \frac{\sigma^2}{2}\right) \Delta t_i}{\sigma} - \frac{\left(Y_{t_i} - \left(\mu - \frac{\sigma^2}{2}\right) \Delta t_i\right)^2}{\sigma^3\Delta t_i} \right).\end{aligned}$$

Las segundas derivadas parciales de $l_n^X(\theta)$ están dadas por

$$\begin{aligned}\frac{\partial^2 l_n^X(\theta)}{\partial \mu^2} &= -\sum_{i=1}^n \frac{\Delta t_i \ln^2(X_{t_{i-1}})}{\sigma^2}, \\ \frac{\partial^2 l_n(\theta)^X}{\partial \sigma^2} &= \frac{n}{\sigma^2} - \sum_{i=1}^n \frac{3\left(\ln(X_{t_i}) - \Delta t_i \left(\mu - \frac{\sigma^2}{2}\right) \ln(X_{t_{i-1}})\right)^2}{\sigma^4\Delta t_i} \\ &\quad - \frac{2\left(-2\mu\sigma\Delta t_i^2 \ln^2(X_{t_{i-1}}) + \sigma^3\Delta t_i^2 \ln^2(X_{t_{i-1}}) + 2\sigma\Delta t_i \ln(X_{t_i}) \ln(X_{t_{i-1}})\right)}{\sigma^3\Delta t_i} \\ &\quad + \frac{-2\mu\Delta t_i^2 \ln^2(X_{t_{i-1}}) + 3\sigma^2\Delta t_i^2 \ln^2(X_{t_{i-1}}) + 2\Delta t_i \ln(X_{t_i}) \ln(X_{t_{i-1}})}{2\sigma^2\Delta t_i}.\end{aligned}$$

Proposición 5.1.4. ³ Los estimadores máximo verosímiles del movimiento browniano geométrico están dados por

$$\begin{aligned}\hat{\mu} &= \frac{\bar{Y}}{\Delta t_i} + \frac{\hat{\sigma}^2}{2}, \\ \hat{\sigma}^2 &= \frac{1}{n} \sum_{i=1}^n \frac{1}{\Delta t_i} (Y_{t_i} - \bar{Y})^2.\end{aligned}$$

Donde $\bar{Y} = \frac{1}{n} \sum_{i=1}^n Y_{t_i}$.

Resultados y análisis. Utilizando $\mu = 0.5$ y $\sigma = 0.2$ con $\Delta t_i = 0.001$ se prueba la consistencia del estimador y la normalidad asintótica.

$$dX_t = 0.5X_t dt + 0.2X_t dB_t.$$

Las Figuras 5.5 y 5.6 reflejan cómo al incrementar el número de observaciones el estimador se aproxima al parámetro real, lo cual es evidencia de que el estimador es consistente y de que el tamaño de paso $\Delta t_i = 0.01$ es adecuado.

³La prueba puede ser consultada en [16] y [4]

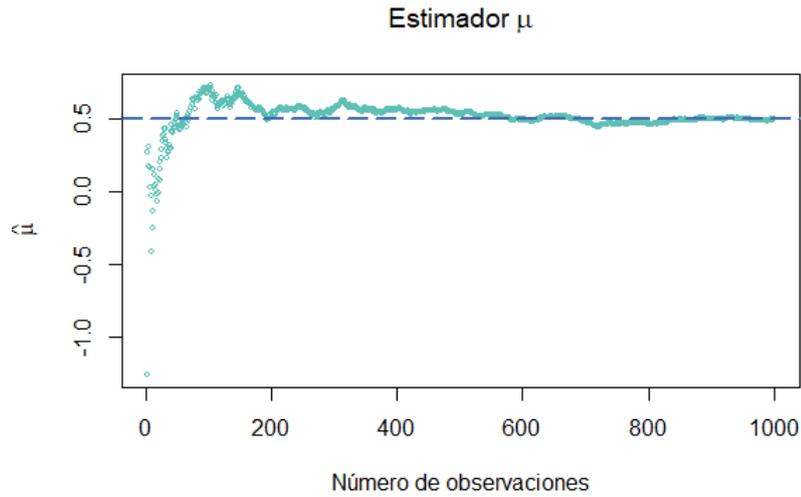


Figura 5.5: Consistencia estimador $\mu = 0.5$

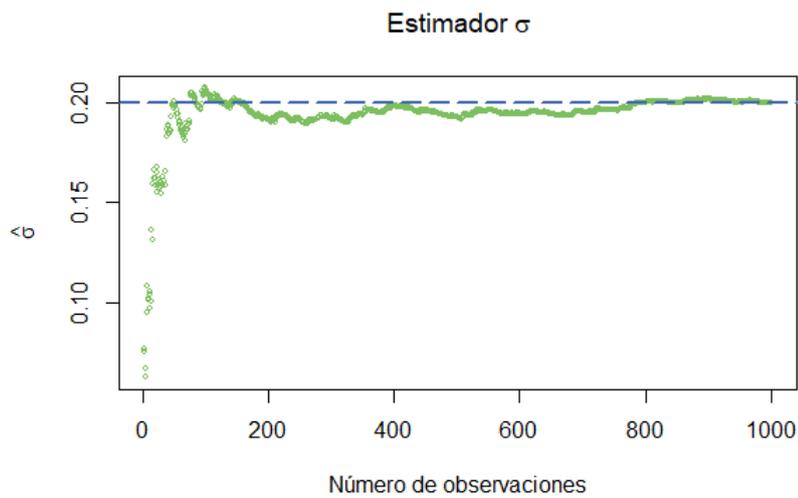


Figura 5.6: Consistencia estimador $\sigma = 0.2$

En la Tabla 5.3 se muestran los estimadores puntuales y por intervalos usando un 95 % de confianza. Para obtenerlos se utilizaron 1000 trayectorias diferentes.

	Estimador puntual	Estimador por intervalos	Error relativo
μ	0.501288	(0.49736, 0.50521)	0.002575
σ	0.199866	(0.19959, 0.20014)	0.000669

Tabla 5.3: Estimadores MLE y error relativo

Los estimadores obtenidos son precisos con un error relativo pequeño.

Al igual que con el proceso de Vasicek se prueba la normalidad asintótica usando la prueba de Mardia en R. Los resultados de la prueba en el Código 5.2 indican que

hay normalidad asintótica en el estimador máximo verosímil para el movimiento browniano geométrico.

Código 5.2: Prueba Mardia: Normalidad estimadores MBG

```
> n = 100
> muestra_estimadores = matrix(NA,n,2)
> for(i in 1:n){
+   X = MBG(mu_real,sigma_real,X0,T,N)
+   muestra_estimadores[i,] = (MLE_gbm(T,X)-c(mu_real,sigma_real))*N
+ }
> mvn(data = muestra_estimadores,mvnTest = "mardia")$multivariateNormality
```

	Test	Statistic	p value	Result
1	Mardia Skewness	4.29178960713916	0.367954335199054	YES
2	Mardia Kurtosis	-0.360998538636744	0.718100539514371	YES
3	MVN	<NA>	<NA>	YES

5.2. Algoritmo EM

En la sección anterior se utiliza la metodología del estimador máximo verosímil. En ella se asume que la información es completa. Es decir, que los datos de la muestra sobre la cual se hace inferencia son suficientes para que haya convergencia entre el estimador y el parámetro real.

En la práctica, los datos y muestras con las que se trabaje pueden no tienen información suficiente para utilizar el estimador máximo verosímil. En ese caso se dice que hay información incompleta.

El algoritmo EM sirve para hacer estimación de parámetros en casos donde hay información incompleta, en particular para difusiones cuya transición no es conocida. Este es un algoritmo que maximiza la verosimilitud iterando sobre dos pasos: esperanza y maximización.

Notación. Consideremos \mathbf{x} un vector aleatorio proveniente de una variable aleatoria X , con datos observados y datos faltantes. La densidad de los datos completos es $f(\mathbf{x}; \theta)$. Sea \mathbf{y} el vector de datos observados con densidad $f_Y(\mathbf{y}; \theta)$, y \mathbf{z} el vector de datos faltantes con densidad condicional a \mathbf{y} $f_{Z|Y}(\mathbf{z}|\mathbf{y}; \theta)$. Entonces $\mathbf{x} = (\mathbf{y}, \mathbf{z})$ y

$$f(\mathbf{x}; \theta) = f_Y(\mathbf{y}; \theta)f_{Z|Y}(\mathbf{z}|\mathbf{y}; \theta).$$

Sea $l(\theta; \mathbf{x})$ la log-verosimilitud de los datos completos. Entonces la log-verosimilitud de los datos observados es

$$l_Y(\theta; \mathbf{y}) = l(\theta; \mathbf{x}) - \ln(f_{Z|Y}(\mathbf{z}|\mathbf{y}; \theta)).$$

El primero paso para maximizar $l(\theta; \mathbf{x})$ es reemplazar los datos faltantes con el valor esperado dado los datos observados. En el segundo paso se busca un parámetro que maximice la log-verosimilitud obtenida en el paso anterior.

Algoritmo 3 : EM

- 1: Condiciones iniciales: $\theta^{(0)}$, $k = 0$.
- 2: **Paso E:** Calcular $Q(\theta; \theta^{(k)})$ donde

$$Q(\theta; \theta^{(k)}) = \mathbb{E}_{\theta^{(k)}}[l(\theta; \mathbf{x})|\mathbf{y}].$$

- 3: **Paso M:** Obtener $\theta^{(k+1)}$ tal que maximice $Q(\theta; \theta^{(k)})$. Es decir

$$Q(\theta^{(k+1)}; \theta^{(k)}) \geq Q(\theta; \theta^{(k)}) \quad \forall \theta \in \Theta.$$

- 4: Parar si

$$l_Y(\theta^{(k+1)}; \mathbf{y}) - l_Y(\theta^{(k)}; \mathbf{y}) < \delta$$

para δ suficientemente pequeña. En otro caso, $k = k + 1$ y volver al Paso E.

Observación 5.2.1. El algoritmo se simplifica cuando $f(\mathbf{x}; \theta)$ pertenece a la familia exponencial de forma canónica. En ese caso se pueden reescribir su densidad y log-verosimilitud como

$$\begin{aligned} f(\mathbf{x}; \theta) &= h(\mathbf{x}) \exp [\theta^T t(\mathbf{x}) - A(\theta)], \\ l(\theta; \mathbf{x}) &= \ln(h(\mathbf{x})) + \theta^T t(\mathbf{x}) - A(\theta). \end{aligned}$$

Por lo tanto

$$Q(\theta; \theta^{(k)}) = \mathbb{E}_{\theta^{(k)}}[\ln(h(\mathbf{x}))|\mathbf{y}] + \theta^T \mathbb{E}_{\theta^{(k)}}[t(\mathbf{x})|\mathbf{y}] - A(\theta).$$

Por ser de la familia exponencial se cumple que

$$\mathbb{E}[t(\mathbf{x}); \theta] = \frac{\partial}{\partial \theta} A(\theta).$$

En el Paso M se maximiza respecto a θ . Notamos que $\mathbb{E}_{\theta^{(k)}}[\ln(h(x))|Y]$.

$$\begin{aligned} \frac{\partial}{\partial \theta} Q(\theta; \theta^{(k)}) &= \mathbb{E}_{\theta^{(k)}}[t(\mathbf{x})|\mathbf{y}] - \frac{\partial}{\partial \theta} A(\theta) \\ &= \mathbb{E}_{\theta^{(k)}}[t(\mathbf{x})|\mathbf{y}] - \mathbb{E}_{\theta}[t(\mathbf{x})]. \end{aligned} \quad (5.2.1)$$

El paso E se reduce a calcular $\mathbb{E}_{\theta^{(k)}}[t(\mathbf{x})|\mathbf{y}]$ y el paso M a encontrar $\theta^{(k+1)}$ tal que (5.2.1) sea igual a 0.

Algoritmo 4 : EM Familia Exponencial

- 1: Condiciones iniciales: $\theta^{(0)}$, $k = 0$.
- 2: **Paso E:** Calcular $t^{(k)}$ donde

$$t^{(k)} = \mathbb{E}_{\theta^{(k)}}[t(\mathbf{x})|\mathbf{y}].$$

- 3: **Paso M:** Obtener $\theta^{(k+1)}$ tal que resuelva la ecuación

$$t^{(k)} = \mathbb{E}_{\theta}[t(\mathbf{x})].$$

- 4: Parar si

$$l_Y(\theta^{(k+1)}; \mathbf{y}) - l_Y(\theta^{(k)}; \mathbf{y}) < \delta$$

para δ suficientemente pequeña. En otro caso, $k = k + 1$ y volver al Paso E.

Teorema 5.2.1. ⁴ Sea $\theta^{(k)}$, $k = 0, 1, \dots$ el parámetro estimado en la k -ésima iteración del algoritmo EM, tal que:

1. La sucesión $\{l_Y(\theta^{(k)}; \mathbf{y})\}_{k=0}^{\infty}$ es acotada.
2. $Q(\theta^{(k+1)}; \theta^{(k)}) - Q(\theta^{(k)}; \theta^{(k)}) \geq \lambda(\theta^{(k+1)} - \theta^{(k)})(\theta^{(k+1)} - \theta^{(k)})^T$ para algún escalar $\lambda > 0$ y para todo k .

Entonces la sucesión $\{\theta^{(k)}\}_{k=0}^{\infty}$ converge a algún $\theta^* \in \Theta$.

⁴Ver [6]

5.2.1. Algoritmo EMMC

En ocasiones el cálculo de la esperanza en el Paso E del algoritmo EM puede ser muy complicado o no existe una forma cerrada para la función $Q(\theta; \theta^{(k)})$. Una solución a este problema es reemplazar la esperanza por el estimador de Monte Carlo. La función $Q(\theta; \theta^{(k)})$ puede ser reescrita de la siguiente manera

$$\begin{aligned} Q(\theta; \theta^{(k)}) &= \mathbb{E}_{\theta^{(k)}}[l(\theta; \mathbf{x})|\mathbf{y}] \\ &= \mathbb{E}_{\theta^{(k)}}[\ln(f(\mathbf{y}, \mathbf{z}; \theta))|Y] \\ &= \int \ln(f(y, z; \theta))f_{Z|Y}(\mathbf{z}|\mathbf{y}; \theta^{(k)})dz. \end{aligned} \quad (5.2.2)$$

En cada iteración se genera una muestra de tamaño m a partir de la densidad $f_{Z|Y}(\mathbf{z}|\mathbf{y}; \theta^{(k)})$.

$$\mathbf{z}^{(k)} = (z_1^{(k)}, \dots, z_m^{(k)}).$$

Entonces el estimador de Monte Carlo para (5.2.2) es

$$\hat{Q}(\theta; \theta^{(k)}) = \frac{1}{m} \sum_{i=1}^m l(\theta; \mathbf{y}, z_i^{(k)}).$$

En el Paso M se maximiza la función $\hat{Q}(\theta; \theta^{(k)})$. Así como en el algoritmo EM, para el EMMC también se tiene la convergencia del estimador a un parámetro. Los resultados pueden ser consultados en [15].

Algoritmo 5 : EMMC

1: Condiciones iniciales: $\theta^{(0)}$, $k = 0$.

2: **Paso MC E:**

a) Simular el vector aleatorio $\mathbf{z}^{(k)} = (z_1^{(k)}, \dots, z_m^{(k)})$ de la función de densidad $f_{Z|Y}(\mathbf{z}|\mathbf{y}; \theta^{(k)})$.

b) Calcular $\hat{Q}(\theta; \theta^{(k)})$ donde

$$\hat{Q}(\theta; \theta^{(k)}) = \frac{1}{m} \sum_{i=1}^m l(\theta; \mathbf{y}, z_i^{(k)}).$$

3: **Paso M:** Obtener $\theta^{(k+1)}$ tal que

$$\hat{Q}(\theta^{(k+1)}; \theta^{(k)}) \geq \hat{Q}(\theta; \theta^{(k)}) \quad \forall \theta \in \Theta.$$

4: Parar si

$$l_Y(\theta^{(k+1)}; \mathbf{y}) - l_Y(\theta^{(k)}; \mathbf{y}) < \delta$$

para δ suficientemente pequeña. En otro caso, $k = k + 1$ y volver al Paso MC E.

5.2.2. Algoritmo EME

Este algoritmo es un caso particular del algoritmo EMMC donde $m = 1$. En el Paso E se simula solo una variable aleatoria $z^{(k)}$ a partir de la densidad $f_{Z|Y}(z|\mathbf{y}; \theta^{(k)})$. El estimador de Monte Carlo es

$$\hat{Q}(\theta; \theta^{(k)}) = l(\theta; \mathbf{y}, z^{(k)}).$$

Algoritmo 6 : EME

- 1: Condiciones iniciales: $\theta^{(0)}$, $k = 0$.
- 2: **Paso E E:** Simular la variable aleatoria $z^{(k)}$ utilizando la función de distribución $f_{Z|Y}(z|\mathbf{y}; \theta^{(k)})$. Calcular $\hat{Q}(\theta; \theta^{(k)})$ donde

$$\hat{Q}(\theta; \theta^{(k)}) = l(\theta; \mathbf{y}, z^{(k)}).$$

- 3: **Paso M:** Obtener $\theta^{(k+1)}$ tal que

$$\hat{Q}(\theta^{(k+1)}; \theta^{(k)}) \geq \hat{Q}(\theta; \theta^{(k)}) \quad \forall \theta \in \Theta.$$

- 4: Parar si

$$l_Y(\theta^{(k+1)}; \mathbf{y}) - l_Y(\theta^{(k)}; \mathbf{y}) < \delta$$

para δ suficientemente pequeña. En otro caso, $k = k + 1$ y volver al Paso MC E.

5.2.3. Algoritmo EME para procesos de difusión

Consideremos $\{X_t\}$ un proceso de difusión definido como en (5.1.1) y $\mathbf{X}_N = \{X_{t_i}\}_{i=0}^N$ una discretización del proceso, la cual representa a los datos completos. Sea $\mathbf{X}_n = \{X_{t_j}\}_{j=0}^n$ una muestra de \mathbf{X}_N , donde $n < N$. Esta submuestra representa a los datos observados.

Se utiliza el algoritmo anterior para estimar los parámetros de los procesos de Vasicek y movimiento browniano geométrico. Esto se puede hacer ya que son procesos de difusión para los cuales se conocen las probabilidades de transición

En el paso E se generan puentes estocásticos utilizando los datos observados y los parámetros $\theta^{(k)}$. Sea

$$Z_i^{(k)} = (X_{t_{i-1}}, t_{i-1}, X_{t_i}, t_i), \quad i = 1, \dots, n$$

la discretización del puente generado entre los datos observados $X_{t_{i-1}}$ y X_{t_i} . Esta consta de $\frac{N}{n} + 1$ datos. Se denota por $Z^{(k)}$ al puente completo entre X_{t_0} y X_{t_n} , generado por los $Z_i^{(k)}$ puentes en la k -ésima iteración. Este puente tiene $N + 1$ datos. Así, se reconstruye una trayectoria que simula los datos completos.

En el paso M se maximiza la verosimilitud, calculando los estimadores máximo verosímiles con los datos del puente generado en el paso anterior.

Para probar los algoritmos se utilizará $N = 1000$ y $n = 100$, tomando 1 de cada 10 observaciones de los datos completos.

EM proceso Vasicek

Algoritmo 7 : EME Vasicek

- 1: Condiciones iniciales: $\theta^{(0)}$, $k = 0$.
- 2: **Paso E:** Generar los puentes de difusión $\{Z_i^{(k)} = (X_{t_{i-1}}, t_{i-1}, X_{t_i}, t_i)\}_{i=1}^n$ con los parámetros $\theta^{(k)}$ y calcular $\hat{Q}(\theta; \theta^{(k)})$.
- 3: **Paso M:** Actualizar los estimadores

$$\hat{a}^{(k+1)} = -\frac{1}{\Delta t_i} \ln \left(\frac{N \sum_{i=1}^N Z_{t_i}^{(k)} Z_{t_{i-1}}^{(k)} - \sum_{i=1}^N Z_{t_i}^{(k)} \sum_{i=1}^N Z_{t_{i-1}}^{(k)}}{N \sum_{i=1}^N (Z_{t_{i-1}}^{(k)})^2 - \left(\sum_{i=1}^N Z_{t_{i-1}}^{(k)} \right)^2} \right),$$

$$\hat{b}^{(k+1)} = \frac{1}{N(1 - e^{-\hat{a}\Delta t_i})} \left(\sum_{i=1}^N Z_{t_i}^{(k)} - e^{-\hat{a}\Delta t_i} \sum_{i=1}^N Z_{t_{i-1}}^{(k)} \right),$$

$$\hat{\sigma}^2{}^{(k+1)} = \frac{2\hat{a}}{N(1 - e^{-2\hat{a}\Delta t_i})} \sum_{i=1}^N \left(Z_{t_i}^{(k)} - Z_{t_{i-1}}^{(k)} e^{-\hat{a}\Delta t_i} - \hat{b}(1 - e^{-\hat{a}\Delta t_i}) \right)^2.$$

- 4: Parar si $k = K$. En otro caso, $k = k + 1$ y volver al Paso MC E.
-

Pruebas y resultados

Se prueba el algoritmo usando los parámetros $a = 1$, $b = 3$ y $\sigma = 1$ con una condición inicial $X_0 = 3$. Los parámetros iniciales son los estimadores máximo verosímiles para la información incompleta. Se itera $K = 1000$ veces, promediando las últimas 100 iteraciones para obtener el estimador.

A partir de 1000 trayectorias, calculan los estimadores puntuales y por intervalos usando el algoritmo EM. Se calcula el error relativo del estimador EM respecto al parámetro real. También el error relativo al parámetro obtenido como estimador puntual usando máxima verosimilitud (ver Tablas 5.1 y 5.2).

Los estimadores y errores se muestran en las Tablas 5.4 y 5.5. Los estimadores obtenidos para las 1000 trayectorias, así como el cálculo de los errores, pueden ser consultados en el repositorio de Github en los archivos Estimadores_Vasicek.xlsx y Estimadores_Vasicek_sin_sesgo.xlsx.

	Estimador puntual	Estimador por intervalos	Error relativo	Error relativo MLE
\hat{a}	1.32565	(1.28454, 1.36675)	0.32565	0.09563
\hat{b}	2.99239	(2.9717, 3.1031)	0.00254	0.00047
$\hat{\sigma}$	0.93789	(0.9327, 0.9431)	0.06210	0.06726

Tabla 5.4: Estimadores EM y errores, \hat{a} sesgado

	Estimador puntual	Estimador por intervalos	Error relativo	Error relativo MLE
\hat{a}	0.83853	(0.79676, 0.88030)	0.16147	0.19527
\hat{b}	3.01179	(2.84746, 3.17613)	0.00393	0.01097
$\hat{\sigma}$	0.90719	(0.90224, 0.91214)	0.09281	0.09596

Tabla 5.5: Estimadores EM y errores, \hat{a} sin sesgo

Una primera observación a partir de los resultados es que en ambos casos el estimador puntual EM del parámetro a es menor que el estimador puntual MLE, y que la diferencia entre EM y MLE es mayor cuando se corrige el sesgo. Esto se puede explicar por la correlación que hay entre el estimador MLE y EM. Los coeficientes de correlación son $\rho = 0.898$ y $\rho = 0.917$ para los estimadores sesgados y corregidos, respectivamente.

En las Figuras 5.7 y 5.8 se muestran los gráficos de dispersión de los estimadores obtenidos para cada una de las 1000 trayectorias simuladas. En azul claro se tiene la línea de regresión y en azul oscuro la recta identidad.

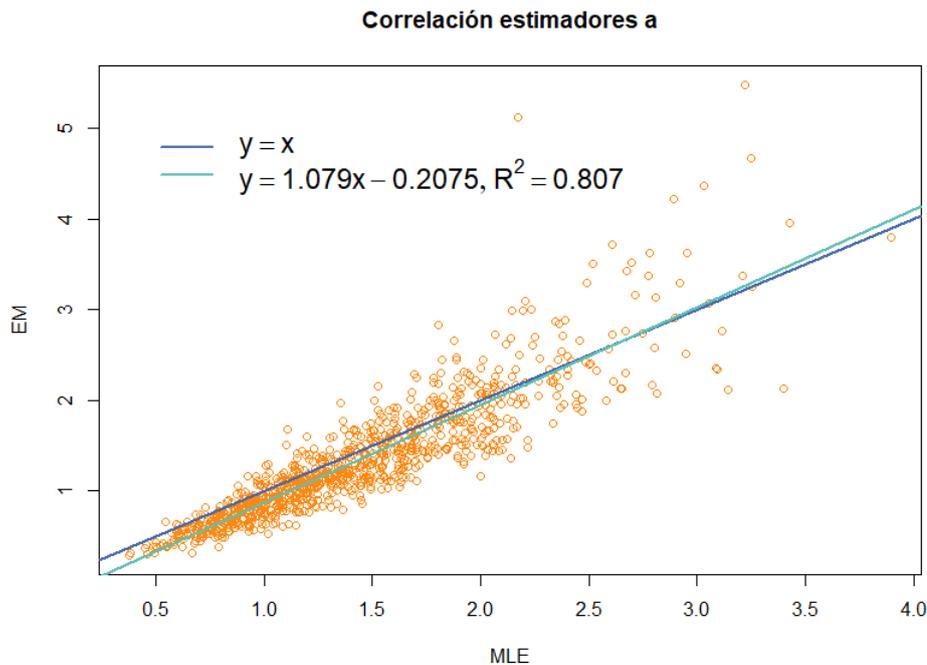


Figura 5.7: Regresión MLE \sim EM: parámetro a

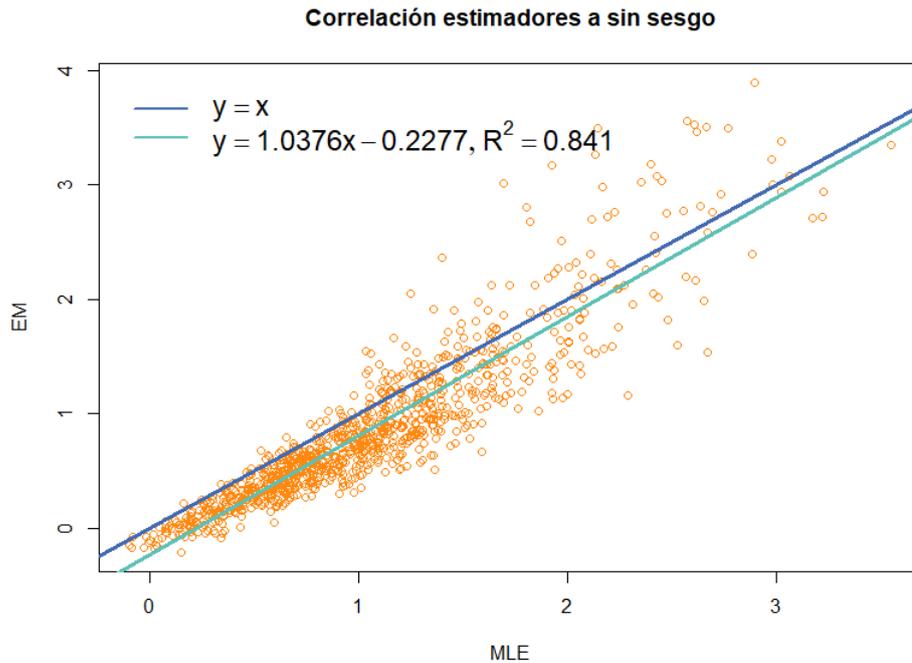


Figura 5.8: Regresión MLE \sim EM: parámetro a sin sesgo

La línea de regresión está por debajo de la recta identidad. Es decir que, por lo general, el estimador MLE para una trayectoria es mayor que el estimador EM.

El mismo resultado se observa con el parámetro σ . Los coeficientes de correlación son $\rho = 0.258$ para el caso sesgado y $\rho = 0.328$ en el otro caso. En las Figuras 5.9 5.10 se muestran los gráficos de dispersión entre ambos estimadores y las líneas de regresión en azul claro.

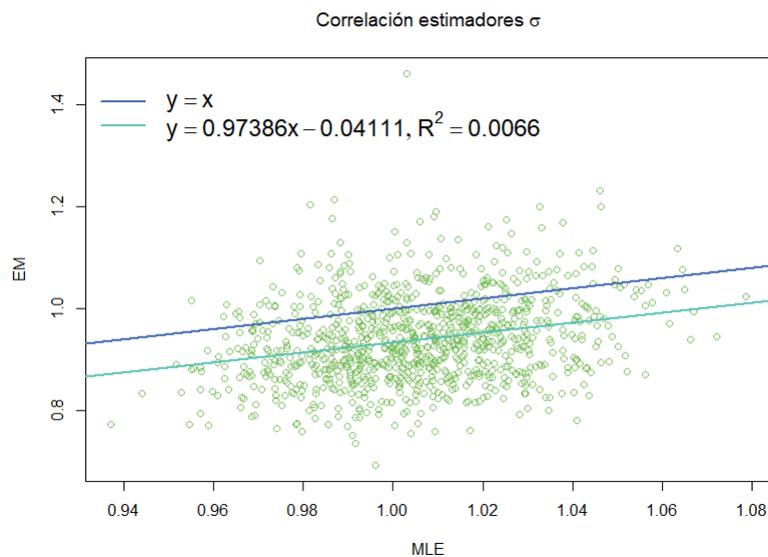


Figura 5.9: Regresión MLE \sim EM: parámetro σ

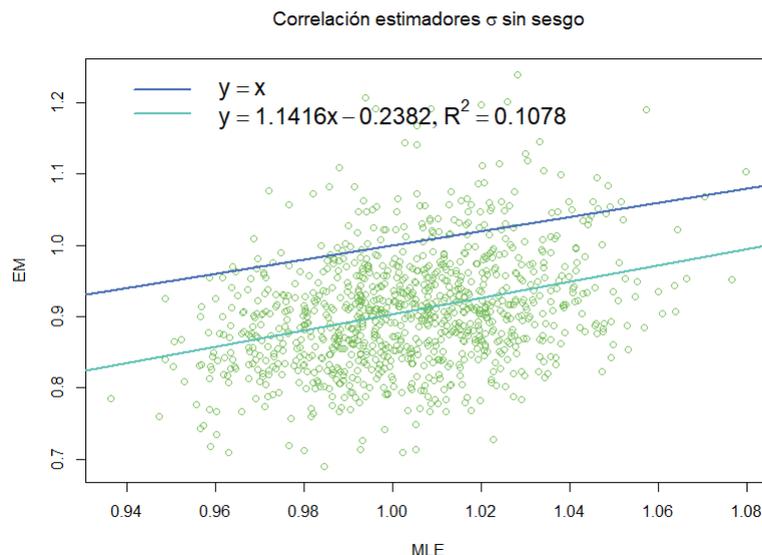


Figura 5.10: Regresión $\text{MLE} \sim \text{EM}$: parámetro σ sin sesgo

De nuevo la recta de regresión se encuentra por debajo de la recta identidad, llegando a una conclusión análoga a la del parámetro a . En este caso la conclusión no es muy fuerte, pues tanto los coeficientes de correlación como los de determinación son valores pequeños. Para el estimador b se obtienen estimadores buenos con un error relativo cercano a 0.

En algunos casos el algoritmo no tuvo convergencia, generando estimadores que tendían a infinito. La divergencia se dio en 23 de 1000 trayectorias para el caso donde a está sesgado, y 6 de 1000 trayectorias en el otro caso. Esto sucedía en trayectorias cuya condición inicial (MLE usando información incompleta) para el parámetro a tomaba valores grandes. En la Tabla 5.6 se muestra la condición inicial en promedio para las trayectorias con EM divergente.

Promedio valor inicial	a	b	σ
a sesgado	3.8448	3.0358	1.0885
a sin sesgo	3.7466	2.8510	1.0370

Tabla 5.6: Trayectorias con EM divergente

Por último, en las Figuras 5.11, 5.12, 5.13 y 5.14 se muestran los estimadores obtenidos en cada iteración para 4 trayectorias con diferentes condiciones iniciales. Los estimadores se aproximan al parámetro independientemente de la condición inicial.

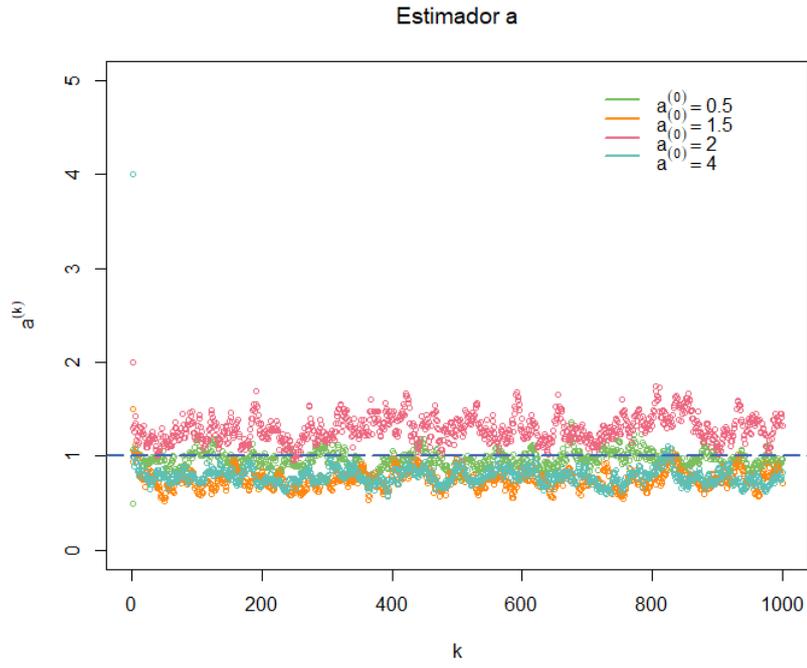


Figura 5.11: Estimador EM $a = 1$

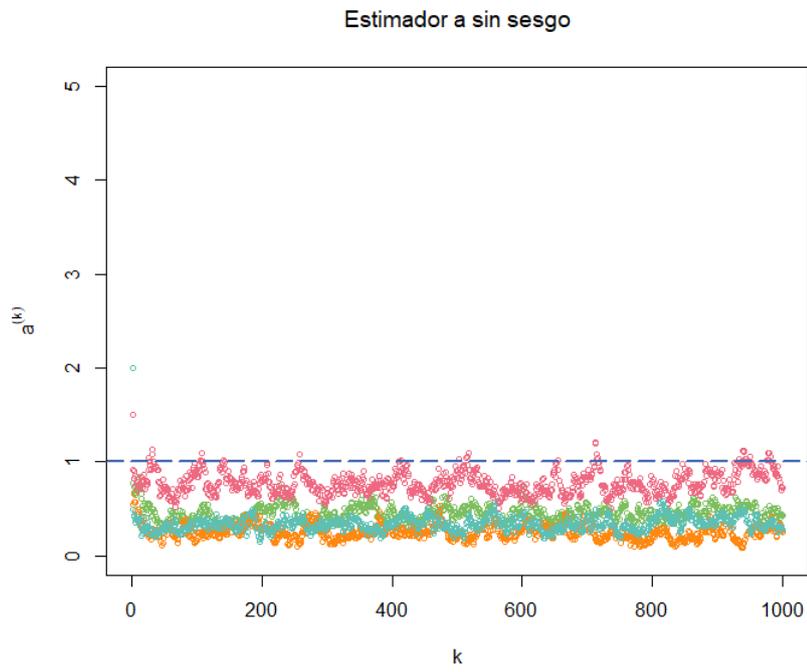


Figura 5.12: Estimador EM $a = 1$ sin sesgo

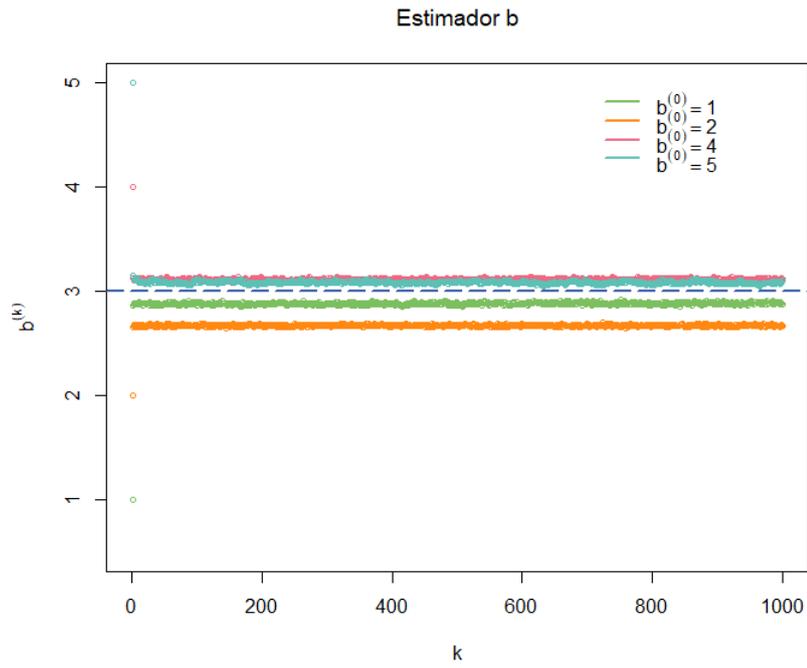


Figura 5.13: Estimador EM $b = 3$

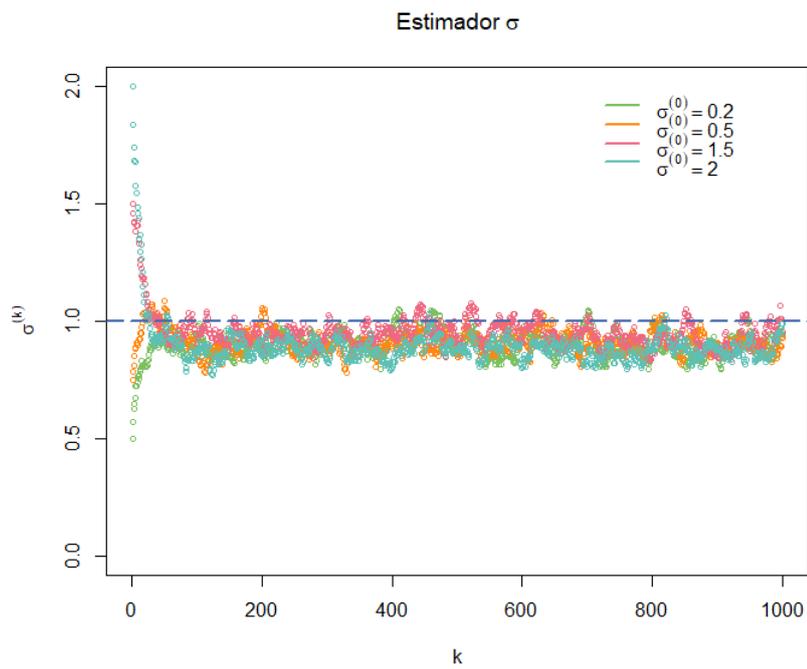


Figura 5.14: Estimador EM $\sigma = 1$

EM movimiento browniano geométrico

Algoritmo 8 : EME MBG

- 1: Condiciones iniciales: $\theta^{(0)}$, $k = 0$.
- 2: **Paso E:** Generar los puentes de difusión $\{Z_i^{(k)} = (X_{t_{i-1}}, t_{i-1}, X_{t_i}, t_i)\}_{i=1}^n$ con los parámetros $\theta^{(k)}$ y calcular $\hat{Q}(\theta; \theta^{(k)})$.
- 3: **Paso M:** Actualizar los estimadores:

$$\hat{\mu}^{(k+1)} = \frac{\overline{Z^{(k)}}}{\Delta t_i} + \frac{\hat{\sigma}^2}{2},$$

$$\hat{\sigma}^2^{(k+1)} = \frac{1}{N\Delta t_i} \sum_{i=1}^N (Z_{t_i}^{(k)} - \overline{Z})^2.$$

Donde $\overline{Z^{(k)}} = \frac{1}{N} \sum_{i=1}^N Z_{t_i}^{(k)}$.

- 4: Parar si hay $k=K$. En otro caso, $k = k + 1$ y volver al Paso MC E.
-

Pruebas y resultados

El algoritmo se prueba con trayectorias de movimiento browniano geométrico cuyos parámetros son $\theta = (0.5, 0.2)$ y valor inicial $X_0 = 1$. Se itera $K = 1000$ y el estimador se obtiene de promediar las últimas 100 iteraciones. Los parámetros iniciales son los obtenidos con el estimador máximo verosímil usando la trayectoria incompleta.

Se calculan los estimadores puntuales y por intervalos usando 1000 trayectorias. También, los errores relativos respecto al parámetro real y respecto al estimador MLE puntual (ver Tabla 5.3). Los estimadores y errores se muestran en la Tabla 5.7. Los estimadores obtenidos y el cálculo de los errores se encuentran en el archivo `Estimadores_MBG.xlsx` dentro del repositorio.

	Estimador puntual	Estimador por intervalos	Error relativo	Error relativo MLE
$\hat{\mu}$	0.503699	(0.49966, 0.50774)	0.007398	0.004810
$\hat{\sigma}$	0.211008125	(0.21014, 0.21188)	0.05504	0.05575

Tabla 5.7: Estimadores EM y errores

Los estimadores obtenidos son buenos, con errores relativos cercanos a cero. Para el movimiento browniano geométrico no hubo trayectorias donde el algoritmo fuera divergente.

En las Figuras 5.15 y 5.16 se grafican los estimadores obtenidos en cada iteración para 4 trayectorias. En cada prueba se empieza el algoritmo con una condición inicial diferente. A partir de las gráficas se observa que hay convergencia del estimador al parámetro y que esta es independiente de los valores iniciales $\theta^{(0)} = (\mu^{(0)}, \sigma^{(0)})$.

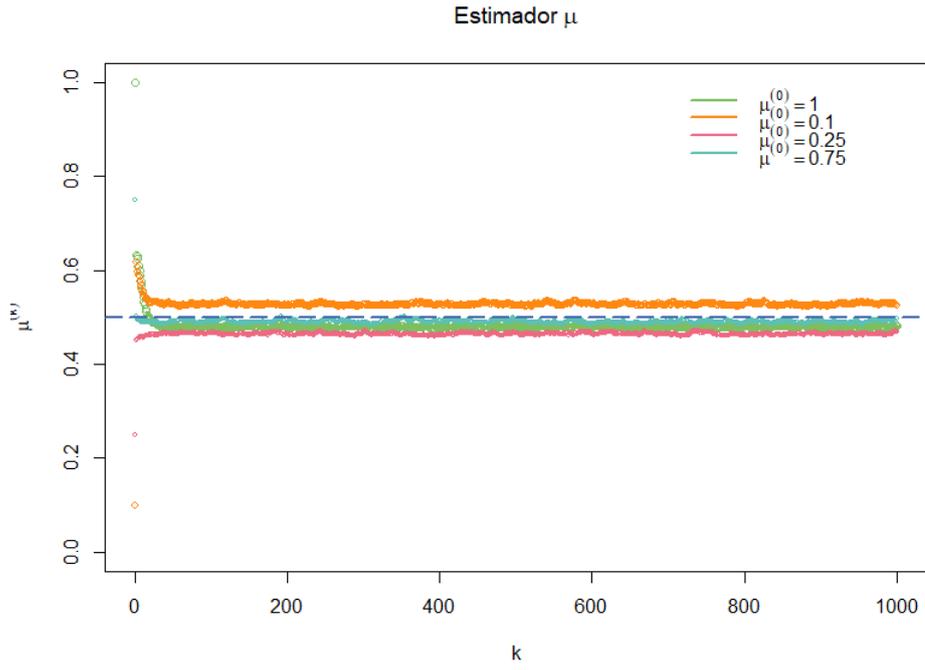


Figura 5.15: Estimador EM $\mu = 0.5$

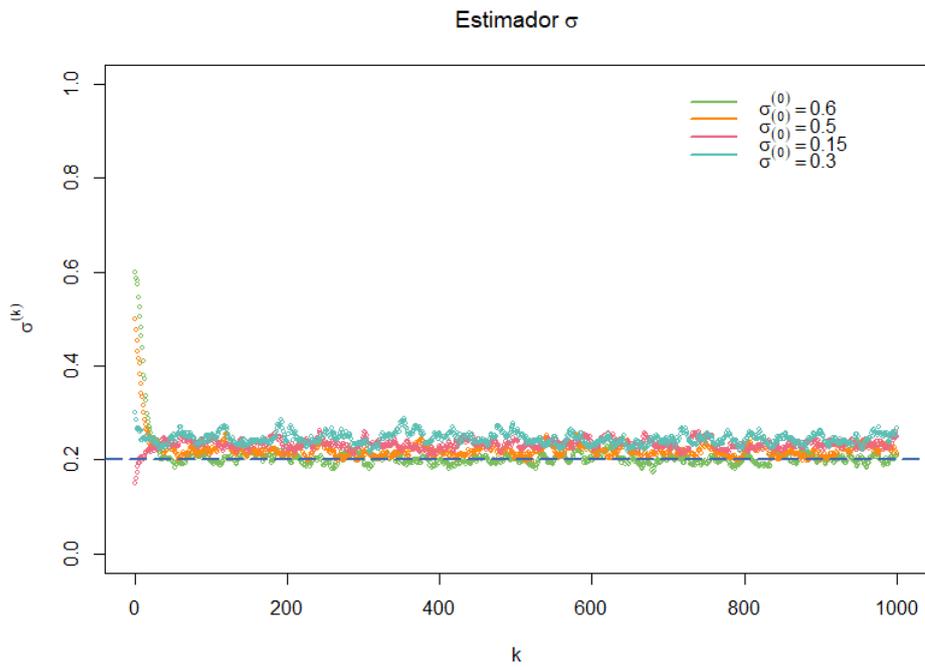


Figura 5.16: Estimador EM $\sigma = 0.2$

Capítulo 6

Aplicaciones de procesos de difusión en finanzas

En este último capítulo se utilizan las bases teóricas y computacionales descritas anteriormente, para trabajar con dos aplicaciones financieras. En la Sección 6.1 se aborda la aplicación correspondiente al proceso de Vasicek. Esta es en modelos de tasa corta y la valuación de bonos cupón cero. En la Sección 6.2 se trabaja con valuación de opciones, usando la fórmula de Black-Scholes y el movimiento browniano geométrico. Ver [25] para detalles sobre modelos de tasa corta y [22] y [18] para detalles sobre Black-Scholes y opciones.

Dentro de cada sección se da un contexto teórico de cómo se usan los procesos de difusión a las aplicaciones correspondientes. También se implementan computacionalmente los modelos. Finalmente, se hace estimación paramétrica a partir de una base de datos reales ajustada a cada modelo.

6.1. Proceso de Vasicek y tasas de interés

6.1.1. Modelos de tasa corta

Una *tasa corta* r_t es una tasa continuamente compuesta, anualizada para un periodo infinitesimalmente corto en un tiempo futuro t . Esta representa el interés por prestar o invertir dinero.

Las tasas cortas no son constantes, sino que dependen del tiempo. Tampoco son deterministas pues están sujetas a factores aleatorios que influyen en el mercado. Por lo tanto la dinámica de las tasas cortas es modelada por medio de un proceso de difusión,

$$dr_t = m(r_t, t)dt + \sigma(r_t, t)dB_t.$$

Las tasas de interés son una pieza fundamental a la hora de valorar un instrumento. En este caso se trabajará en la valuación de bonos cupón 0 con fecha de maduración T . Dada una tasa corta r_t y un valor de redención unitario, la valuación neutral al riesgo del bono al tiempo $t \leq T$ está dada por

$$P(t, T) = \mathbb{E} \left[\exp \left(- \int_t^T r_s ds \right) \middle| \mathcal{F}_t \right],$$

donde \mathcal{F}_t es la filtración natural del proceso de tasa corta al tiempo t . En otras palabras, la información disponible sobre la tasa hasta ese momento.

Los modelos de tasa corta pueden ser *modelos de equilibrio*, en los cuales el resultado del modelo es la tasa, o *modelos de arbitraje* donde la tasa es un parámetro del modelo. También se pueden clasificar como modelos de un factor o modelos multifactor. Los modelos de un factor tienen como supuesto que hay sola variable afectando a la tasa de interés, mientras que los multifactor toman en cuenta 2 o más variables ¹.

6.1.2. Modelo de Vasicek

En 1977 Oldrich Vasicek propuso modelar la dinámica de las tasas de interés de acuerdo a la ecuación (3.3.3). Este es un modelo de equilibrio y de un solo factor. Un supuesto importante es que r_t es un proceso de Markov continuo; es un proceso que no da saltos y donde el comportamiento para un tiempo futuro es independiente de la dinámica en tiempos pasados. También asume que el mercado es eficiente; no hay costos de transacción, los inversionistas actúan de forma racional y la información es completa y simultánea para todos los participantes.

Recordamos algunos resultados enunciados en la Sección 3.3.1. La ecuación diferencial estocástica que satisface r_t es

$$dr_t = a(b - r_t)dt + \sigma dB_t, \quad (6.1.1)$$

donde $a, b, \sigma \in \mathbb{R}$ y $a, \sigma > 0$. La solución a la ecuación dado un valor inicial r_s está dada por

$$r_t = e^{-a(t-s)}r_s - b(e^{-a(t-s)} - 1) + \sigma \int_s^t e^{-a(t-u)} dB_u. \quad (6.1.2)$$

La esperanza y varianza dada la información al tiempo s son

$$\mathbb{E}[r_t | \mathcal{F}_s] = r_s e^{-a(t-s)} - b(e^{-a(t-s)} - 1),$$

$$\text{Var}(r_t | \mathcal{F}_s) = \frac{\sigma^2}{2a}(1 - e^{-2a(t-s)}).$$

Además, sabemos que $r_t | r_s$ sigue una distribución normal

$$r_t | r_s \sim \mathcal{N}(\mathbb{E}[r_t | \mathcal{F}_s], \text{Var}(r_t | \mathcal{F}_s)).$$

Se dice que r_t es un proceso con reversión a la media. Recordamos que el parámetro b representa la media a largo plazo y a la velocidad de reversión a la media. Si las tasas están por encima de la media entonces $r_t > b$ y el factor de deriva en dr_t es negativo.

$$r_t > b \Rightarrow a(b - r_t) < 0.$$

Por lo tanto la tasa decrece, acercándose a la media. Análogamente si $r_t < b$ la tasa crecerá hacia b . También notamos que

$$\lim_{t \rightarrow \infty} \mathbb{E}[r_t | r_s] = b,$$

¹Ver [25] para más detalles sobre modelos de estructura temporal de tasas.

es decir, el valor esperado de r_t al largo plazo es precisamente la media a largo plazo. Una ventaja del modelo es la reversión a la media. Este es un comportamiento que se ha visto empíricamente e históricamente en el mercado. Al seguir una distribución normal las tasas pueden llegar a tomar valores negativos. Aunque este no es un escenario deseable, se ha llegado a observar en momentos de crisis. Existen alternativas al modelo donde no se admite el supuesto de posibles tasas negativas, como Cox-Ingersoll-Ross y Vasicek exponencial.

6.1.3. Valuación bonos cupón cero

Como se mencionó anteriormente, el precio justo al tiempo t de un bono con periodo de maduración T , valor de redención unitaria y con una tasa corta r_t está dado por

$$P(t, T) = \mathbb{E} \left[\exp \left(- \int_t^T r_u du \right) \middle| \mathcal{F}_t \right].$$

Buscamos expresar este precio usando la dinámica de tasas propuesta por Vasicek. Recordamos que la función generadora de momentos de una variable aleatoria X con distribución normal está dada por $\mathbb{E}[e^X] = \exp \left(\mathbb{E}[X] + \frac{1}{2} \text{Var}(X) \right)$. Por lo tanto

$$P(t, T) = \exp \left(\mathbb{E} \left[- \int_t^T r_u du \middle| \mathcal{F}_t \right] + \frac{1}{2} \text{Var} \left(- \int_t^T r_u du \middle| \mathcal{F}_t \right) \right).$$

Se calcula $\int_t^T r_u du$ usando la solución (6.1.2).

$$\begin{aligned} \int_t^T r_u du &= \int_t^T r_t e^{-a(u-t)} du - b \int_t^T (e^{-a(u-t)} - 1) du + \sigma \int_t^T \int_t^u e^{-a(u-s)} dB_s du \\ &= -\frac{r_t}{a} [e^{-a(u-t)}]_t^T + b \left[\frac{e^{-a(u-t)}}{a} - u \right]_t^T + \sigma \int_t^T \int_t^u e^{-a(u-s)} dB_s du \\ &= \frac{r_t}{a} (1 - e^{-a(T-t)}) + b \left((T-t) - \frac{1 - e^{-a(T-t)}}{a} \right) + \sigma \int_t^T \int_t^u e^{-a(u-s)} dB_s du. \end{aligned}$$

A partir del resultado anterior se calcula $\mathbb{E} \left[- \int_t^T r_u du \middle| \mathcal{F}_t \right] = -\mathbb{E} \left[\int_t^T r_u du \middle| \mathcal{F}_t \right]$.

$$\begin{aligned} \mathbb{E} \left[\int_t^T r_u du \middle| \mathcal{F}_t \right] &= \mathbb{E} \left[\frac{r_t}{a} (1 - e^{-a(T-t)}) \right] + \mathbb{E} \left[b \left((T-t) - \frac{1 - e^{-a(T-t)}}{a} \right) \right] \\ &\quad + \mathbb{E} \left[\sigma \int_t^T \int_t^u e^{-a(u-s)} dB_s du \right] \\ &= \frac{r_t}{a} (1 - e^{-a(T-t)}) + b \left((T-t) - \frac{1 - e^{-a(T-t)}}{a} \right) + \sigma \int_t^T \mathbb{E} \left[\int_t^u e^{-a(u-s)} dB_s \right] du. \end{aligned}$$

Los primeros dos sumandos son deterministas, por lo que salen como constantes de la esperanza. La tercera esperanza corresponde a la integral estocástica de una función determinista. Por lo tanto su esperanza es cero. Entonces

$$\mathbb{E} \left[- \int_t^T r_u du \middle| \mathcal{F}_t \right] = -\frac{r_t}{a} (1 - e^{-a(T-t)}) - b \left((T-t) - \frac{1 - e^{-a(T-t)}}{a} \right).$$

Se calcula la varianza. Como los dos primeros sumandos de $\int_t^T r_u du$ son deterministas, por lo que su varianza es cero. Entonces

$$\text{Var} \left(- \int_t^T r_u du | \mathcal{F}_t \right) = \text{Var} \left(-\sigma \int_t^T \int_t^u e^{-a(u-s)} dB_s du \right).$$

Por la isometria de Itô la varianza se reduce a

$$\text{Var} \left(-\sigma \int_t^T \int_t^u e^{-a(u-s)} dB_s du \right) = \mathbb{E} \left[\left(-\sigma \int_t^T \int_t^u e^{-a(u-s)} dB_s du \right)^2 \right].$$

Utilizando el Teorema de Fubini y la isometria de Itô se obtiene

$$\begin{aligned} \text{Var} \left(-\sigma \int_t^T \int_t^u e^{-a(u-s)} dB_s du \right) &= \mathbb{E} \left[\left(-\sigma \int_t^T \int_s^T e^{-a(u-s)} du dB_s \right)^2 \right] \\ &= \mathbb{E} \left[\sigma^2 \int_t^T \left(\int_s^T e^{-a(u-s)} du \right)^2 ds \right]. \end{aligned}$$

Se llega a la esperanza de una función determinista, por lo que sale como constante y se integra el resultado.

$$\begin{aligned} \text{Var} \left(-\sigma \int_t^T \int_t^u e^{-a(u-s)} dB_s du \right) &= \sigma^2 \int_t^T \left(\int_s^T e^{-a(u-s)} du \right)^2 ds \\ &= \sigma^2 \int_t^T \left(-\frac{e^{-a(u-s)}}{a} \Big|_s^T \right)^2 ds \\ &= \sigma^2 \int_t^T \left(\frac{1 - e^{-a(T-s)}}{a} \right)^2 ds \\ &= \frac{\sigma^2}{a^2} \int_t^T (1 - 2e^{-a(T-s)} + e^{-2a(T-s)}) ds \\ &= \frac{\sigma^2}{a^2} \left(s - \frac{2e^{-a(T-s)}}{a} + \frac{e^{-2a(T-s)}}{2a} \right) \Big|_t^T \\ &= \frac{\sigma^2}{a^2} \left((T-t) + 2 \left(\frac{e^{-a(T-t)} - 1}{a} \right) + \frac{1 - e^{-2a(T-t)}}{2a} \right). \end{aligned}$$

Por lo tanto el precio $P(t, T)$ para un bono que rinde de acuerdo a la tasa modelada por (6.1.2) es

$$\begin{aligned} P(t, T) &= \exp \left\{ \left(-\frac{r_t}{a} (1 - e^{-a(T-t)}) - b \left((T-t) - \frac{1 - e^{-a(T-t)}}{a} \right) \right) \right. \\ &\quad \left. + \frac{\sigma^2}{2a^2} \left((T-t) + 2 \left(\frac{e^{-a(T-t)} - 1}{a} \right) + \frac{1 - e^{-2a(T-t)}}{2a} \right) \right\}. \end{aligned}$$

Se define $B(t, T) = \frac{1 - e^{-a(T-t)}}{a}$ y se sustituye en $P(t, T)$.

$$P(t, T) = \exp \left\{ (-r_t B(t, T) - b((T-t) - B(t, T))) + \frac{\sigma^2}{2a^2} \left((T-t) - 2B(t, T) + \frac{1 - e^{-2a(T-t)}}{2a} \right) \right\}.$$

El sumando $\frac{1-e^{-2a(T-t)}}{2a}$ se puede expresar en términos de $B(t, T)$.

$$\begin{aligned} B(t, T)^2 &= \frac{1 - 2e^{-a(T-t)} + e^{-2a(T-t)}}{a^2} \\ a^2 B(t, T)^2 &= 1 - 2e^{-a(T-t)} + e^{-2a(T-t)} \\ 2 - a^2 B(t, T)^2 - 2e^{-a(T-t)} &= 1 - e^{-2a(T-t)} \\ \frac{2 - a^2 B(t, T)^2 - 2e^{-a(T-t)}}{2a} &= \frac{1 - e^{-2a(T-t)}}{2a} \\ \frac{1 - e^{-a(T-t)}}{a} - \frac{aB(t, T)^2}{2} &= \frac{1 - e^{-2a(T-t)}}{2a} \\ B(t, T) - \frac{aB(t, T)^2}{2} &= \frac{1 - e^{-2a(T-t)}}{2a}. \end{aligned}$$

Sustituyendo en $P(t, T)$.

$$\begin{aligned} P(t, T) &= \exp \left\{ (-r_t B(t, T) - b((T-t) - B(t, T))) + \frac{\sigma^2}{2a^2} \left((T-t) - B(t, T) - \frac{aB(t, T)^2}{2} \right) \right\} \\ &= \exp \left\{ -r_t B(t, T) + \left(\frac{\sigma^2}{2a} - b \right) [(T-t) - B(t, T)] - \frac{\sigma^2 B(t, T)^2}{4a} \right\}. \end{aligned}$$

Por lo tanto el precio $P(t, T)$ para la tasa r_t está dado por

$$P(t, T) = A(t, T)e^{-r_t B(t, T)},$$

donde

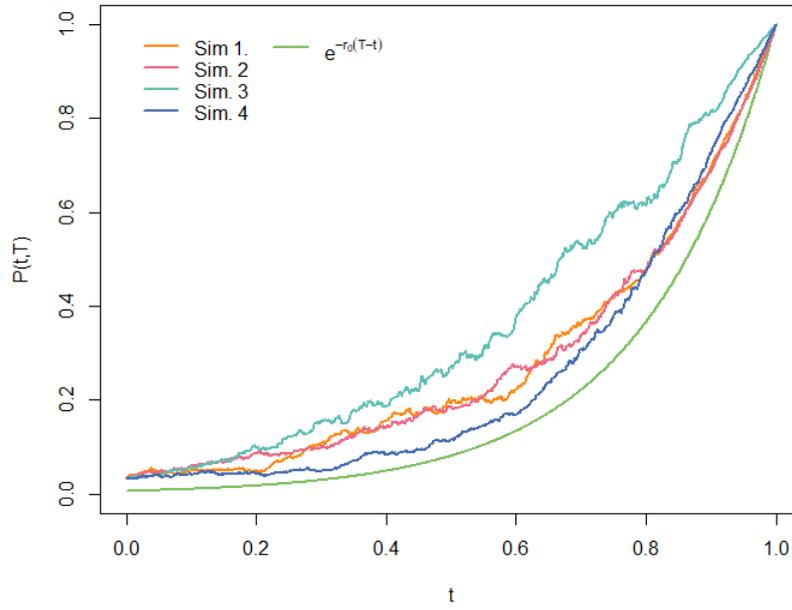
$$A(t, T) = \exp \left\{ \left(\frac{\sigma^2}{2a} - b \right) [(T-t) - B(t, T)] - \frac{\sigma^2 B(t, T)^2}{4a} \right\}.$$

6.1.4. Implementación computacional

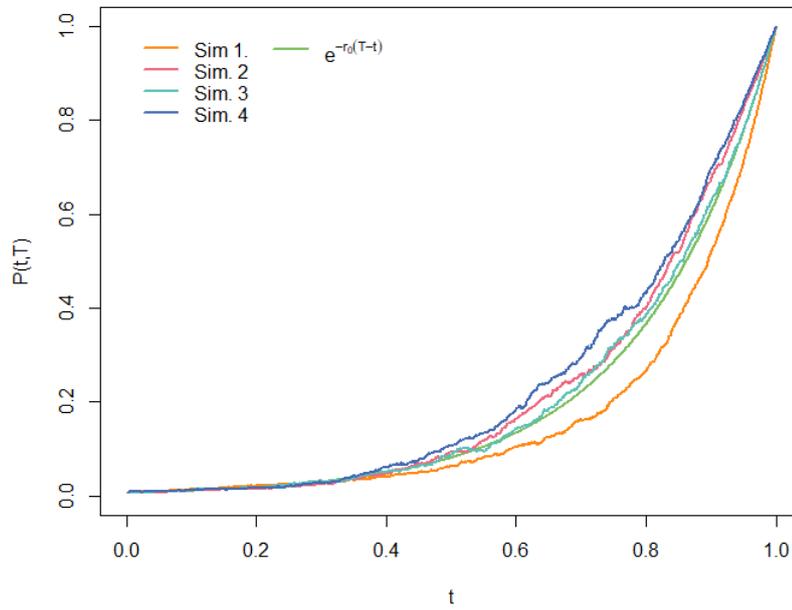
A continuación se muestran diferentes gráficas comparando el precio determinista del cupón contra el precio $P(t, T)$. El precio determinista está dado por $t \mapsto e^{-r_0(T-t)}$. En la Figura 6.1 se muestran trayectorias variando b , la media a largo plazo. Los parámetros usados son $r_0 = 5\%$, $T = 1$, $a = 0.5$ y $\sigma = 1$.

Notamos que si $b < r_0$ los precios estocásticos toman valores mayores a los del precio determinista, y al contrario cuando $b > r_0$. Si $b = r_0$ entonces hay un mayor parecido entre los precios deterministas y aleatorios. En el caso donde $b = 1\%$ y $r_0 = 5\%$ el proceso r_t decrecerá para tender a la media b . Por ser un instrumento valuado a descuento al bajar las tasas el precio incrementa. Es por ello que las trayectorias de precios simulados se encuentran por encima del precio determinista.

Trayectoria Precios: b = 1%



Trayectoria Precios: b = 5%



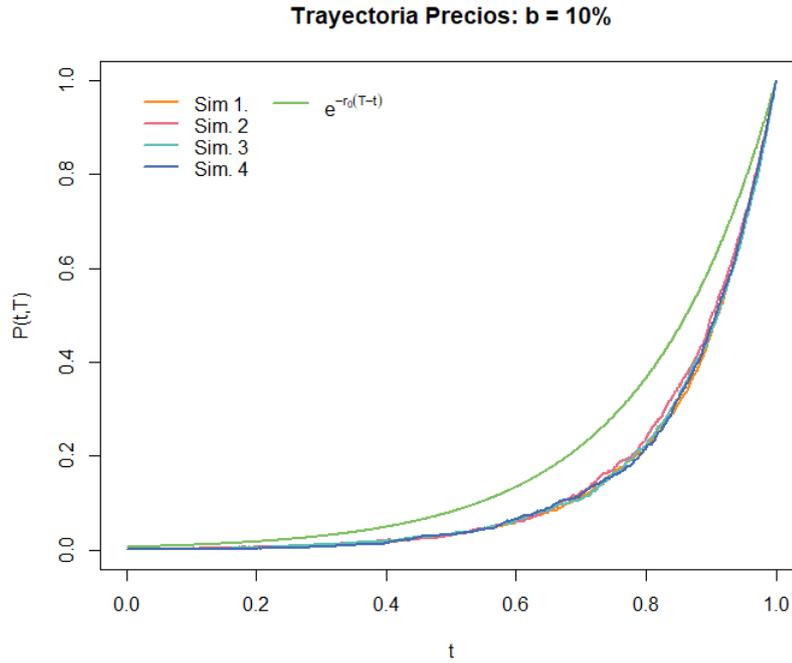


Figura 6.1: precios teóricos vs simulados

6.1.5. Estimación paramétrica

A partir de las herramientas descritas en el Capítulo 5, se hará estimación paramétrica sobre una serie de tasas correspondientes a bonos cupón-zero del gobierno de Canadá a 1 año.² Se utilizará una muestra de 251 datos diarios del año 2019, considerando únicamente días hábiles. En la Tabla 6.1 se presenta estadística descriptiva sobre la base de datos

Min.	Q1	Mediana	Q3	Max.	Media	Var.	SD
1.51	1.68	1.72	1.76	1.93	1.73	0.007	0.086

Tabla 6.1: Estadística descriptiva tasas cupón 0.

En las Tablas 6.2 y 6.3 se presentan los estimadores obtenidos usando tanto estimación MLE como EM.

	a	b	σ
MLE	10.6416	1.7108	0.3538
EM	8.7281	1.7054	0.3066

Tabla 6.2: Estimadores para proceso r_t

²La información fue obtenida en <https://www.bankofcanada.ca/rates/interest-rates/bond-yield-curves/>. La base de datos puede ser consultada en el archivo `Tasas_Canada.csv` dentro del repositorio de Github en la carpeta Bases de datos

	a	b	σ
MLE	6.5889	1.7014	0.3516
EM	4.3618	1.6855	0.3013

Tabla 6.3: Estimadores para proceso r_t con corrección de sesgo

Para obtener los estimadores se hace una partición uniforme del intervalo de tiempo $[0, 1]$. Las particiones son de tamaño 251 y 51 para los estimadores MLE y EM, respectivamente. Es decir, se divide 1 año entre el número de datos observados. Por lo tanto, los tamaños de paso constantes son $\Delta t_{MLE} = \frac{1}{250}$ y $\Delta t_{EM} = \frac{1}{50}$.

Para la estimación EM se utilizó un dato semanal, o 1 de cada 5, pensando en semanas de días hábiles. Por lo tanto se tienen 50 datos, equivalente a 80 % de información faltante. Los parámetros iniciales son los estimadores máximo verosímiles para la información incompleta. El algoritmo EM ejecutó con 1000 iteraciones, y los estimadores obtenidos son un promedio de las últimas 100 iteraciones.

En las Figuras 6.2 y 6.3 se grafican las tasas de la muestra de datos, así como trayectorias simuladas usando los parámetros máximo verosímiles. Similarmente, en las Figuras 6.4 y 6.5 se muestra la información incompleta de las tasas y trayectorias simuladas a partir de los estimadores EM. En ambos casos se ve que las trayectorias simuladas tienen un comportamiento similar a la muestra de datos.

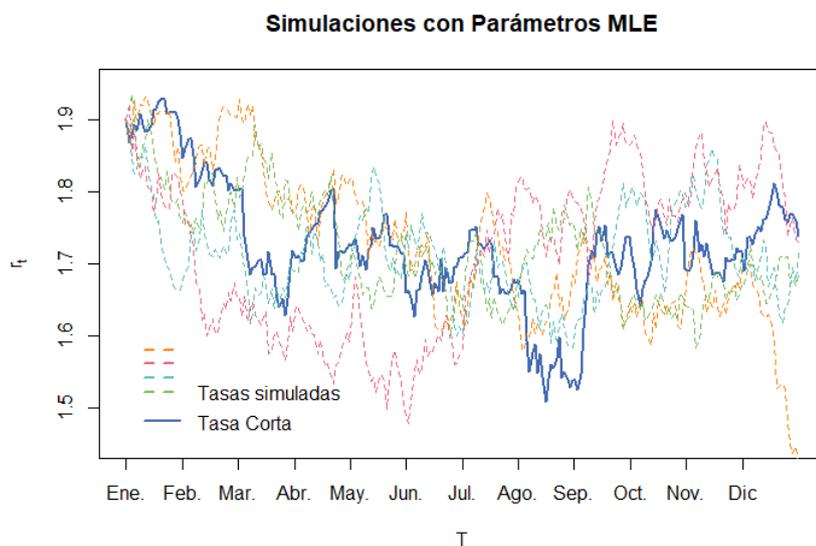


Figura 6.2: Tasa corta: real y simuladas

Simulaciones con Parámetros MLE sin sesgo

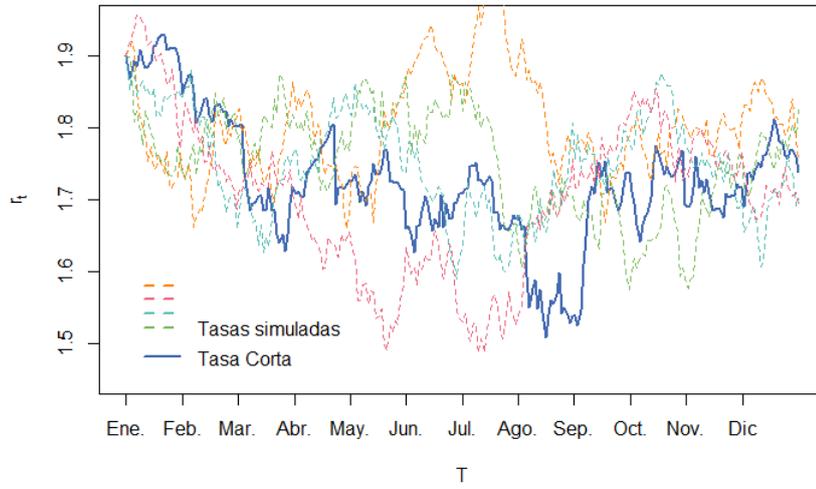


Figura 6.3: Tasa corta: real y simuladas

Simulaciones con Parámetros EM

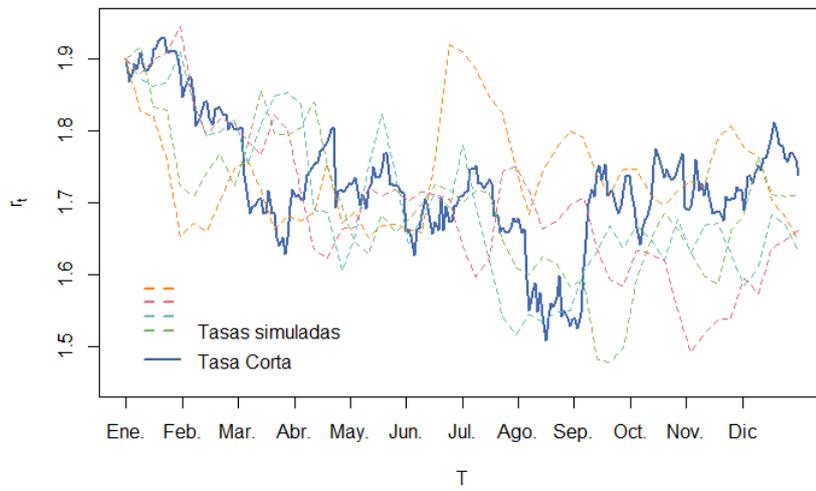


Figura 6.4: Tasa corta: real (incompleta) y simuladas

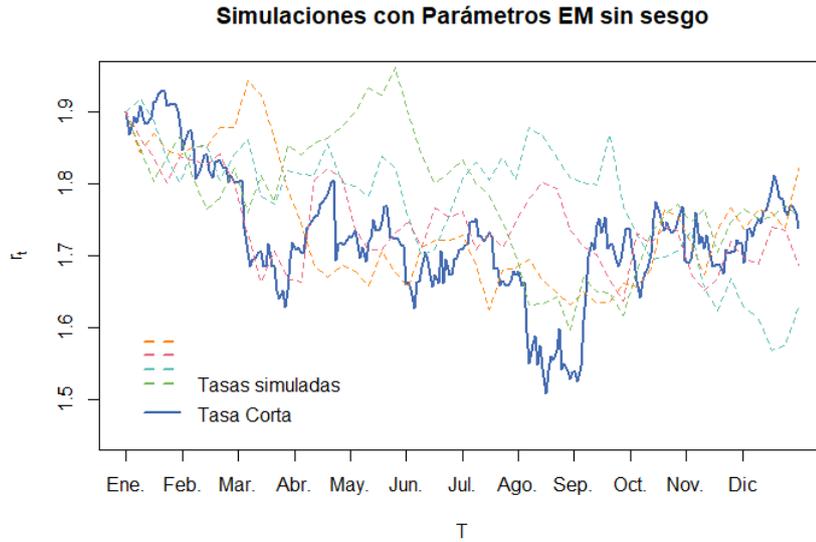


Figura 6.5: Tasa corta: real (incompleta) y simuladas

Por último, se grafican los precios del bono utilizando la base de datos de tasas. Este precio está dado por $P(t) = e^{-\int_t^T r(s)ds}$. También se grafican dos trayectorias de precios utilizando trayectorias de tasas simuladas, utilizando los parámetros estimados. Como se puede ver en las Figuras 6.6 y 6.7 los precios simulados se ajustan a la trayectoria real de precios.

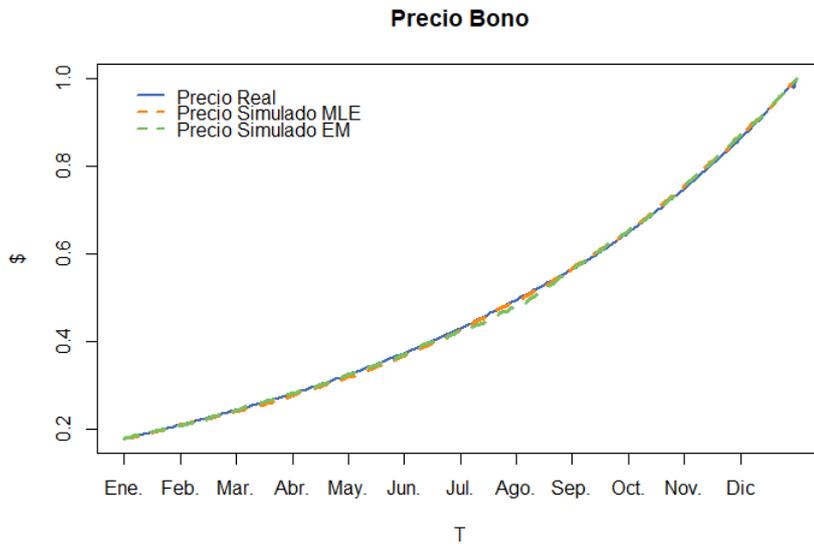


Figura 6.6: Precios: real y simulados

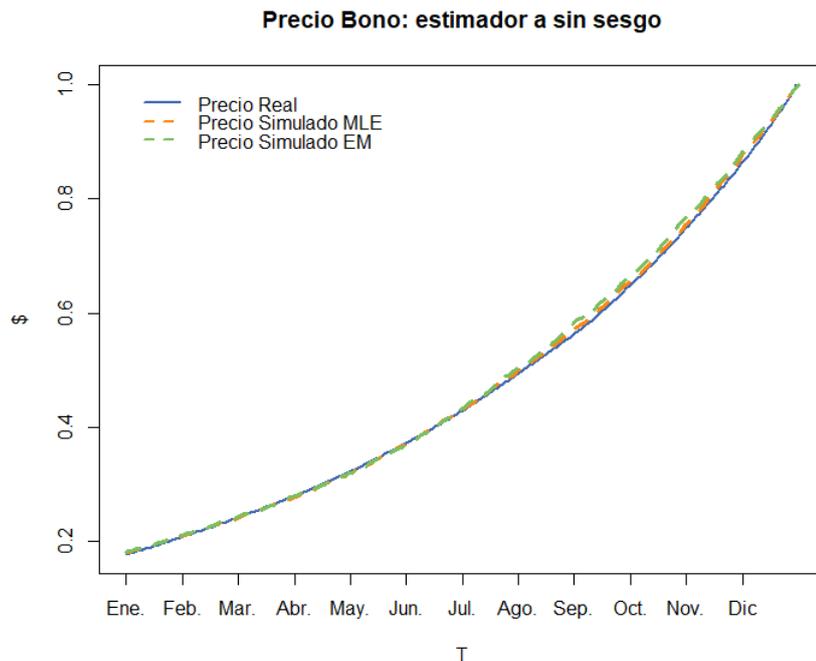


Figura 6.7: Precios: real y simulados

Validación del modelo

Se valida que el proceso de Vasicek con los parámetros estimados realmente modelen los datos reales. Se aplica la transformación de Lamperti a la ecuación 6.1.1.

$$dY_t = \frac{a(b - r_t)}{\sigma} dt + dB_t,$$

donde $Y_t = \frac{r_t}{\sigma}$. Se define a los residuales $\widehat{\Delta B}_{t_i}$ como

$$\widehat{\Delta B}_{t_i} = \frac{1}{\widehat{\sigma}} \left[r_{t_i} - r_{t_{i-1}} - \widehat{a}(\widehat{b} - r_{t_i})\Delta t \right], \quad i = 2, \dots, 251.$$

Donde $\widehat{a}, \widehat{b}, \widehat{\sigma}$ son los parámetros estimados.

Para verificar la bondad de ajuste de los residuales a la distribución normal, se generaron pruebas de Kolmogorov-Smirnov. En el caso de los residuales con estimadores máximo verosímiles, se generaron 1000 muestras de vectores normales de tamaño 250 (número de residuales), con media 0 y varianza $\Delta t_{MLE} = \frac{1}{250}$. Las 1000 muestras se compararon con los residuales y se promediaron los *p-values* obtenidos con cada muestra.

El mismo proceso se hizo para comparar los residuales de los estimadores EM y la distribución normal. En este caso las 1000 muestras de vectores normales eran de tamaño 50 con una varianza de $\Delta t_{EM} = \frac{1}{50}$. En la Tabla 6.4 se muestran los *p-values* obtenidos en cada caso.

	MLE	EM
a con sesgo	0.1992	0.7282
a sin sesgo	0.1832	0.6377

Tabla 6.4: p -values pruebas K-S

En cada caso, los 1000 vectores normales fueron sumados entrada a entrada y el vector resultante multiplicado por $\frac{1}{\sqrt{1000}}$ (para recuperar la varianza). Con ese vector se generaron gráficos QQ que comparan la muestra normal con los residuales.

Las gráficas (a) y (b) de la Figura 6.8 muestran los gráficos QQ para los estimadores MLE y EM, respectivamente, haciendo el cálculo de los estimadores con a sesgado.

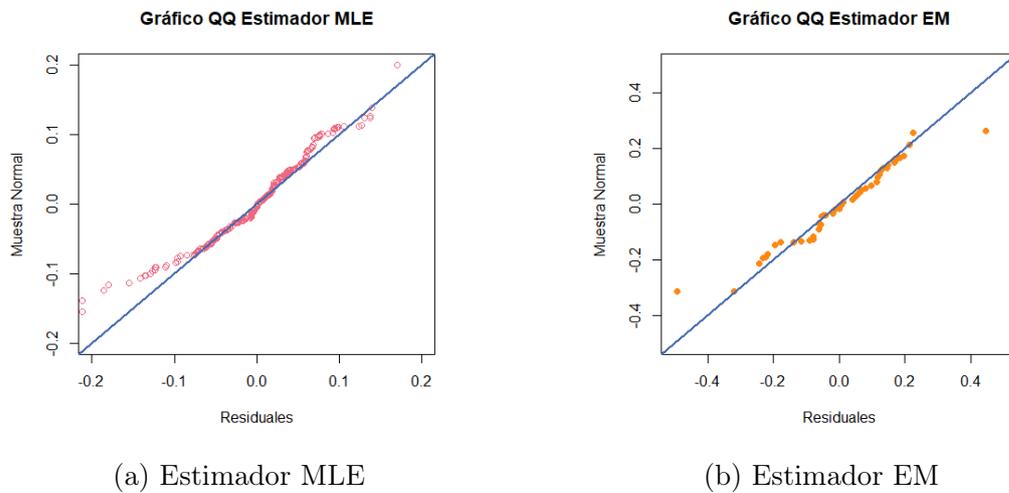
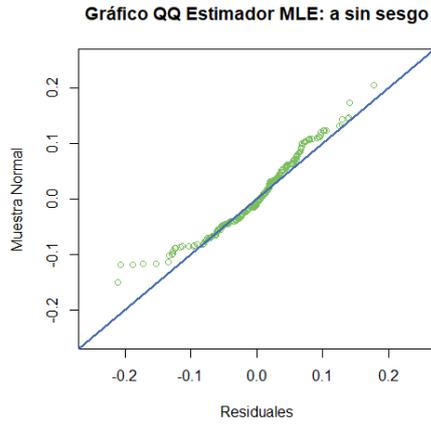


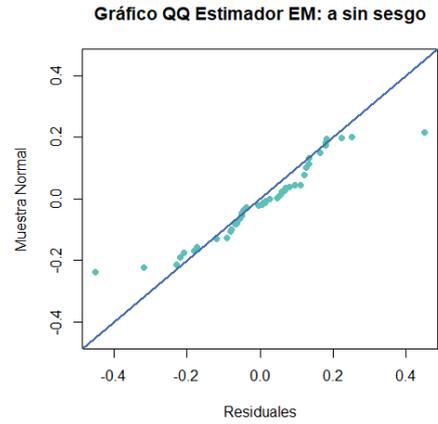
Figura 6.8: gráficos QQ: tasa corta

Las gráficas de la Figura 6.9 muestran la misma información que la Figura 6.8 pero usando los estimadores de la Tabla 6.3. Es decir, los estimadores obtenidos haciendo la corrección del sesgo para el parámetro a .³

³Las muestras utilizadas para la validación, así como los p -values obtenidos para cada muestra se encuentran en el archivo `Muestras_Validación_Tasa_Corta.xlsx` en el repositorio de Github en la carpeta Bases de datos



(a) Estimador MLE



(b) Estimador EM

Figura 6.9: gráficos QQ: tasa corta, a sin sesgo

A partir de los p -values y los gráficos QQ se tiene evidencia de que los datos reales se ajustan al modelo de Vasicek con los parámetros estimados. También se concluye que hay buen ajuste haya sesgo o no, pero que es mejor en el primer caso.

6.2. Movimiento browniano geométrico y valuación de opciones

6.2.1. Opciones financieras

Las opciones financieras son instrumentos derivados que otorgan el derecho, pero no la obligación, de comprar o vender un activo subyacente a un precio establecido y en una fecha determinada. Al precio establecido se le llama *precio de ejercicio* y se denota con la letra K . A la fecha final se le conoce como *fecha de vencimiento* y se denota con la letra T . Las opciones se pueden clasificar de acuerdo a cuándo se pueden ejercer:

- **Europeas:** Sólo pueden ser ejercidas en la fecha de vencimiento.
- **Americanas:** Pueden ser ejercidas en cualquier momento entre la compra y la fecha de vencimiento.

También se pueden clasificar como *call* y *put*:

- **Call:** Quien toma la posición larga de la opción tiene derecho a comprar el activo subyacente. En ese caso, la posición corta tiene la obligación de venderlo.
- **Put:** El comprador del contrato tiene derecho a vender el activo subyacente. El emisor tendría la obligación de comprarlo.

Se define S_t como el precio del activo subyacente para $t \in [0, T]$, S_0 y S_T representan el precio inicial y final, respectivamente.

Consideremos una opción *call* Europea ⁴. Si al tiempo T se tiene $S_T > K$ el comprador decidirá ejercer la opción de compra, pues supondría una ganancia de $S_T - K$. En caso contrario no ejerce una opción, por lo que no tiene ganancia pero tampoco pérdida. Se define como *valor intrínseco* a la diferencia positiva entre el precio de mercado del activo subyacente y el precio de ejercicio de la opción. Es decir

$$(S_T - K)^+ = \text{máx}\{S_T - K, 0\}. \quad (6.2.1)$$

6.2.2. Valuación libre de arbitraje y Black-Scholes

La teoría y pruebas correspondientes a esta sección pueden ser consultadas en [18].

Sea $\{A_t\}$ el precio de un activo libre de riesgo con tasa de rendimiento r .

$$dA_t = rA_t dt, \quad A_t = A_0 e^{rt}, \quad t \geq 0.$$

Sea $\{S_t\}$ el precio de un activo riesgoso, por ejemplo una acción. Este representará el activo subyacente de la opción. Suponemos que el precio sigue un movimiento browniano geométrico con media μ y varianza σ^2 .

$$dS_t = \mu S_t dt + \sigma S_t dB_t. \quad (6.2.2)$$

⁴De ahora en adelante al hacer referencia a una opción será de tipo *call* y Europea.

Un portafolio tiene invertido al tiempo t las cantidades η_t y ξ_t en los activos A_t y S_t , respectivamente. Por lo tanto, el valor del portafolio al tiempo t está dado por

$$V_t = \eta_t A_t + \xi_t S_t.$$

A continuación se enuncian algunas definiciones y proposiciones necesarias para definir el precio libre de arbitraje de una opción.

Definición 6.2.1. Un *reclamo contingente* es una variable aleatoria no negativa denotada $C \geq 0$. Este caso C representa el precio de la opción al tiempo T .

Definición 6.2.2. El reclamo contingente C es *alcanzable* si existe un portafolio con estrategia $\{\eta_t, \xi_t\}_{t \in [0, T]}$ tal que $C = V_T$. Se dice entonces que el portafolio *cubre* al reclamo contingente.

Definición 6.2.3. El portafolio con la estrategia $\{\eta_t, \xi_t\}$ es *autofinanciado* si el valor del portafolio es constante al pasar de $\{\eta_t, \xi_t\}$ a $\{\eta_{t+dt}, \xi_{t+dt}\}$

$$A_{t+dt}\eta_t + S_{t+dt}\xi_t = A_{t+dt}\eta_{t+dt} + S_{t+dt}\xi_{t+dt}.$$

En ese caso

$$A_{t+dt}d\eta_t = -S_{t+dt}d\xi_t.$$

Es decir, la compra (venta) del activo libre de riesgo entre t y $t + dt$ se financia con la venta (compra) del activo riesgoso.

Definición 6.2.4. El portafolio con la estrategia $\{\eta_t, \xi_t\}$ es una oportunidad de arbitraje si:

1. $V_0 \leq 0$,
2. $V_T \geq 0$,
3. $\mathbb{P}(V_T > 0) > 0$.

Definición 6.2.5. Una medida de probabilidad \mathbb{P}^* en Ω es una *medida neutral al riesgo* si

$$\mathbb{E}^*[S_t | \mathcal{F}_u] = e^{r(t-u)} S_u, \quad 0 \leq u \leq t.$$

Proposición 6.2.1. Sea $\{X_t\}$ el proceso de precios a descuento del activo riesgoso, es decir $X_t = e^{-rt} S_t$. La medida de probabilidad \mathbb{P}^* es neutral al riesgo si y sólo si $\{X_t\}$ es martingala bajo \mathbb{P}^* .

Teorema 6.2.1. (1er teorema fundamental de valuación de activos). Un mercado no tiene oportunidades de arbitraje si y sólo si admite al menos una medida de probabilidad neutral al riesgo.

En los modelos matemáticos para mercados financieros se utiliza el principio de no arbitraje, al cual se hace referencia en el Teorema 6.2.1. Matemáticamente se tiene que la ganancia esperada es cero, y que el rendimiento a valor presente del activo riesgoso es una martingala.

Intuitivamente en este mercado se tiene un “juego justo” para los diferentes competidores. Black, Scholes y Merton propusieron el precio $C(t, S_t)$ que debe valer una opción usando este principio.

Dado que el precio del activo riesgoso es un movimiento browniano geométrico, el proceso $\{X_t\}$ satisface

$$dX_t = (\mu - r)X_t dt + \sigma X_t dB_t.$$

Recordamos que se ha trabajado en el espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$. Si $\mu \neq r$ entonces $\{X_t\}$ no es martingala bajo \mathbb{P} . Se define \tilde{B}_t como

$$\tilde{B}_t := \frac{\mu - r}{\sigma}t + B_t.$$

Entonces $dX_t = \sigma X_t d\tilde{B}_t$. Se define \mathbb{P}^* tal que \tilde{B}_t sea un movimiento browniano. Así, por la propiedad de martingala de la integral estocástica, $\{X_t\}$ sería martingala. La medida \mathbb{P}^* es neutral al riesgo si se define por

$$\frac{d\mathbb{P}^*}{d\mathbb{P}} = \exp\left(-\frac{\mu - r}{\sigma}B_T - \frac{(\mu - r)^2}{2\sigma^2}T\right). \quad (6.2.3)$$

Entonces, por el Teorema 6.2.1, se cumple el principio de no arbitraje. Para hacer el cambio de medida se utiliza el Teorema de Girsanov. Los detalles de cómo se define \mathbb{P}^* pueden ser consultados en el Capítulo 6 de [18]. En adelante al usar \mathbb{P}^* o \mathbb{E}^* será la medida definida en (6.2.3).

Proposición 6.2.2. Sea $\{\eta_t, \xi_t\}$ la estrategia de un portafolio autofinanciado que cubre al reclamo contingente C . Sean $C(t, x)$ y $\phi(x)$ tales que

$$V_t = C(t, S_t), \quad C = V_T = \phi(S_T).$$

Entonces el precio bajo el principio de no arbitraje del reclamo está dado por

$$C(t, S_t) = e^{-r(T-t)}\mathbb{E}^*[\phi(S_T)|\mathcal{F}_t]$$

Además, la función $C(t, x)$ resuelve la ecuación diferencial parcial de Black-Scholes.

$$\begin{cases} rC(t, x) &= \frac{\partial C}{\partial t}(t, x) + \frac{1}{2}x^2\sigma^2\frac{\partial^2 C}{\partial x^2}(t, x) + rx\frac{\partial C}{\partial x}(t, x) \\ C(T, x) &= \phi(x), \end{cases} \quad (6.2.4)$$

y $\xi_t = \frac{\partial C}{\partial x}(t, S_t)$. A esta derivada se le denomina la Delta del precio de la opción.

En el siguiente teorema se enuncia la solución a la ecuación (6.2.4), dada por la fórmula de Black-Scholes. Esta representa el precio justo de la opción.

Teorema 6.2.2. (Fórmula de valuación de Black-Scholes) Consideramos una opción *call* europea con función de paga (valor intrínseco) $\phi(x) = (x - K)^+$, precio de ejercicio K y vencimiento T . El precio de la opción al tiempo t está dado por

$$C(t, S_t) = S_t\Phi(\omega) - Ke^{-r(T-t)}\Phi(\omega - \sigma\sqrt{T-t}).$$

$$\omega = \frac{(r + \sigma^2/2)(T-t) + \log(S_t/K)}{\sigma\sqrt{T-t}}.$$

donde Φ es la función de distribución normal estándar bajo la medida de probabilidad \mathbb{P}^* .

Proposición 6.2.3. La Delta de la opción descrita en el Teorema 6.2.2 está dada por

$$\xi_t = \Phi(\omega) \in [0, 1].$$

Al tiempo t el portafolio tiene invertido $S_t \xi_t = S_t \Phi(\omega)$ en el activo riesgoso. Simultáneamente la cantidad invertida en el activo libre de riesgo es

$$\eta_t A_t = -K A_0 e^{-r(T-t)} \Phi(\omega - \sigma \sqrt{T-t}).$$

Como $\Phi(\omega) \geq 0, S_t \geq 0$ siempre se toma una posición larga en el activo subyacente. En cambio la posición en el activo libre de riesgo siempre es corta.

El tener una fórmula cerrada hace que el cálculo sea mucho más rápido, ya que no hay necesidad de simular. El modelo puede ser usado también para la valuación de opciones con otros subyacentes como opciones o *commodities*. Una desventaja del modelo es que asume volatilidad y tasas de interés constantes, o que no considera el pago de dividendos de la acción subyacente.

6.2.3. Implementación computacional

Se simula el precio de una opción usando los siguientes parámetros

$$A_0 = 1, r = 15\%, S_0 = 1, \mu = 0.15, \sigma = 0.05, K = 4, T = 10.$$

A partir de 20 diferentes trayectorias de S_t , en la Figura 6.10 se simulan diferentes trayectorias de $C(t, S_t)$. Notemos que

$$\mathbb{E}[S_T] = e^{1.5} \approx 4.48.$$

El valor final esperado del activo subyacente es mayor al precio de ejercicio. Por ello en la mayoría de las simulaciones sí se ejerce la opción.

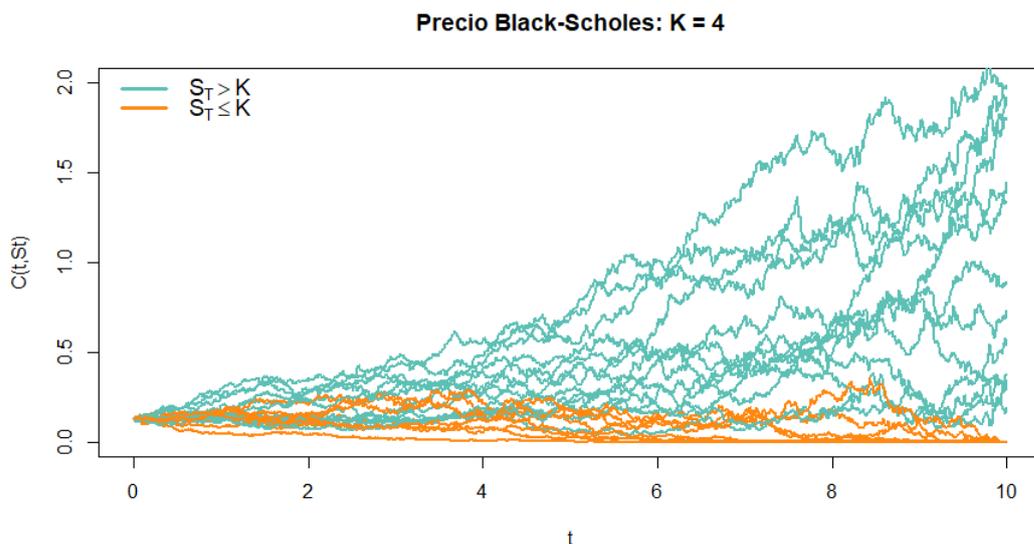


Figura 6.10: Simulación precio opción con $K = 4$.

En la Figura 6.11 se hace la misma simulación pero usando un precio de ejercicio $K = 5$. En la mayoría de las trayectorias $C(T, S_T) = 0$. Esto se puede explicar porque $\mathbb{E}[S_T] < K$.

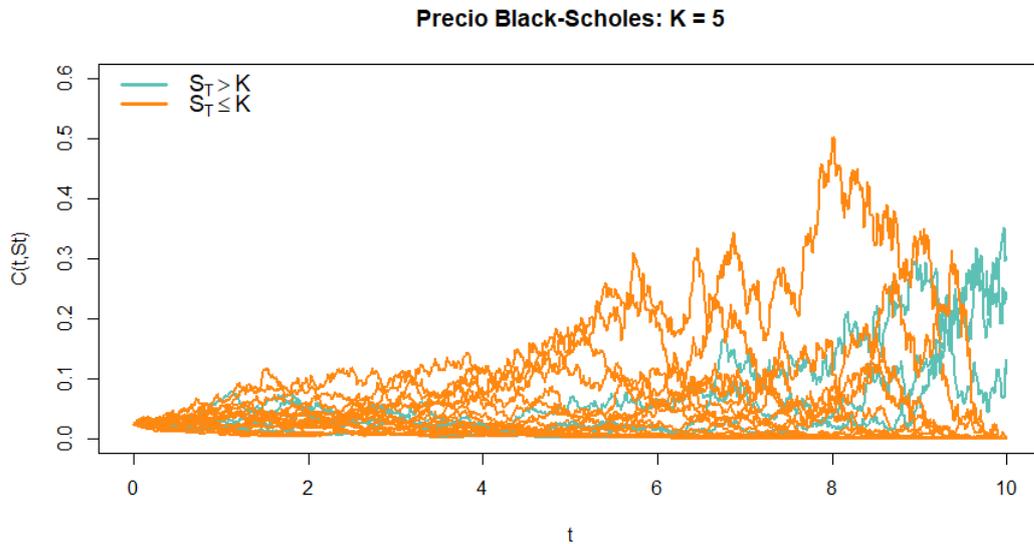


Figura 6.11: Simulación precio opción con $K = 5$.

6.2.4. Estimación paramétrica

Consideremos como activo subyacente al ETF VOO, el cual replica al índice S&P 500. Se hará estimación paramétrica para una muestra de datos del precio de cierre diario del 01/01/2019 al 31/12/2019⁵ suponiendo que se modelan con un movimiento browniano geométrico. Se consideran únicamente días hábiles, contando con 251 datos para hacer la estimación. En la Tabla 6.5 se presenta estadística descriptiva sobre la base de datos.

Min.	Q1	Mediana	Q3	Max.	Media	Var.	SD
224.5	259.4	268.0	275.3	296.7	267.3	189.2	13.8

Tabla 6.5: Estadística descriptiva datos VOO.

Al igual que en la Subsección 6.1.5 para el estimador EM usa 1 de cada 5 datos (50 datos en total). Los parámetros iniciales son los estimadores máximo verosímiles para la información incompleta. El algoritmo EM se ejecutó con 1000 iteraciones y los parámetros obtenidos son el promedio de las últimas 100 iteraciones.

Los tamaños de paso usados en la estimación fueron de $\Delta t_{MLE} = \frac{1}{251}$ y $\Delta t_{EM} = \frac{1}{49}$. La Tabla 6.6 muestra los estimadores obtenidos usando ambos métodos de estimación paramétrica: máximo verosímil y EM.

⁵Los datos se encuentran en el archivo `V00.csv` en la carpeta Bases de datos del repositorio de Github.

	μ	σ
MLE	0.256812	0.124395
EME	0.254534	0.104331

Tabla 6.6: Estimadores precio índice S&P500

En la Figuras 6.12 y 6.13 se grafican los precios reales, así como simulaciones de precios hechas a partir de los parámetros estimados. El valor inicial de las simulaciones es el valor inicial de los datos reales. El comportamiento de las trayectorias simuladas es similar a los precios verdaderos.

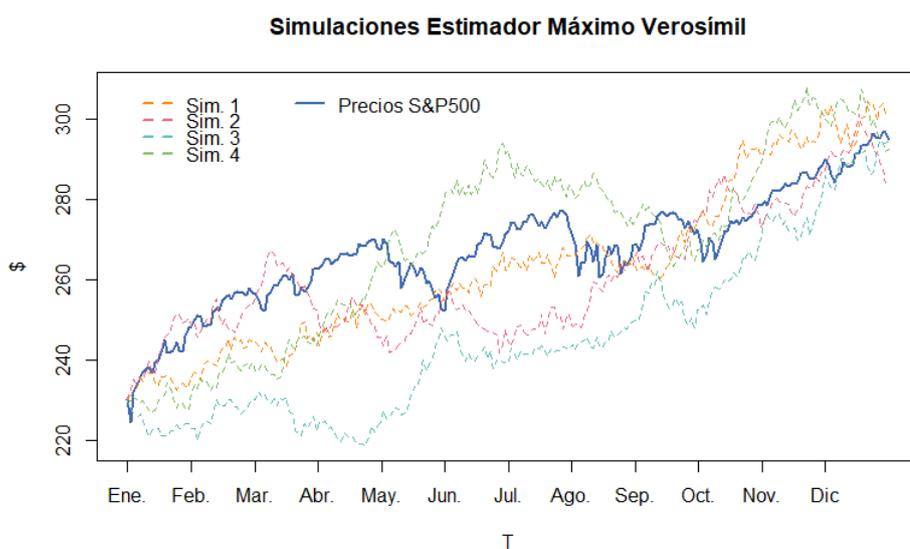


Figura 6.12: Precios índice: real y simulados.

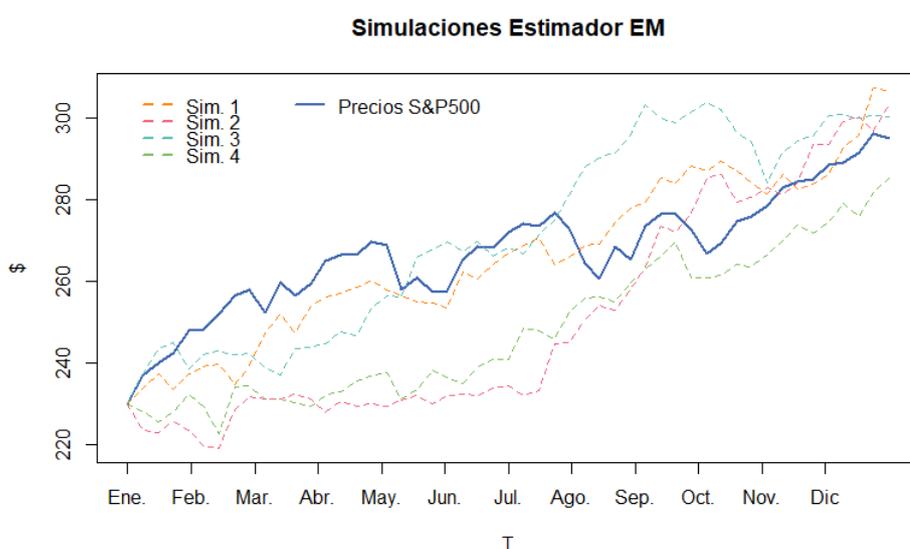


Figura 6.13: Precios índice: real (incompleto) y simulados.

Utilizando los estimadores máximo verosímiles se simulan precios del ETF. Con los precios se valúa una opción con precio de ejercicio $K = 300$ y tasa $r = 10\%$. En la Figura 6.14 se muestra una simulación donde se ejerce la opción, en la Figura 6.15 un caso en el que no.

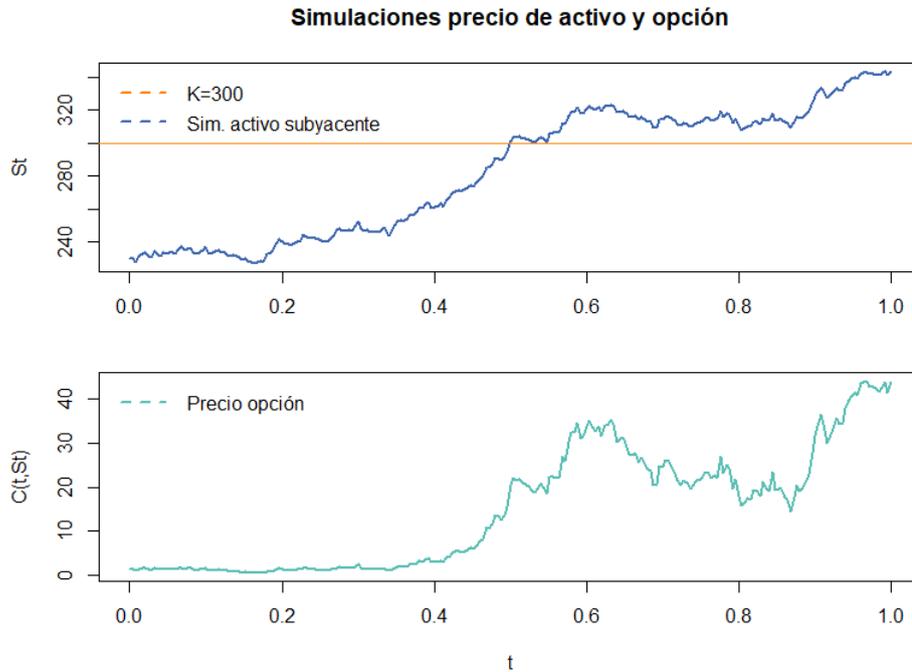


Figura 6.14: Opción con valor final positivo

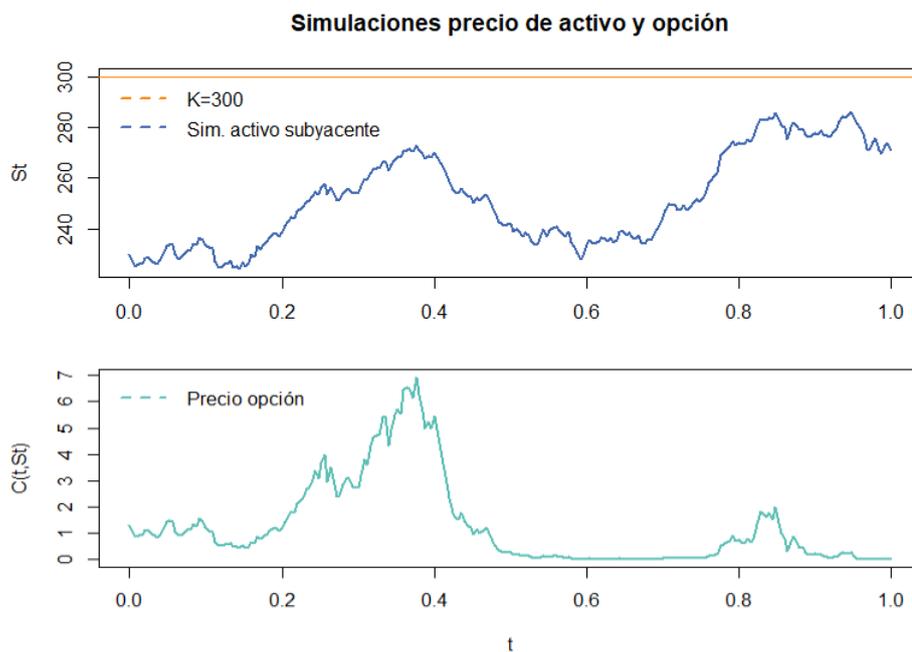


Figura 6.15: Opción con valor final cero

Validación del modelo

Se valida que los datos de precios efectivamente sean modelados con el proceso y parámetros propuestos. Aplicando la transformación de Lamperti a 6.2.2 se tiene que

$$d \ln \left(\frac{S_t}{S_0} \right) = \left(\frac{\mu}{\sigma} - \frac{1}{2} \sigma \right) + \sigma dB_t.$$

Por lo tanto se definen los residuales como

$$\widehat{\Delta B_{t_i}} = \frac{1}{\sigma} \ln \left(\frac{S_{t_i}}{S_{t_{i-1}}} \right) - \left(\mu - \frac{1}{2} \sigma^2 \right) \Delta t_i, \quad i = 2, \dots, 252.$$

Donde $\hat{\mu}$ y $\hat{\sigma}$ son parámetros estimados.

Para comparar los residuales obtenidos con los estimadores MLE se generan 1000 vectores normales de tamaño 250, con media 0 y varianza $\Delta t_{MLE} = \frac{1}{250}$. Para cada muestra se realiza una prueba de Kolmogorov-Smirnov y los p -values obtenidos se promedian obteniendo p -value = 0.10496.

El mismo proceso se hace para los residuales resultantes de los estimadores EM, pero con un vector normal de tamaño 50 y varianza $\Delta t_{EM} = \frac{1}{50}$. Se obtiene p -value = 0.13023.

Los 1000 vectores normales se suman y se multiplican por $\frac{1}{\sqrt{1000}}$. El vector normal resultante se compara con los residuales por medio de gráficos QQ, los cuales se pueden ver en la Figura 6.16

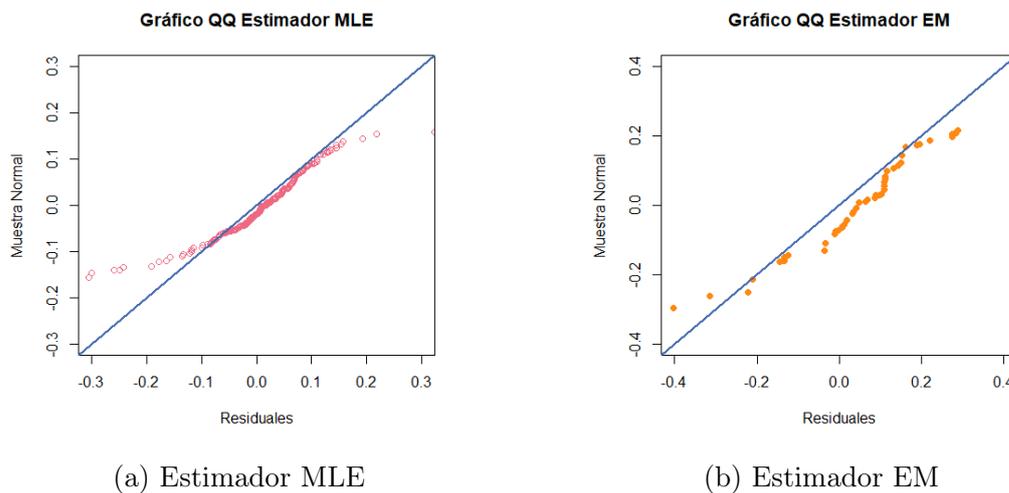


Figura 6.16: gráficos QQ: precio VOO

Por medio de los gráficos QQ de la Figura y de las pruebas Kolmogorov-Smirnov mostradas a continuación, se puede concluir que el modelo y los parámetros estimados ajustan a la base de datos.⁶

⁶Las muestras utilizadas para la validación, así como los p -values obtenidos en cada muestra se encuentran en el archivo `Muestras_Validación_Opciones.xlsx` en el repositorio de Github en la carpeta Bases de datos

Conclusiones

Este trabajo tenía como objetivo poder dar un contexto general, tanto teórico como computacional, sobre procesos de difusión, ecuaciones diferenciales estocásticas y métodos de estimación paramétrica para los procesos de difusión. El contexto fue dado y utilizado en aplicaciones relacionadas a finanzas.

El principal logro fue poder ajustar bases de datos reales a los modelos de Vasicek y Black-Scholes. Además del ajuste se hizo estimación paramétrica con los datos. Se evidenció que es posible trabajar con bases de datos incompletas. Por lo tanto se puede concluir que el estimador EM para procesos de difusión es útil en casos de información faltante. Los estimadores implementados son precisos y eficientes.

Los dos modelos presentados en el Capítulo 6 son muy básicos pero útiles como primer acercamiento. Como siguiente paso se podría trabajar con modelos y problemas más complejos, por ejemplo con el modelo CIR donde las tasas de interés son siempre no negativas. Para la valuación de opciones un siguiente paso serían las opciones americanas; que pueden ser ejercidas en cualquier momento previo a la fecha de vencimiento. Otro posible camino es estudiar y hacer estimación paramétrica para procesos de difusión continuos distintos al proceso de Vasicek y al movimiento browniano geométrico, o trabajar con difusiones con saltos.

Bibliografía

- [1] URL: <https://search.r-project.org/CRAN/refmans/mvnormalTest/html/mardia.html>.
- [2] Fernando Baltazar-Larios y Sergio Iván López Ortega. *Simulación Estocástica*. 2023.
- [3] Mogens Bladt y Michael Sørensen. “Simple simulation of diffusion bridges with application to likelihood inference for diffusions”. En: *Bernoulli* 20.2 (mayo de 2014). DOI: 10.3150/12-bej501. URL: <https://doi.org/10.3150%5C%2F12-bej501>.
- [4] Damiano Brigo et al. “A stochastic processes toolkit for Risk Management”. En: *SSRN Electronic Journal* (nov. de 2007). DOI: 10.2139/ssrn.1109160.
- [5] Jakob Croghan, John Jackman y K. Jo Min. “Estimation of Geometric Brownian Motion Parameters for Oil Price Analysis”. En: *IISE Annual Conference Proceedings* (ene. de 2017), págs. 1858-1863.
- [6] Arthur Pentland Dempster, Nan McKenzie Laird y Donald Bruce Rubin. “Maximum Likelihood from Incomplete Data via the EM Algorithm”. En: *Journal of the Royal Statistical Society*. 39.1 (1977), págs. 1-38.
- [7] Rick Durrett. *Probability: Theory and Examples*. 5.^a ed. Cambridge, 2019.
- [8] José Carlos García Franco y Onward. “An Analysis of Example”. En: (2023). URL: <https://api.semanticscholar.org/CorpusID:15749742>.
- [9] Axel Gerebrink et al. “Maximum Likelihood calibration of the Vasicek model to the Swedish interest rate market”. En: (2019).
- [10] Stefano M. Iacus. *Simulation and Inference for Stochastic Differential Equations*. Springer, 2008.
- [11] Peter E. Kloeden y Eckhard Platen. *Numerical Simulation of Stochastic Differential Equations*. 3.^a ed. Springer, 1999.
- [12] Gregory F. Lawler. *Stochastic Calculus: An Introduction with Applications*. 2023.
- [13] Richard A Levine y George Casella. “Implementations of the monte carlo em algorithm”. En: *Journal of Computational and Graphical Statistics* 10.3 (2001), págs. 422-439. DOI: 10.1198/106186001317115045.
- [14] Kantilal Vardichand Mardia. “Measures of Multivariate Skewness and Kurtosis with Applications”. En: *Biometrika* 57.3 (dic. de 1970), págs. 519-530.
- [15] Søren Feodor Nielsen. “The stochastic EM algorithm: Estimation and asymptotic results”. En: *Bernoulli* 6.3 (2000), págs. 457-489. DOI: 10.2307/3318671.
- [16] Bridget Nkemnole y Olaide Abass. “Estimation of geometric brownian motion model with a T-distribution-based particle filter”. En: *Journal of Economic and Financial Sciences* 12.1 (feb. de 2019). DOI: 10.4102/jef.v12i1.159.

- [17] Bernt Øksendal. *Stochastic Differential Equations An Introduction with Applications*. 5.^a ed. Springer, 2003.
- [18] Nicholas Privault. *Notes on Stochastic Finance*. 2013.
- [19] Luis Rincón. *Introducción a la Probabilidad*. 2.^a ed. Las prensas de ciencias, 2016.
- [20] Luis Rincón. *Introducción a los Procesos Estocásticos*. 3.^a ed. Las prensas de ciencias, 2012.
- [21] Lola Serrano. “Proyecto II: Cálculo Estocástico Aplicado”. URL: https://github.com/lolaserrano28/ProyectoII/blob/main/Notas_Exposici%C3%B3n_Proyecto_II.pdf.
- [22] Steven Shreve. *Stochastic Calculus and Finance*. Prentice Hall, jul. de 1997.
- [23] Marco Taboga. *Maximum likelihood estimation*. URL: <https://www.statlect.com/fundamentals-of-statistics/maximum-likelihood>.
- [24] Cheng Yong Tang. *Parameter estimation and bias correction for diffusion processes, and, A nonparametric approach to census population size estimation*. 2008.
- [25] Oldrich Alfons Vasicek y Francisco Venegas-Martínez. “Modelos de la Estructura de plazos de las tasas de interés: Revisión, Tendencias y Perspectivas”. En: *Revista Mexicana de Economía y Finanzas* 16.2 (abr. de 2021). DOI: 10.21919/remef.v16i2.587.
- [26] Jaime Vázquez Alamilla et al. *Inferencia estadística para estudiantes de ciencias*. 1.^a ed. Las prensas de ciencias, 2019.