



UNIVERSIDAD NACIONAL AUTÓNOMA  
DE MÉXICO

---

---

FACULTAD DE CIENCIAS

MODELOS DE VOLATILIDAD  
ESTOCÁSTICA BIVARIADOS APLICADOS A  
DATOS FINANCIEROS: TIPO DE CAMBIO Y  
RESERVAS INTERNACIONALES

T E S I S

QUE PARA OBTENER EL TÍTULO DE:

ACTUARIA

PRESENTA:

ITZEL LÓPEZ HERNÁNDEZ



DIRECTORA DE TESIS

DRA. ELIANE REGINA RODRIGUES

CIUDAD UNIVERSITARIA, CD. MX., 2021



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.



# Índice general

<b>1. Introducción</b>	<b>7</b>
<b>2. Resultados preliminares</b>	<b>9</b>
2.1. Cadenas de Markov . . . . .	9
2.2. Inferencia Bayesiana . . . . .	14
2.3. Selección de modelos . . . . .	16
<b>3. Muestreo de Gibbs y Metropolis-Hastings</b>	<b>19</b>
3.1. El muestreo de Gibbs . . . . .	19
3.1.1. Descripción del algoritmo. . . . .	20
3.2. El algoritmo Metropolis-Hastings . . . . .	21
3.2.1. Descripción del algoritmo . . . . .	21
<b>4. Modelos de volatilidad estocástica</b>	<b>23</b>
4.1. El modelo univariado . . . . .	23
4.2. Modelos de Volatilidad Estocástica Bivariados . . . . .	26
4.3. Modelos utilizados . . . . .	28
4.3.1. Modelo I . . . . .	28
4.3.2. Modelo II . . . . .	28
4.3.3. Modelo III . . . . .	28
4.4. Análisis Bayesiano . . . . .	29
<b>5. Aplicación</b>	<b>31</b>
5.1. Introducción . . . . .	31
5.2. Descripción de los datos . . . . .	31
5.3. Distribuciones a priori . . . . .	35
5.3.1. Modelo I . . . . .	35
5.3.2. Modelo II . . . . .	35
5.3.3. Modelo III . . . . .	36
5.4. Estimación de los parámetros . . . . .	36

5.5. Resultados obtenidos . . . . .	38
5.6. Discusión . . . . .	41
<b>6. Conclusiones</b>	<b>45</b>
<b>Apéndices</b>	<b>47</b>
<b>Bibliografía</b>	<b>63</b>

# Agradecimientos

Agradezco a mi madre, Alma, por ser una mujer extraordinaria, por estar a mi lado en los buenos y en los malos momentos, por todo el amor y la dedicación que me ha dado. Gracias mami, esto no hubiera sido posible sin ti.

Al Sr. Alfonso por ser más un padre que un tío, por darme amor, por consentirme y por ver en mí a una hija. Gracias tío, allá donde estés te llevo conmigo.

A mi familia, porque a pesar de las dificultades nos hemos mantenido unidos. Gracias por brindarme un hogar hermoso y todo ese apoyo incondicional.

Agradezco a Hugo por llegar a mi vida y quedarse. Gracias por todo el apoyo y todo el amor que me has dado, por creer en mí siempre, porque a tu lado soy una mejor persona.

A mis amigos, especialmente Ana y Lucia, recorrimos juntas este camino y por fin hemos llegado. Gracias por siempre estar.

Agradezco especialmente a la Dra. Eliane Rodrigues por aceptar la dirección de la tesis, por su tiempo y dedicación para hacer este trabajo posible. Gracias por sus consejos, su esfuerzo y su apoyo incondicional.

Agradezco a la Dra. Ana Meda Guardiola, a la Dra. María Begoña Fernández Fernández, a la Dra. Ruth Selene Fuentes García y al Mat. Carlos Nathanael Chávez Saab, sinodales de esta tesis, por el tiempo dedicado a la revisión de este trabajo y por el esfuerzo extra que supone trabajar en estos tiempos tan difíciles.

Gracias a la UNAM, especialmente a la E.N.P. 9 “Pedro de Alba” y a la Facultad de Ciencias. Formar parte de esta gran Universidad es un honor y un gran orgullo.



# Capítulo 1

## Introducción

El objetivo de esta tesis es realizar el análisis de datos financieros a partir de la aplicación de modelos de volatilidad estocástica. Comúnmente este análisis se realiza utilizando como herramienta los modelos construidos a partir de series de tiempo. En el presente trabajo se optó por presentar un enfoque bayesiano para determinar la correlación existente entre el tipo de cambio peso-dólar y las reservas internacionales de México.

La elección de estas variables financieras fue motivada por el complejo entorno económico global que actualmente se vive, y la forma en que los sucesos financieros en un país en particular pueden afectar a gran escala la situación financiera de muchos otros. Tal es el caso de México y Estados Unidos, que se acentúa por la gran relación comercial que existe entre ambas naciones. Actualmente los estudios sobre la relación entre el tipo de cambio y las reservas internacionales de un país son muy limitados, y están enfocados a economías desarrolladas, o bien, a amplios conjuntos de países. Es aquí donde surge la pregunta ¿Existe relación entre estas variables en la economía mexicana?.

Se tomará como base el desarrollo presentado en Achcar et. al. (2010), Zozolotto (2010) y De Jesus Romo (2013), donde los modelos de volatilidad estocástica son aplicados a datos de contaminación atmosférica.

La principal contribución del presente trabajo es la aplicación de los modelos a los datos de tipo de cambio peso-dólar y las reservas internacionales de México.

Esta tesis está organizada de la siguiente forma. En el capítulo 2 se presentan los principales resultados teóricos y matemáticos utilizados para el desarrollo del tema central. En la primera sección se presentan los resultados básicos de cadenas de Markov. En la segunda sección se describen los

conceptos fundamentales de la inferencia bayesiana. En la tercera sección se presenta el criterio para la selección del modelo que mejor describe el comportamiento de los datos.

En el capítulo 3 se presentan los dos algoritmos que fueron implementados para generar los valores que permitieron la estimación de los parámetros en los modelos utilizados, el muestreo de Gibbs y el algoritmo de Metropolis-Hastings.

En el capítulo 4 se presentan los modelos de volatilidad estocástica. En la primera sección se describe el caso univariado, que sirve como base de la segunda sección, donde se introducen los modelos bivariados. En la tercera sección se muestran los modelos que serán utilizados en la parte aplicada, y en la última sección se realiza el análisis bayesiano de los mismos.

El capítulo 5 contiene la descripción de los datos, así como los resultados obtenidos al aplicar los modelos y su interpretación. Se presentan algunos datos y gráficos de interés para comprender el comportamiento del modelo. También se realiza la discusión relativa a los resultados producidos por el modelo que fue seleccionado a partir de los resultados numéricos, y se presenta una estimación futura del comportamiento del tipo de cambio y las reservas internacionales para un periodo de 12 meses, correspondientes al año 2019.

En el capítulo 6 se encuentran las conclusiones a las que se llegó después de realizar la aplicación de los modelos teóricos y se presentan otras alternativas para el estudio de estos datos.

Finalmente, en los apéndices se presentan algunos gráficos relevantes obtenidos mediante el software Winbugs (<https://www.mrc-bsu.cam.ac.uk/software/bugs/the-bugs-project-winbugs/>) y la forma en la que se incorporaron los datos en los códigos presentados en Zozzoloto (2010).

Adicionalmente, se presentan los códigos de los programas utilizados para obtener las volatilidades estocásticas estimadas, así como los valores de las mediciones futuras predecidas utilizando el modelo seleccionado y sus parámetros estimados.

## Capítulo 2

# Resultados preliminares

En este capítulo serán presentados algunos conceptos básicos relacionados con cadenas de Markov e inferencia bayesiana.

### 2.1. Cadenas de Markov

Los conceptos tratados a continuación serán de utilidad al trabajar con los métodos Monte Carlo vía cadenas de Markov (MCMC) que permiten estimar los parámetros que aparecen en los modelos de volatilidad. Los temas tratados en esta sección fueron consultados en Ross (1996), Grimmett y Stirzaker (2001), Karlin y Taylor (1975), entre otros.

**Definición 1.** *Se define un proceso estocástico como una colección de variables aleatorias  $X = \{X_t | t \in T\}$ , con  $T$  un conjunto de índices, donde  $X_t$  toma valores en un mismo conjunto  $S$  llamado espacio de estados.*

**Definición 2.** *Dado  $T = \{0, 1, \dots\}$ , una cadena de Markov  $X = \{X_n : n \geq 0\}$  es un proceso estocástico con espacio de estados  $S$  numerable o finito que satisface la propiedad de Markov, esto es, para todo  $n \geq 0$ ,  $x_0, \dots, x_{n-1}, x \in S$  y  $A \subset S$  se cumple:*

$$\mathbb{P}(X_{n+1} \in A | X_n = x, X_{n-1} = x_{n-1}, \dots, X_0 = x_0) = \mathbb{P}(X_{n+1} \in A | X_n = x). \quad (2.1)$$

En otras palabras, la ecuación (2.1) dice que, dado el estado actual, los estados pasados y futuros de la cadena son independientes entre sí. Cuando se tiene  $A = \{y\}$ , esta propiedad puede ser expresada como:

$$\mathbb{P}(X_{n+1} = y | X_n = x, X_{n-1} = x_{n-1}, \dots, X_0 = x_0) = \mathbb{P}(X_{n+1} = y | X_n = x), \quad (2.2)$$

con  $x_0, \dots, x_{n-1}, x, y \in S$ ,  $n \geq 0$ .

La ecuación (2.2) dice que la probabilidad de estar en el estado  $y$  al instante  $(n + 1)$ , dado que la cadena siguió la trayectoria  $\{x_0, x_1, \dots, x_n\}$ , sólo depende del estado en el tiempo inmediato anterior, es decir, el estado en el instante  $n$ . Entonces se dice que el proceso no tiene memoria. En general, las probabilidades en la propiedad de Markov dependen de  $x$ ,  $A$  y  $n$ . Si no dependen de  $n$  entonces se dice que la cadena es homogénea, en tal caso es conveniente definir el siguiente concepto.

**Definición 3.** Se definen las probabilidades de transición del estado  $x$  al estado  $y$  en  $m$  pasos como:

$$P_{xy}^{(m)} = \mathbb{P}(X_{n+m} = y | X_n = x), \quad m \geq 0, \quad x, y \in S. \quad (2.3)$$

**Observación 1.**  $P_{xy}^{(m)}$  satisface:

1.  $P_{xy} \geq 0$  con  $x, y \in S$ .
2.  $\sum_{z \in S} P_{xz} = 1$ ,  $x \in S$ .

Cuando  $m = 1$ , se tiene la probabilidad de transición de la cadena de Markov y se expresa por  $P_{xy} = \mathbb{P}(X_{n+1} = y | X_n = x)$ .

De la ecuación (2.3) se derivan las llamadas ecuaciones de Chapman-Kolmogorov.

**Proposición 2.1.1.** Para una cadena de Markov  $X = \{X_n : n \geq 0\}$  con espacio de estados  $S$  y para toda  $n, m \in \mathbb{N}$  y toda pareja  $x, y \in S$  se cumple:

$$P_{xy}^{(m+n)} = \mathbb{P}[X_{n+m} = y | X_0 = x] = \sum_{z \in S} P_{xz}^{(m)} P_{zy}^{(n)}. \quad (2.4)$$

*Demostración.* Para  $m + n = 1$  se supone, sin pérdida de generalidad,  $n = 1$  y  $m = 0$ . Notar que:

$$P_{xy}^{(0)} = \begin{cases} 1, & x = y, \\ 0, & \text{en otro caso.} \end{cases}$$

Por lo tanto, se tiene que:

$$P_{xy}^{(1)} = \mathbb{P}[X_1 = y | X_0 = x] = P_{xy} = \sum_{z \in S} P_{xz}^{(1)} P_{zy}^{(0)}.$$

Suponga cierto para toda  $k \in \mathbb{N}$  con  $k < n + m$ . Se quiere demostrar que se cumple para  $k = n + m$ . Sea

$$P_{xy}^{(m+n)} = \mathbb{P}[X_{n+m} = y | X_0 = x],$$

por definición de probabilidad condicional esto es:

$$\begin{aligned} P_{xy}^{(m+n)} &= \frac{\mathbb{P}[X_{n+m} = y, X_0 = x]}{\mathbb{P}[X_0 = x]} \\ &= \sum_{z \in S} \mathbb{P}[X_{n+m} = y, X_n = z, X_0 = x] \left[ \frac{1}{\mathbb{P}[X_0 = x]} \right] \\ &= \sum_{z \in S} \frac{\mathbb{P}[X_{n+m} = y | X_n = z, X_0 = x] \mathbb{P}[X_n = z, X_0 = x]}{\mathbb{P}[X_0 = x]}. \end{aligned}$$

Por propiedad de Markov se sigue que:

$$\begin{aligned} P_{xy}^{(m+n)} &= \sum_{z \in S} \mathbb{P}[X_{n+m} = y | X_n = z] \mathbb{P}[X_n = z | X_0 = x] \\ &= \sum_{z \in S} P_{xz}^{(n)} P_{zy}^{(m)}. \end{aligned}$$

□

**Definición 4.** La distribución de la cadena al tiempo  $n$ , es decir,  $\mathbb{P}[X_n = x]$  se puede definir como el vector renglón  $\pi^{(n)}$  con componentes  $\pi^{(n)}(x)$ ,  $x \in S$ . En el caso del espacio de estados finito con  $r$  elementos,  $S = \{x_1, x_2, \dots, x_r\}$ , este es un vector  $r$ -dimensional:

$$\pi^{(n)} = (\pi^{(n)}(x_1), \dots, \pi^{(n)}(x_r)).$$

**Observación 2.** Cuando  $n = 0$ , se tiene la distribución inicial de la cadena.

**Proposición 2.1.2.** La distribución al tiempo  $n$  de la cadena está dada por:

$$\pi^{(n)}(y) = \sum_{x \in S} P_{xy}^{(n)} \pi^{(0)}(x).$$

*Demostración.* Utilizando la propiedad de probabilidad total y propiedades de probabilidad condicional se tiene:

$$\begin{aligned} \pi^{(n)}(y) &= \mathbb{P}(X_n = y) \\ &= \sum_{x \in S} \mathbb{P}(X_n = y, X_0 = x) \\ &= \sum_{x \in S} \mathbb{P}(X_n = y | X_0 = x) \mathbb{P}(X_0 = x) \\ &= \sum_{x \in S} P_{xy}^{(n)} \pi^{(0)}(x). \end{aligned}$$

□

Otro concepto que será de utilidad cuando se utilicen los algoritmos para estimar los parámetros de los modelos es la clasificación de los estados de una cadena de Markov. Se definen los siguientes conceptos:

- La probabilidad de que una cadena que comienza en el estado  $x$  llegue al estado  $y$  en el futuro se denota como:

$$\rho_{xy} = \mathbb{P}[T_y < \infty | X_0 = x],$$

donde  $T_y$  es el primer tiempo de llegada al estado  $y$ .

- El número de visitas de una cadena al estado  $y$  es:

$$N(y) = |\{n > 0 : X_n = y\}|,$$

donde, para  $A$  un conjunto,  $|A|$  es su cardinalidad.

**Definición 5.** Sean  $x, y \in S$ , se dice que  $x$  se comunica con  $y$ , denotado por  $x \leftrightarrow y$ , si existen  $m, n \geq 0$ , tales que  $P_{xy}^{(n)} > 0$  y  $P_{yx}^{(m)} > 0$ .

**Definición 6.** Se dice que una cadena es irreducible si  $x \leftrightarrow y$  para toda  $x, y \in S$ . Por tanto, una cadena es irreducible si todos los estados en un espacio de estados se comunican entre si.

**Definición 7.** Una distribución  $\pi$  es llamada distribución estacionaria de una cadena de Markov con probabilidad de transición  $P_{xy}$ ,  $x, y \in S$  si:

$$\sum_{x \in S} \pi(x) P_{xy} = \pi(y),$$

con  $\sum_{x \in S} \pi(x) = 1$ .

**Definición 8.** Sea  $X = \{X_n : n \geq 0\}$  una cadena de Markov homogénea con probabilidad de transición  $P_{xy}$ ,  $x, y \in S$ , y distribución estacionaria  $\pi$ . Se dice que la cadena es reversible si cumple:

$$\pi(x) P_{xy} = \pi(y) P_{yx}, \tag{2.5}$$

para  $x, y \in S$ .

**Proposición 2.1.3.** Sea  $X = \{X_n : n \geq 0\}$  una cadena de Markov homogénea con probabilidad de transición  $P_{xy}$ ,  $x, y \in S$ . Si la cadena es reversible, es decir, cumple la ecuación (2.5), entonces  $\pi$  es una distribución estacionaria para  $X$ .

*Demostración.* A partir de la ecuación (2.5) se tiene que:

$$\begin{aligned} \sum_{x \in S} \pi(x) P_{xy} &= \sum_{x \in S} \pi(y) P_{yx} \\ &= \pi(y) \sum_{x \in S} P_{yx} \\ &= \pi(y), \end{aligned}$$

pues  $\sum_{x \in S} P_{yx} = 1$ . □

**Definición 9.** Se dice que un estado  $x$  tiene periodo  $d_x$  si:

$$d_x = \text{mcd}\{n > 0 \mid P_{xx}^{(n)} > 0\}.$$

**Definición 10.** Un estado  $x$  es aperiódico si  $d_x = 1$ . Una cadena de Markov será aperiódica si todos sus estados son aperiódicos.

**Definición 11.** Un estado es recurrente si la cadena de Markov, que inicia en  $y$ , regresa a  $y$  con probabilidad 1, es decir,  $\rho_{yy} = 1$ . Se dice que es transitorio si tiene una probabilidad positiva de no regresar a  $y$ , es decir,  $\rho_{yy} < 1$ .

**Definición 12.** Dada una cadena de Markov  $X = \{X_n : n \geq 0\}$ , el estado  $x \in S$  es transitorio si y sólo si:

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

Se sigue entonces que, para un estado  $x$  transitorio,  $P_{xx}^{(n)} \rightarrow 0$  cuando  $n \rightarrow \infty$ .

**Definición 13.** Sea  $\mu_{xx}$ , que denota el número esperado de pasos necesarios para volver al estado  $x$ . Entonces:

$$\mu_{xx} = \begin{cases} \infty, & \text{si } x \text{ es transitorio,} \\ \sum_{n=1}^{\infty} n f_{xx}^{(n)}, & \text{si } x \text{ es recurrente,} \end{cases} \quad (2.6)$$

donde  $f_{xx}^{(n)} = \mathbb{P}[X_n = x, X_k \neq x, k = 1, 2, \dots, n-1 \mid X_0 = x]$  es la probabilidad de que partiendo de  $x$ , la cadena regrese a  $x$  por primera vez en  $n$  pasos.

**Definición 14.** Si el estado  $x$  es recurrente, entonces se dice que es un estado positivo recurrente si  $\sum_{n=1}^{\infty} n f_{xx}^{(n)} < \infty$ , y será recurrente nulo si  $\sum_{n=1}^{\infty} n f_{xx}^{(n)} = \infty$ .

**Teorema 2.1.4.** *Una cadena de Markov irreducible y aperiódica pertenece a alguna de las siguientes clases:*

- (i) *Todos los estados son transitorios o recurrentes nulos, entonces no existe una distribución estacionaria;*
- (ii) *Todos los estados son positivos recurrentes, esto es, existe el límite:*

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

*y es estrictamente positivo, para toda  $x, y \in S$ . En este caso, existe una distribución estacionaria única  $\pi(x)$ , tal que:*

$$\lim_{n \rightarrow \infty} P_{yx}^{(n)} = \pi(x).$$

*Demostración.* Ver Ross (1996). □

**Definición 15.** *Cuando la cadena de Markov irreducible y aperiódica es de la clase (ii) se dice que es una cadena ergódica.*

## 2.2. Inferencia Bayesiana

El enfoque bayesiano tiene por objetivo describir el comportamiento de las observaciones  $D = \{x_1, \dots, x_n\}$  que se supone son resultado de un fenómeno aleatorio que puede ser descrito a través de un modelo matemático. Este modelo considera una función de probabilidad  $f(D|\theta)$  que refleja la probabilidad de que las observaciones  $D$  sean derivadas de dicho modelo. En este escenario  $\theta$  es el parámetro de una familia de posibles distribuciones que describen las observaciones, y a la cual pertenece  $f(\cdot|\theta)$ .

El caso típico es aquel donde se tiene una muestra  $\mathbf{x} = (x_1, \dots, x_n)$  extraída de una población que se distribuye de acuerdo a  $f(\mathbf{x}|\theta)$ , donde las observaciones  $x_i$  son independientes e idénticamente distribuidas.

Una descripción más detallada será dada en este capítulo, cuyo contenido se consultó en Gamerman y Lopes (2006).

**Definición 16.** *Sea  $\mathbf{x} = (x_1, \dots, x_n)$  una muestra aleatoria y  $\theta \in \Theta$  el parámetro desconocido del modelo que describe al experimento que produjo la muestra  $\mathbf{x}$ . La distribución a priori  $p(\theta)$  es la distribución del parámetro  $\theta$  que se le asigna sin tomar en cuenta la información proporcionada por la muestra. Generalmente, esta distribución puede ser descrita con la ayuda de parámetros conocidos llamados hiperparámetros.*

**Definición 17.** La función de verosimilitud del modelo, indicada por  $L(\mathbf{x}|\theta)$ , es proporcional a la densidad conjunta  $f_{x_1, \dots, x_n}(\cdot|\theta)$  de la muestra  $\mathbf{x}$  condicionada a  $\theta$ . Si las observaciones son independientes, la función de verosimilitud está dada por:

$$L(\mathbf{x}|\theta) \propto f(x_1, \dots, x_n|\theta) = \prod_{i=1}^n f_i(x_i|\theta), \quad (2.7)$$

donde  $f_i(\cdot|\theta)$  es la distribución marginal de la  $i$ -ésima coordenada de  $\mathbf{x}$ .

**Definición 18.** La distribución a posteriori, denotada por  $p(\theta|\mathbf{x})$ , es la distribución del parámetro después de incorporar la información proporcionada por la muestra  $\mathbf{x}$  que se quiere describir usando el modelo con vector de parámetros  $\theta$ .

**Proposición 2.2.1.** Sean  $\mathbf{x} = (x_1, \dots, x_n)$  una muestra aleatoria,  $\theta \in \Theta$  el parámetro desconocido del modelo que describe al experimento que produjo  $\mathbf{x}$ ,  $p(\theta)$  la distribución a priori de  $\theta$  y  $L(\mathbf{x}|\theta)$  la función de verosimilitud del modelo. La distribución a posteriori  $p(\theta|\mathbf{x})$ , que es la distribución de  $\theta$  considerando la información proporcionada por la muestra  $\mathbf{x}$ , satisface:

$$p(\theta|\mathbf{x}) = \frac{f(\mathbf{x}|\theta)p(\theta)}{f(\mathbf{x})} \propto L(\mathbf{x}|\theta)p(\theta), \quad (2.8)$$

donde  $f(\mathbf{x}) = \int f(\mathbf{x}|\theta)p(\theta)d\theta$ .

*Demstración.* Sea  $f(\mathbf{x}, \theta)$  la función de densidad conjunta de  $\mathbf{x}$  y  $\theta$ . Por propiedades de probabilidad condicional se tiene:

$$\begin{aligned} p(\theta|\mathbf{x}) &= \frac{f(\theta, \mathbf{x})}{f(\mathbf{x})} \\ &= \frac{f(\mathbf{x}|\theta)p(\theta)}{f(\mathbf{x})} \\ &\propto f(\mathbf{x}|\theta)p(\theta) \propto L(\mathbf{x}|\theta)p(\theta). \end{aligned}$$

□

**Definición 19.** Sea  $P$  una familia de distribuciones a priori y sea  $F$  una clase de funciones de densidad  $f(\mathbf{x}|\theta)$ . Se dice que una familia  $P$  es conjugada para  $F$  si para todo  $p(\theta) \in P$  y para toda  $f \in F$  la distribución a posteriori  $p(\theta|\mathbf{x})$  pertenece a la familia de distribuciones  $P$ .

A continuación se presenta un ejemplo que ilustra las definición del concepto de familia conjugada.

**Ejemplo 1.** Suponga que  $X$  se distribuye  $Poisson(\theta)$  y que  $\theta$ , que es el parámetro de la distribución, se distribuye a priori  $Gamma(\alpha, \beta)$ , con  $\alpha, \beta > 0$  conocidos. El objetivo es verificar que la distribución gamma es una familia conjugada para la función de verosimilitud de un modelo Poisson.

Por un lado se tiene que:

$$\begin{aligned} p(\theta|\mathbf{x}) \propto L(\mathbf{x}|\theta)p(\theta) &= e^{-\theta} \frac{\theta^x}{x!} \frac{\beta^\alpha}{\Gamma(\alpha)} \theta^{\alpha-1} e^{-\beta\theta} \\ &= \frac{1}{x!} \frac{\beta^\alpha}{\Gamma(\alpha)} \theta^x e^{-\theta} \theta^{\alpha-1} e^{-\beta\theta} \\ &= \frac{\beta^\alpha}{x! \Gamma(\alpha)} \theta^{x+\alpha-1} e^{-(1+\beta)\theta}. \end{aligned}$$

Sean  $C = x + \alpha$  y  $D = 1 + \beta$  entonces:

$$p(\theta|\mathbf{x}) \propto \frac{\beta^\alpha \Gamma(C)}{D^C x! \Gamma(\alpha) \Gamma(C)} \theta^{C-1} e^{-D\theta}.$$

Tomando  $R = \frac{\beta^\alpha \Gamma(C)}{D^C x! \Gamma(\alpha)}$ , y dado que esta cantidad no depende de  $\theta$  es una constante con respecto a  $\theta$  debido a que es función de cantidades conocida, como son  $x, \alpha$  y  $\beta$  se tiene:

$$p(\theta|\mathbf{x}) \propto \frac{D^C}{\Gamma(C)} \theta^{C-1} e^{-D\theta}.$$

Por lo tanto, la distribución a posteriori pertenece a la familia de distribución gamma.

### 2.3. Selección de modelos

Para la selección del mejor modelo en la aplicación de los datos se utilizará el criterio DIC (Deviance Information Criterion). La desviación se define como:

$$D(\theta) = -2 \ln(L(\mathbf{x}|\theta)) + C, \quad (2.9)$$

donde  $\theta$  es el vector de parámetros desconocidos del modelo, y  $L(\mathbf{x}|\theta)$  su respectiva función de verosimilitud. El criterio introducido por Spiegelhalter et. al. (2002) está dado por:

$$DIC = D(\bar{\theta}) + 2p_D, \quad (2.10)$$

donde  $\bar{\theta}$  es la media de  $\theta$ .

Al componente  $p_D$  se le conoce como número efectivo de parámetros, y se calcula como:

$$p_D = \bar{D} - D(\bar{\theta}), \quad (2.11)$$

donde  $\bar{D} = \mathbb{E}(D(\theta))$  es la desviación media a posteriori que mide la bondad de ajuste del modelo a los datos. Valores menores del DIC indican mejores modelos.



## Capítulo 3

# El muestreo de Gibbs y el algoritmo de Metropolis-Hastings

El muestreo de Gibbs y el algoritmo de Metropolis-Hastings son algoritmos que permiten generar valores a partir de una distribución dada cuando no es posible muestrearlos directamente de ella. La forma de generar los valores a partir de la distribución de interés es construir una cadena de Markov ergódica cuya distribución límite es esta distribución. En este capítulo se presenta la descripción de estos dos algoritmos. La información fue consultada en Gamerman y Lopes (2006).

### 3.1. El muestreo de Gibbs

El muestreo de Gibbs es un algoritmo utilizado para generar una muestra aleatoria a partir de funciones de distribución multidimensionales. Sea  $f(\mathbf{x})$  la distribución de interés, donde  $\mathbf{x} = (x_1, \dots, x_d)$ ,  $d \geq 1$ .

**Definición 20.** *Se definen las distribuciones marginales condicionales completas como:*

$$f_i(x_i) = f(x_i|x_{-i}), i = 1, \dots, d,$$

donde  $x_{-i}$  es el vector  $\mathbf{x}$  sin la  $i$ -ésima coordenada, es decir:

$$x_{-i} = (x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_d), i = 1, 2, \dots, d$$

### 3.1.1. Descripción del algoritmo.

Se construirá una cadena de Markov  $X = \{X^{(n)} = (X_1^{(n)}, \dots, X_d^{(n)}) : n \geq 0\}$  ergódica cuya distribución límite es  $f(\cdot)$ , de la cual se quiere generar valores. A continuación se presenta la descripción del algoritmo mediante el cual se generarán los valores deseados.

1. Iniciar el contador de iteración de la cadena, es decir, tomar  $n = 0$  y dar valores iniciales arbitrarios a  $X$  de acuerdo a alguna distribución inicial, es decir, tomar:

$$X^{(0)} = (x_1^{(0)}, \dots, x_d^{(0)}).$$

2. Para  $n \geq 1$  suponga que el estado de la cadena es  $X^{(n-1)} = (x_1^{(n-1)}, \dots)$ , obtener un nuevo valor  $X^{(n)} = (x_1^{(n)}, \dots, x_d^{(n)})$  de la siguiente forma:

- Genere  $x_1^{(n)}$  a partir de  $f(\cdot | x_2^{(n-1)}, \dots, x_d^{(n-1)})$ .
- Genere  $x_i^{(n)}$  a partir de  $f(\cdot | x_1^{(n)}, x_2^{(n)}, \dots, x_{i-1}^{(n)}, x_{i+1}^{(n-1)}, \dots, x_d^{(n-1)})$ , para  $i = 2, \dots, d-1$ .
- Finalmente se genera  $x_d^{(n)}$  a partir de  $f(\cdot | x_1^{(n)}, \dots, x_{d-1}^{(n)})$ .

3. Actualizar el contador  $n$  a  $(n+1)$  y volver al paso 2 hasta alcanzar la convergencia.

La forma natural de obtener una muestra de tamaño  $N$  a partir de  $f$  es esperar que la cadena converja y después tomar  $N$  valores generados.

Es importante observar que el muestreo de Gibbs define una cadena de Markov, pues la iteración  $n$  depende únicamente de los valores de la cadena en el tiempo  $(n-1)$ . En este caso la probabilidad de transición es:

$$P_{x^{(n-1)}x^{(n)}} = \prod_{i=1}^d f_i(x_i^{(n)}),$$

donde  $f_i(x_i^{(n)})$  es tal que:

$$\begin{cases} f_1(x_1^{(n)}) = f(x_1^{(n)} | x_2^{(n-1)}, \dots, x_d^{(n-1)}), \\ f_i(x_i^{(n)}) = f(x_i^{(n)} | x_1^{(n)}, \dots, x_{i-1}^{(n)}, x_{i+1}^{(n-1)}, \dots, x_d^{(n-1)}), i = 1, \dots, d-1, \\ f_d(x_d^{(n)}) = f(x_d^{(n)} | x_1^{(n)}, \dots, x_{d-1}^{(n)}). \end{cases} \quad (3.1)$$

**Observación 3.** *Bajo condiciones adecuadas (ver Robert y Casella, 2004) la cadena  $X$  definida de esta forma es ergódica y converge a la distribución de interés.*

## 3.2. El algoritmo Metropolis-Hastings

En el algoritmo de Metropolis-Hastings (ver Metropolis et. al. 1953) también se considera una distribución  $f(\cdot)$  a partir de la cual se quiere extraer una muestra. Nuevamente, el objetivo es construir una cadena de Markov ergódica con probabilidad de transición  $P_{xy}$ ,  $x, y \in S$ , de tal forma que  $f(\cdot)$  sea la distribución de equilibrio de la cadena.

La probabilidad de transición  $P_{xy}$  está formada por dos elementos:

1. Un kernel de transición  $Q_{xy}$ , tal que  $Q_{xy} > 0$  para toda  $x, y \in S$  del cual sea fácil generar valores;
2. Una probabilidad  $\alpha(x, y)$  de aceptación del valor generado que se define como:

$$\alpha(x, y) = \min \left\{ \frac{f(y)Q_{yx}}{f(x)Q_{xy}}, 1 \right\}.$$

### 3.2.1. Descripción del algoritmo

1. Iniciar el contador de iteración en  $n = 0$  y elegir un valor inicial arbitrario  $x^{(0)}$ , es decir  $X^{(0)} = x^{(0)}$ .
2. Para  $n \geq 1$  Suponga que la cadena está en el paso  $(n-1)$  y su valor es  $X^{(n-1)} = x^{(n-1)}$ . Generar un valor  $y$  a partir de la densidad  $Q_{x^{(n-1)}}$ .
3. Evaluar la probabilidad de aceptación  $\alpha(x^{(n-1)}, y)$ . Generar  $u$  a partir de una distribución uniforme en  $(0, 1)$ . Si  $u \leq \alpha(x^{(n-1)}, y)$  aceptar el valor generado, en otro caso mantener  $x^{(n-1)}$ . Si el nuevo valor es aceptado, entonces  $X^{(n)} = y$ . Si se rechaza, entonces  $X^{(n)} = X^{(n-1)} = x^{(n-1)}$  y la cadena no se mueve.
4. Cambiar el contador a  $(n+1)$  y volver al paso 2 hasta alcanzar la convergencia.

Se observa que el kernel de transición  $(Q_{xy})_{x, y \in S}$  arroja un posible valor para mover la cadena, sin embargo, es aceptado o no de acuerdo a una probabilidad  $\alpha(\cdot, \cdot)$ , y por ello la densidad  $Q$  es conocida como densidad propuesta.

**Proposición 3.2.1.** *La cadena de Markov  $X = \{X_n : n \geq 0\}$  construida a través del algoritmo de Metropolis-Hastings tiene como matriz de transición*

a  $P = (P_{xy})_{x,y \in S}$ , dada por:

$$P_{xy} = \begin{cases} Q_{xy}\alpha(x, y), & \text{si } x \neq y, \\ Q_{xx} + \sum_{y \neq x} Q_{xy}(1 - \alpha(x, y)), & \text{si } x = y. \end{cases} \quad (3.2)$$

*Demostración.* Para  $x \neq y$ , se define:

$$\begin{aligned} P_{xy} &= \mathbb{P}[\text{proponer } y, \text{aceptar } y] \\ &= \mathbb{P}[\text{proponer } y] \mathbb{P}[\text{aceptar } y] \\ &= Q_{xy}\alpha(x, y) \end{aligned}$$

Si  $x = y$ , se consideran dos casos: proponer  $y = x$  y aceptarlo o proponer  $y \neq x$  y no aceptar. Esto es:

$$\begin{aligned} P_{xy} &= Q_{xx}\mathbb{P}[\text{aceptar } y = x] + \sum_{y \neq x} Q_{xy}(1 - \alpha(x, y)) \\ &= Q_{xx} + \sum_{y \neq x} Q_{xy}(1 - \alpha(x, y)) \end{aligned}$$

□

**Proposición 3.2.2.** *La cadena de Markov con probabilidad de transición (3.2) es ergódica y tiene como distribución estacionaria a  $f(\cdot)$ .*

*Demostración.* Observar que por la definición 8 del capítulo 2 y por la proposición 2.1.3 basta probar para  $x \neq y$ . La ecuación (2.5) se satisface, por lo tanto  $P_{xy} = Q_{xy}\alpha(x, y)$ , se tiene que:

$$\begin{aligned} f(y)P_{xy} &= f(y)Q_{yx}\alpha(y, x) \\ &= f(y)Q_{yx} \left( \frac{f(x)Q_{xy}}{f(y)Q_{yx}} \right) \\ &= f(x)Q_{xy} \\ &= f(x)P_{xy}. \end{aligned}$$

Por lo tanto la cadena  $f(\cdot)$  es la distribución estacionaria de la cadena.

Para mostrar la ergodicidad se toma  $S$  discreto. Note que por definición  $Q_{xx} \geq 0$ , y por lo tanto toda  $x \in S$  tiene periodo uno. Note también que dado que  $Q_{xy} \geq 0$  se puede ir de cualquier estado a otro en el espacio  $S$ . De esta forma la cadena es irreducible. También se cumple la recurrencia positiva, por lo tanto la cadena será ergódica.

Para el caso en que  $S$  es continuo ver, por ejemplo, Robert y Casella (2004). □

## Capítulo 4

# Modelos de volatilidad estocástica

En este capítulo se presentan los modelos de volatilidad estocástica que serán utilizados en la aplicación a los datos financieros. Se utilizarán modelos bivariados, pero se presenta también la versión univariada. El tema presentado aquí puede ser consultado en Achcar et. al. (2010) y De Jesus Romo (2013).

### 4.1. El modelo univariado

**Definición 21.** *La volatilidad,  $\sigma^2(t)$ , de cualquier serie de tiempo  $Y(t)$ ,  $t = 1, \dots, N$ , por definición, es el segundo momento condicional de la serie  $Y(t)$  dado su valor hasta el tiempo  $t - 1$ , es decir:*

$$\sigma^2(t) = \mathbb{E}[Y^2(t)|Y(1), Y(2), \dots, Y(t-1)].$$

Suponga que la serie puede ser modelada de la siguiente manera:

$$Y(t) = \sigma(t)\epsilon(t),$$

donde  $\epsilon(t)$  es el vector del error, independiente de  $\sigma(t)$ , y tal que su distribución sea  $N(0, 1)$ . Suponga que la volatilidad de  $Y(t)$  está dada por:

$$\sigma^2(t) = e^{h(t)}, \tag{4.1}$$

donde  $h(t)$  es una variable latente que explica cualquier fenómeno no modelado, pero que influye en el valor de  $Y(t)$ .

De (4.1) se tiene que

$$\sigma(t) = e^{\frac{h(t)}{2}}. \quad (4.2)$$

Suponga  $h(t)$  que satisface un modelo autorregresivo de primer orden AR(1), definido por:

$$\begin{cases} h(1) = \mu + \eta(1), \\ h(t) = \mu + \phi[h(t-1) - \mu] + \eta(t), \quad \text{para } t = 2, \dots, N, \end{cases} \quad (4.3)$$

donde, para asegurar que  $\{h(t)\}_{t=1, \dots, N}$  sea estacionario, se supone que  $|\phi| < 1$ . Adicionalmente, se supone que  $\eta(t)$  tiene distribución  $N(0, \sigma_\eta^2)$ ,  $t = 1, \dots, N$  y son variables aleatorias independientes. Así, el modelo univariado queda definido como:

$$Y(t) = e^{\frac{h(t)}{2}} \epsilon(t), \quad (4.4)$$

para  $t = 1, \dots, N$ .

El siguiente lema será de utilidad en la demostración de algunos resultados.

**Lema 4.1.1.** *Sea  $Y$  una variable aleatoria con distribución  $N(0, \sigma^2)$ . Si se define  $Z$  variable aleatoria tal que  $Z = aY + \mu$ , entonces  $Z$  tiene distribución  $N(\mu, a^2\sigma^2)$ , con  $a, \mu \in \mathbb{R}$ .*

*Demostración.* En primer lugar, se obtendrá la esperanza y la varianza. Para la esperanza se tiene que:

$$\begin{aligned} \mathbb{E}[Z] &= \mathbb{E}[aZ + \mu] \\ &= a\mathbb{E}[Z] + \mu \\ &= \mu \end{aligned}$$

Para la varianza se tiene lo siguiente:

$$\begin{aligned} \text{Var}(Z) &= \text{Var}(aZ + \mu) \\ &= a^2 \text{Var}(Z) \\ &= a^2 \sigma^2 \end{aligned}$$

Por otro lado, sea  $M_Y(u)$  la función generadora de momentos de la variable aleatoria  $Y$ , dada por:

$$M_Y(u) = e^{\frac{\sigma^2 u^2}{2}}.$$

Por definición de  $Z$  se tiene que:

$$\begin{aligned} M_Z(u) &= e^{\mu u} M_Y(au) \\ &= e^{\mu u} e^{\frac{a^2 \sigma^2 u^2}{2}} \\ &= e^{\mu u + \frac{a^2 \sigma^2 u^2}{2}} \end{aligned}$$

Por lo anterior se concluye que  $Z$  tiene distribución  $N(\mu, a^2 \sigma^2)$   $\square$

**Proposición 4.1.2.** *Bajo los supuestos considerados hasta el momento, se cumple lo siguiente:*

- (a) *La variable aleatoria  $h(1)$  tiene una distribución  $N(\mu, \sigma_\eta^2)$ .*
- (b) *Dado  $h(t-1)$ , la variable aleatoria  $h(t)$  tiene una distribución  $N(\mu + \phi[h(t-1) - \mu], \sigma_\eta^2)$ , con  $t = 2, \dots, N$ .*
- (c)  $\mathbb{E}[Y(t)|h(t)] = 0$ .
- (d)  $\text{Var}(Y(t)|h(t)) = e^{h(t)}$ .
- (e) *Dado  $h(t)$ , la variable aleatoria  $Y(t)$  tiene distribución  $N(0, e^{h(t)})$ , con  $t = 1, \dots, N$ .*

*Demostración.* (a) Se sabe que  $\eta(t)$  tiene una distribución  $N(0, \sigma_\eta^2)$ . Como  $h(1) = \mu + \eta(t)$ , aplicando el lema 4.1.1 se tiene que  $h(1)$  tiene distribución  $N(\mu, \sigma_\eta^2)$ .

(b) Por hipótesis se tiene que  $h(t) = \mu + \phi[h(t-1) - \mu] + \eta(t)$ . Por el lema 4.1.1 se concluye que dada  $h(t-1)$ ,  $h(t)$  tiene una distribución  $N(\mu + \phi[h(t-1) - \mu], \sigma_\eta^2)$ .

(c) Se definió  $Y(t) = e^{\frac{h(t)}{2}} \epsilon(t)$ , donde  $\epsilon(t)$  tiene distribución  $N(0, 1)$ . Por lo tanto:

$$\mathbb{E}[Y(t)|h(t)] = \mathbb{E}[e^{\frac{h(t)}{2}} \epsilon(t)|h(t)] = e^{\frac{h(t)}{2}} \mathbb{E}[\epsilon(t)|h(t)] = 0.$$

(d) Por definición de  $Y(t)$  dada en (4.4) y por el inciso (c), se tiene que:

$$\begin{aligned} \text{Var}[Y(t)|h(t)] &= \mathbb{E}[Y(t)^2|h(t)] \\ &= \mathbb{E}[e^{h(t)} \epsilon(t)^2|h(t)] \\ &= e^{h(t)} \mathbb{E}[\epsilon(t)^2|h(t)] \\ &= e^{h(t)}. \end{aligned}$$

- (e) Se tiene que  $\epsilon(t)$  tiene una distribución  $N(0, 1)$ . Del lema 4.1.1 y con los resultados obtenidos en (c) y (d) se tiene que  $Y(t)$  tiene una distribución  $N(0, e^{h(t)})$ . □

Como consecuencia de la proposición 4.1.2, se tiene que la función de verosimilitud del modelo está dada por:

$$L(Y|h(t), \theta) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi e^{h(t)}}} \exp\left\{-\frac{Y^2(t)}{2e^{h(t)}}\right\}. \quad (4.5)$$

*Demostración.* Ver, por ejemplo, De Jesús Romo (2013). □

## 4.2. Modelos de Volatilidad Estocástica Bivariados

Sean  $N > 1$  un número entero fijo que representa el tamaño del conjunto de datos, y  $\mathbf{Z}(t) = (Z_1(t), Z_2(t))$ ,  $t = 1, \dots, N$ , la serie de tiempo bivariada que será objeto de estudio. Consideremos el vector de variables latentes  $h(t) = (h_1(t), h_2(t))$ ,  $t = 1, \dots, N$  con cada  $h_i(t)$  definida por un modelo AR(t) donde, para  $t = 1$ , se tiene:

$$\begin{cases} h_1(1) = \mu_1 + \eta_1(1), \\ h_2(1) = \mu_2 + \eta_2(1), \end{cases} \quad (4.6)$$

y para  $t = 2, \dots, N$ :

$$\begin{cases} h_1(t) = \mu_1 + \phi_{11}[h_1(t-1) - \mu_1] + \eta_1(t), \\ h_2(t) = \mu_2 + \phi_{22}[h_2(t-1) - \mu_2] + \eta_2(t), \end{cases} \quad (4.7)$$

donde  $0 < \phi_{11}, \phi_{22} < 1$  y  $\eta(t) = (\eta_1(t), \eta_2(t))$  tiene una distribución normal bivariada con vector de medias  $\mathbf{0} = (0, 0)$  y matriz de varianzas y covarianzas dada por:

$$\sigma_\eta^2 = \begin{pmatrix} \sigma_{\eta_1}^2 & 0 \\ 0 & \sigma_{\eta_2}^2 \end{pmatrix}.$$

Consideremos  $\mathbf{H}(t)$  con  $t = 1, \dots, N$ , una matriz diagonal dada por:

$$\mathbf{H}(t) = \begin{pmatrix} e^{\frac{h_1(t)}{2}} & 0 \\ 0 & e^{\frac{h_2(t)}{2}} \end{pmatrix}. \quad (4.8)$$

Si para  $\mathbf{v}$  un vector o matriz, se indica por  $\mathbf{v}'$  a su transpuesto, entonces sea  $\mathbf{Y}(t) = (Y_1(t), Y_2(t))'$ ,  $t = 1, \dots, N$  modelada por:

$$\mathbf{Y}(t) = \mathbf{H}(t)\epsilon(t),$$

donde  $\epsilon(t) = (\epsilon_1(t), \epsilon_2(t))'$  es el vector de errores, con distribución normal bivariada con vector de medias  $\mathbf{0} = (0, 0)$  y matriz de varianzas y covarianzas:

$$\Sigma_\epsilon = \begin{pmatrix} 1 & \rho_\epsilon \\ \rho_\epsilon & 1 \end{pmatrix}, \quad (4.9)$$

con  $-1 \leq \rho_\epsilon \leq 1$ . Entonces,  $\mathbf{Y}(t)$ ,  $t = 1, \dots, N$  está dada por:

$$\begin{cases} Y_1(t) = e^{\frac{h_1(t)}{2}} \epsilon_1(t), \\ Y_2(t) = e^{\frac{h_2(t)}{2}} \epsilon_2(t), \end{cases} \quad (4.10)$$

donde, en la aplicación,  $Y_1(t)$  y  $Y_2(t)$  serán los log-retornos de las observaciones  $Z_1(t)$  y  $Z_2(t)$  respectivamente, es decir:

$$Y_i(t) = \log \left( \frac{Z_i(t)}{Z_i(t-1)} \right), i = 1, 2, t = 2, 3, \dots, N.$$

Por definición,  $\mathbb{E}[\mathbf{Y}(t)] = \mathbf{0}$  y la matriz de varianzas y covarianzas será:

$$\Sigma_Y = \text{Var}(\mathbf{Y}(t)) = \mathbf{H}(t)' \Sigma_\epsilon \mathbf{H}(t) = \begin{pmatrix} e^{\frac{h_1(t)}{2}} & \rho_\epsilon e^{\frac{h_1(t)}{2}} e^{\frac{h_2(t)}{2}} \\ \rho_\epsilon e^{\frac{h_1(t)}{2}} e^{\frac{h_2(t)}{2}} & e^{\frac{h_2(t)}{2}} \end{pmatrix}. \quad (4.11)$$

Además  $\mathbf{Y}(t) = (Y_1(t), Y_2(t))'$  tiene distribución normal bivariada con función de densidad:

$$f(Y_1(t), Y_2(t) | h_1(t), h_2(t)) = \frac{1}{2\pi \sqrt{(1-\rho_\epsilon^2) e^{h_1(t)+h_2(t)}}} \exp \left\{ -\frac{1}{2(1-\rho_\epsilon^2)} \left[ \frac{Y_1^2(t)}{e^{h_1(t)}} + \frac{Y_2^2(t)}{e^{h_2(t)}} - \frac{2\rho_\epsilon Y_1(t)Y_2(t)}{e^{\frac{h_1(t)}{2}} e^{\frac{h_2(t)}{2}}} \right] \right\}.$$

Cada  $h_i(1)$ ,  $i = 1, 2$ , tiene distribución  $N(\mu_i, \sigma_{\eta_i}^2)$  y dada  $h_i(t-1)$ , la variable  $h_i(t)$  tiene distribución  $N(\mu_i + \phi_{ii}[h_i(t-1) - \mu_i], \sigma_{\eta_i}^2)$  para  $i = 1, 2$ ,  $t = 2, \dots, N$ .

La demostración de estos resultados puede consultarse en De Jesús Romo (2013).

### 4.3. Modelos utilizados

En esta sección se presentarán los casos particulares de los modelos que serán utilizados en la aplicación. Estos modelos fueron tomados de Zozolotto (2010).

#### 4.3.1. Modelo I

En este modelo, las coordenadas del vector de errores  $\epsilon_i(t)$ , con  $i = 1, 2$  son independientes, es decir,  $\rho_\epsilon = 0$ . Entonces,  $\mathbf{Y}(t)$  tiene distribución normal con vector de medias  $\mathbf{0} = (0, 0)$  y matriz de varianzas y covarianzas dada por:

$$\mathbf{H}(t) = \begin{pmatrix} e^{\frac{h_1(t)}{2}} & 0 \\ 0 & e^{\frac{h_2(t)}{2}} \end{pmatrix}.$$

Con  $h_1(t)$  y  $h_2(t)$  dados por (4.6) y (4.7).

#### 4.3.2. Modelo II

La covarianza  $\rho_\epsilon$  entre las coordenadas de  $\epsilon(t)$  es un valor desconocido que debe estimarse. Entonces  $\mathbf{Y}(t)$  tiene distribución normal con vector de medias  $\mathbf{0} = (0, 0)$  y matriz de varianzas y covarianzas dada por:

$$\Sigma_Y = Var(\mathbf{Y}(t)) = \mathbf{H}(t)' \Sigma_\epsilon \mathbf{H}(t) = \begin{pmatrix} e^{\frac{h_1(t)}{2}} & \rho_\epsilon e^{\frac{h_1(t)}{2}} e^{\frac{h_2(t)}{2}} \\ \rho_\epsilon e^{\frac{h_1(t)}{2}} e^{\frac{h_2(t)}{2}} & e^{\frac{h_2(t)}{2}} \end{pmatrix}.$$

Con  $h_1(t)$  y  $h_2(t)$  dados por (4.6) y (4.7).

#### 4.3.3. Modelo III

En este modelo se toman las hipótesis del modelo II, donde la covarianza  $\rho_\epsilon$  entre las coordenadas de  $\epsilon(t)$  es un valor desconocido que debe estimarse, pero en este caso se define  $h_2(t)$ ,  $t = 1, \dots, N$ , suponiendo causalidad de Granger, es decir:

$$h_2(t) = \mu_2 + \phi_{21}[h_1(t-1) - \mu_1] + \phi_{22}[h_2(t-1) - \mu_2] + \eta_2(t), \quad (4.12)$$

donde  $t = 2, \dots, N$  y  $0 < \phi_{21}, \phi_{22} < 1$ . La definición de  $h_1(t)$  es la dada en (4.6) y (4.7).

## 4.4. Análisis Bayesiano

La estimación de los parámetros se realizará mediante análisis bayesiano, y se presenta en esta sección.

En el caso del modelo I el vector de parámetros a ser estimado es  $\theta_I = (\mu_1, \mu_2, \phi_{11}, \phi_{22}, \sigma_{\eta_1}^2, \sigma_{\eta_2}^2)$ .

Cuando tomamos en cuenta el modelo II, el vector de parámetros es  $\theta_{II} = (\theta_I, \rho_\epsilon)$ . Finalmente, en el caso del modelo III se tiene  $\theta_{III} = (\theta_{II}, \phi_{21})$ .

Se busca determinar la densidad conjunta a posteriori de  $\theta = \theta_I, \theta_{II}, \theta_{III}$  y  $\mathbf{h}(t) = (h_1(t), h_2(t)), t = 1, 2, \dots, N$ , dada  $\mathbf{Y}(t) = (Y_1(t), Y_2(t))$ . Para esto es necesario definir en primer lugar cada una de las funciones que componen la densidad a posteriori del modelo. Sea  $\varphi = (\theta, \mathbf{h})$ , donde, dependiendo del caso,  $\theta = \theta_I, \theta_{II}, \theta_{III}$  y  $\mathbf{h} = (\mathbf{h}(1), \dots, \mathbf{h}(N))$ . De esta forma:

- La función de verosimilitud del modelo,  $L(D|\varphi)$ , está dada por:

$$L(D|\varphi) \propto \exp\left(-\frac{1}{2} \sum_{t=1}^N \sum_{j=1}^2 [h_j(t) + e^{-h_j(t)} [Y_j(t)]^2]\right). \quad (4.13)$$

- La densidad condicional de las variables latentes dados los parámetros del modelo, indicada por  $f(\mathbf{h}|\theta)$ , es:

$$f(\mathbf{h}|\theta) \propto \left[ \prod_{j=1}^2 (\sigma_{\eta_j}^2)^{-\frac{1}{2}} \exp\left(-\frac{1}{2\sigma_{\eta_j}^2} [h_j(1) - \mu_j]^2\right) \right] \prod_{t=2}^N \left[ \prod_{j=1}^2 (\sigma_{\eta_j}^2)^{-\frac{1}{2}} \exp\left(-\frac{1}{2\sigma_{\eta_j}^2} [h_j(t) - \mu_j - \phi_j(h_j(t-1)) - \mu_j]^2\right) \right]. \quad (4.14)$$

- La densidad a priori conjunta de los parámetros,  $P(\theta)$ , está dada por:

$$P(\theta) = \prod_{j=1}^2 P(\mu_j) P(\phi_j) P(\sigma_{\eta_j}^2). \quad (4.15)$$

En el siguiente capítulo se darán explícitamente las distribuciones a priori de los parámetros.

**Proposición 4.4.1.** *La densidad a posteriori del vector  $\mathbf{Y}$ , indicada por*

$p(\varphi|D)$ , está dada por: .

$$\begin{aligned}
p(\varphi|D) &\propto L(D|\varphi)f(\mathbf{h}|\theta)P(\theta) \\
&\propto \exp\left(-\frac{1}{2}\sum_{t=1}^N\sum_{j=1}^2[h_j(t)+e^{-h_j(t)}[Y_j(t)]^2]\right) \\
&\quad \times \left[\prod_{j=1}^2(\sigma_{\eta_j}^2)^{-\frac{1}{2}}\exp\left(-\frac{1}{2\sigma_{\eta_j}^2}[h_j(1)-\mu_j]^2\right)\right] \\
&\quad \times \prod_{t=2}^N \left[\prod_{j=1}^2(\sigma_{\eta_j}^2)^{-\frac{1}{2}}\exp\left(-\frac{1}{2\sigma_{\eta_j}^2}[h_j(t)-\mu_j-\phi_j(h_j(t-1))-\mu_j]^2\right)\right] \\
&\quad \times \prod_{j=1}^2 P(\mu_j)P(\phi_j)P(\sigma_{\eta_j}^2).
\end{aligned} \tag{4.16}$$

*Demostración.* Ver, por ejemplo, De Jesus Romo (2013).  $\square$

**Observación 4.** Una versión más elaborada del modelo Bayesiano utilizado en este trabajo, es considerar un nivel adicional en el modelo jerárquico, por ejemplo, aleatorizando tanto las varianzas como las medias de las  $\mu_i$ , así como aleatorizar los parámetros de las distribuciones de las otras variables. Sin embargo, se ha decidido, por simplicidad, no utilizar este nivel jerárquico. Este tipo de modificación sería incorporada en una etapa posterior donde se pretende preparar un artículo comparando los diversos modelos y formulaciones aplicados a los conjuntos de datos considerados en este trabajo.

## Capítulo 5

# Aplicación

### 5.1. Introducción

En este capítulo se utilizan los modelos de volatilidad estocástica presentados en el capítulo 4 para estimar la volatilidad entre dos series de datos: el tipo de cambio peso-dólar y las reservas internacionales de México. Los parámetros presentados en los modelos son estimados utilizando el punto de vista bayesiano, a través de los algoritmos MCMC. La tarea es facilitada al utilizar el programa *Winbugs* (que puede ser descargado de forma gratuita de la página <https://www.mrc-bsu.cam.ac.uk/software/bugs/the-bugs-project-winbugs/>), dado que sólo se debe especificar la función de verosimilitud del modelo y las distribuciones a priori de los parámetros.

### 5.2. Descripción de los datos

Se estudiará el comportamiento de las reservas internacionales de México en relación a la variación del tipo de cambio. Los datos utilizados, correspondientes a las reservas internacionales y al tipo de cambio peso-dólar interbancario, se obtuvieron de la base de datos proporcionada por el Instituto Nacional de Estadística y Geografía (INEGI) a través del Banco de Información Económica (<https://www.inegi.org.mx/sistemas/bie/>) y del Sistema de Información Económica del Banco de México (<https://www.banxico.org.mx/SieAPIRest/service/v1/?jsessionid=e26d01353f6720f5df8fe20add16>), y corresponden a datos de periodicidad mensual comprendidos entre enero de 1986 y hasta diciembre de 2018, expresados en unidades de Peso Mexicano. Por lo tanto, tenemos dos series de datos, con 384 elementos cada una.

En el cuadro 5.1 se presenta un resumen de las estadísticas básicas de los datos utilizados.

Estadísticas de los datos		
Estadística	Tipo de cambio	Reservas internacionales
Media	9.6256	951,485,583,044
Desviación estándar	4.988362	1,089,217,087,493
Intervalo 95 % de credibilidad	(9.125104; 10.126129)	(842,301,970,425; 1,060,877,471,705)

Cuadro 5.1: Resumen estadístico

El tipo de cambio es una referencia utilizada en el mercado cambiario para conocer el número de unidades de moneda nacional que deben pagarse para obtener una moneda extranjera.

La figura 5.1 representa el comportamiento del tipo de cambio peso-dólar, durante el periodo enero 1986 a diciembre de 2018.

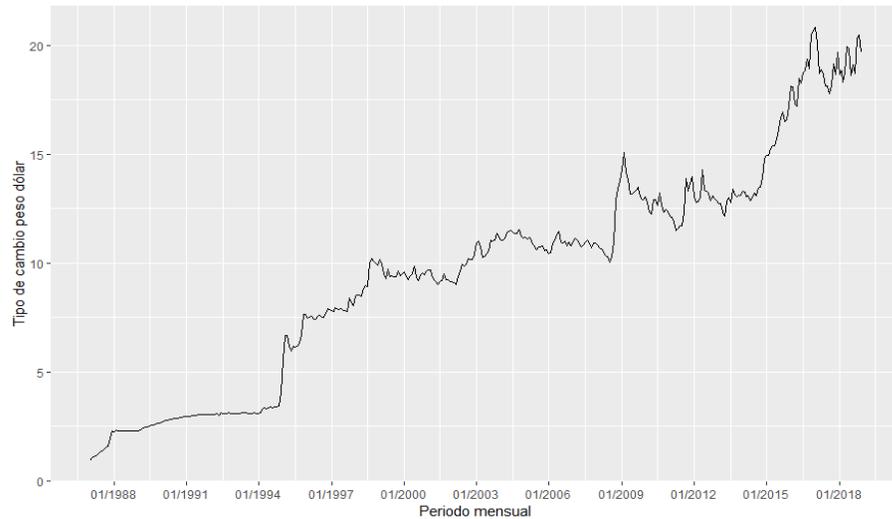


Figura 5.1: Tipo de cambio peso-dólar durante el periodo enero 1986 - diciembre 2018

De la figura 5.1 se puede notar que hasta marzo de 1994 el precio del dólar presentaba un aumento constante pero controlado. Sin embargo, a finales de 1994, tras la devaluación del peso, el precio del dólar tuvo un gran incremento. Desde entonces ha habido una serie de incrementos y decrementos, algunos de forma controlada y otros con grandes saltos. Otra gran devaluación ocurrió a finales del año 2008 con la crisis financiera global. A

finales de 2014 se presenta una tasa de crecimiento bastante alta hasta finales de 2017, donde se estabilizó nuevamente en periodos de crecimiento y decrecimiento intercalados.

Por otro lado, las reservas internacionales son activos financieros que el banco central invierte en el exterior. Su principal característica es la liquidez, es decir, la facilidad que poseen para saldar las obligaciones de pago de un país. Contribuyen a la estabilidad de precios cuando se presenta una disminución de los flujos comerciales o de capital de la balanza de pagos, causados por desequilibrios macroeconómicos o financieros. En México las reservas internacionales se integran por activos financieros en las divisas de mayor importancia en los mercados internacionales, en este caso los dólares.

La figura 5.2 representa el comportamiento de las reservas internacionales expresadas en millones de pesos durante el periodo enero 1986 a diciembre 2018.

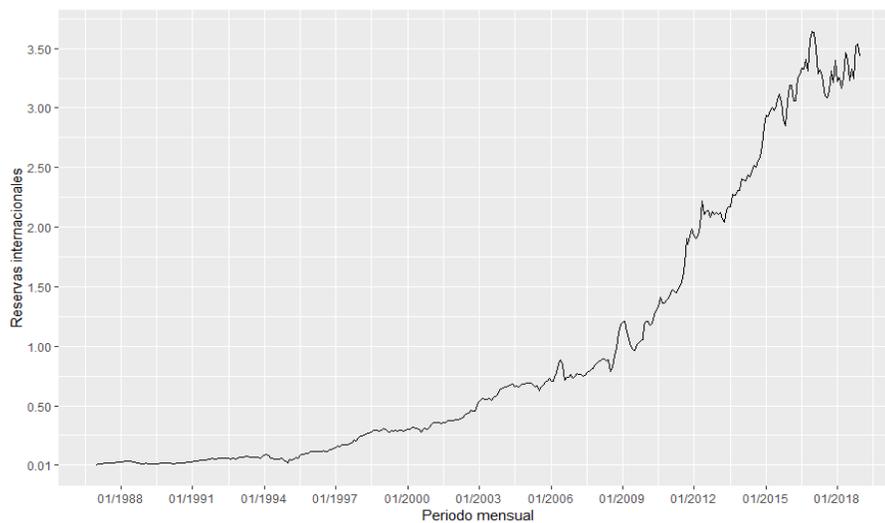


Figura 5.2: Reservas internacionales de México durante el periodo enero 1986 - diciembre 2018

Como se puede observar de la figura 5.2, las reservas internacionales han tenido un aumento persistente a través del tiempo, y sólo en los últimos años han presentado un comportamiento irregular. La tasa de crecimiento se asemeja a una función exponencial con un rápido aumento durante el periodo entre 2011 y 2017.

**Observación 5.** *En ciertos periodos ambas series presentan crecimientos*

abruptos, tal es el caso del año 2009 y el año 2017.

En las figuras 5.3 y 5.4 se muestran las gráficas de los log-retornos  $Y_i(t)$ ,  $i = 1, 2, t = 2, \dots, N$ , centrados alrededor de la media, para los datos de tipo de cambio y reservas internacionales.

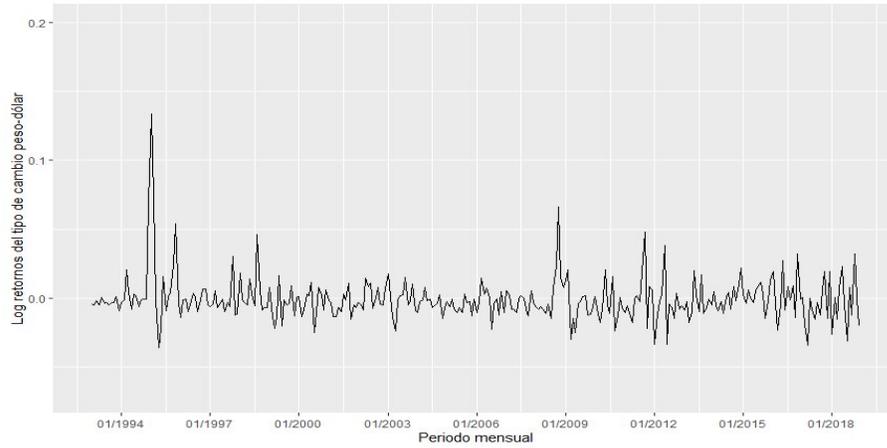


Figura 5.3: Log-retornos del tipo de cambio peso-dólar durante el periodo enero 1986 - diciembre 2018

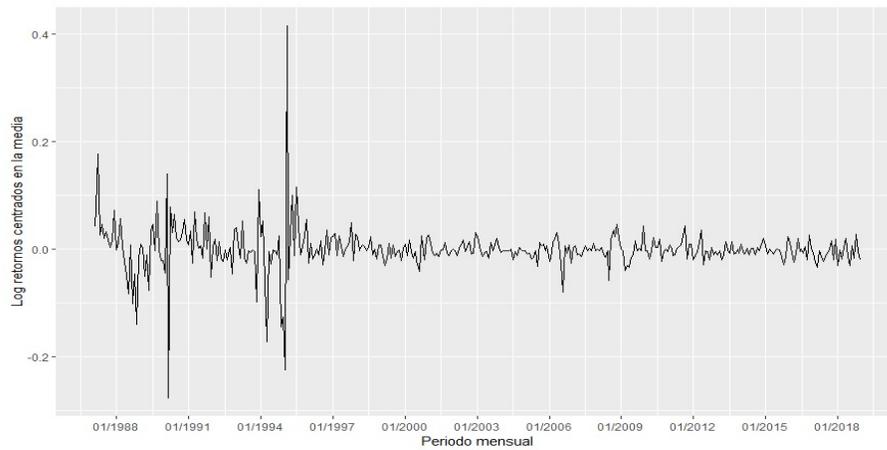


Figura 5.4: Log-retornos de las reservas internacionales de México durante el periodo enero 1986 - diciembre 2018

Al comparar las figuras 5.3 y 5.4 con sus correspondientes figuras 5.1

y 5.2 se puede observar que las variaciones más acentuadas se presentan alrededor de los tiempos donde ocurrieron cambios abruptos en el tipo de cambio y en el volumen de las reservas internacionales.

Como se está utilizando el software Winbugs para estimar los parámetros de los modelos es necesario especificar las distribuciones a priori de estos parámetros. Esto se presenta en la siguiente sección.

### 5.3. Distribuciones a priori

Dado que el punto de vista bayesiano será utilizado para estimar los parámetros de los modelos de volatilidad aplicados a los datos de tipo de cambio y reservas internacionales, se deben especificar las distribuciones a priori para este caso específico.

A continuación se presentan los valores asignados a los hiperparámetros de las distribuciones a priori para cada uno de los modelos desarrollados.

En lo que sigue se utilizará la siguiente notación. Se denotarán las distribuciones normales con media  $\mu$  y varianza  $\sigma^2$  como  $N(\mu, \sigma^2)$ , las distribuciones beta con parámetros  $\alpha, \beta$  como  $Beta(\alpha, \beta)$ , las distribuciones gamma con parámetros  $c, d$  como  $Gamma(c, d)$  y las distribuciones uniformes con parámetros  $a, b$  como  $U(a, b)$ .

#### 5.3.1. Modelo I

En este modelo el vector de parámetros a ser estimado será  $\theta_I = (\phi_{11}, \phi_{22}, \sigma_{\eta_1}^2, \sigma_{\eta_2}^2, \mu_1, \mu_2)$ . Se supusieron las siguientes distribuciones a priori:

- $\mu_i$  tiene distribución  $N(0, 0.01)$ ,  $i = 1, 2$ .
- $\phi_{ii}$  tiene distribución  $Beta(1, 1)$ ,  $i = 1, 2$ .
- $\sigma_{\eta_i}^2$  tiene distribución  $Gamma(1, 1)$ ,  $i = 1, 2$ .

#### 5.3.2. Modelo II

En este caso, el vector de parámetros será  $\theta_{II} = (\phi_{11}, \phi_{22}, \tau_{\eta_1}^2, \tau_{\eta_2}^2, \mu_1, \mu_2, \rho_\epsilon)$  y se tomaron las siguientes distribuciones a priori:

- $\mu_i$  tiene distribución  $N(-3, 0.01)$ ,  $i = 1, 2$ .
- $\phi_{ii}$  tiene distribución  $Beta(1, 1)$ ,  $i = 1, 2$ .

- Se define el parámetro  $\tau_{\eta_i} = \frac{1}{\sigma_{\eta_i}^2}$ ,  $i = 1, 2$ . En este caso  $\tau_{\eta_1} = \frac{1}{\sigma_{\eta_1}^2}$  tiene distribución  $Gamma(5, 1)$ , y  $\tau_{\eta_2} = \frac{1}{\sigma_{\eta_2}^2}$  tiene distribución  $Gamma(4, 1)$ .
- $\rho_\epsilon$  tiene distribución  $U(-1, 1)$ .

### 5.3.3. Modelo III

En este modelo el vector de parámetros a ser estimados será  $\theta_{III} = (\phi_{11}, \phi_{21}, \phi_{22}, \tau_{\eta_1}^2, \tau_{\eta_2}^2, \mu_1, \mu_2, \rho_\epsilon)$ . Se supusieron las distribuciones a priori siguientes:

- $\mu_i$  tiene distribución  $N(-3, 0.01)$ ,  $i = 1, 2$ .
- $\phi_{ii}, \phi_{21}$  tiene distribución  $Beta(1, 1)$ ,  $i = 1, 2$ .
- $\tau_{\eta_1} = \frac{1}{\sigma_{\eta_1}^2}$  tiene distribución  $Gamma(9, 1)$ , y  $\tau_{\eta_2} = \frac{1}{\sigma_{\eta_2}^2}$  tiene distribución  $Gamma(7, 1)$ .
- $\rho_\epsilon$  tiene distribución  $U(-1, 1)$ .

**Observación 6.** *Note que en los tres modelos las varianzas de las  $\mu'_i$ s son las mismas. Se decidió usar la misma varianza por conveniencia. Dado que las medias de las  $\mu_i$  son diferentes, los resultados no se deben ver afectados. Las varianzas fueron tomadas muy pequeñas para acelerar la convergencia de algoritmos. Al utilizar el programa Winbugs esta es una desventaja que se puede presentar, pues en pruebas preliminares no se logró llegar a la convergencia en un tiempo razonable. Cabe destacar que aún con la elección actual de varianzas el número de pasos hasta la convergencia fue bastante alto.*

## 5.4. Estimación de los parámetros

Para estimar los parámetros del modelo I fueron utilizadas 5 cadenas de las cuáles se tomó una muestra de tamaño 100,000, obtenida después de un periodo de calentamiento de 600,000 iteraciones con valores tomados a cada 10 pasos.

Para estimar los parámetros del modelo II fueron utilizadas 5 cadenas de las cuáles se tomó una muestra de tamaño 75,000, obtenida después de un periodo de calentamiento de 650,000 iteraciones con valores tomados a cada 10 pasos.

Finalmente, para estimar los parámetros del modelo III fueron utilizadas 5 cadenas de las cuáles se tomó una muestra de tamaño 100,000, obtenida después de un periodo de calentamiento de 700,000 iteraciones con valores tomados a cada 10 pasos.

En el cuadro 5.2 se presentan algunas estadísticas de interés así como los valores DIC y errores Monte Carlo para cada parámetro de cada uno de los modelos implementados.

Modelo	Parámetro	Media	Desv. Est.	Int. de conf.	MC Error	DIC
Modelo I	$\mu_1$	-5.502	0.7587	(-6.455; -3.381)	0.00825	-3016.64
	$\mu_2$	-7.658	0.3991	(-8.359; -6.794)	0.001804	
	$\phi_1$	0.9128	0.04676	(0.8112; 0.9903)	0.000628	
	$\phi_2$	0.8574	0.04142	(0.7703; 0.9333)	0.000294	
	$\sigma_{\eta_1}$	2.509	0.9467	(1.186; 4.85)	0.01556	
	$\sigma_{\eta_2}$	1.2	0.3079	(0.7328; 1.924)	0.003603	
Modelo II	$\mu_1$	-4.872	0.8158	(-6.182 ; -3.05)	0.007292	-3153
	$\mu_2$	-7.534	0.3748	(-8.179; -6.708)	0.002438	
	$\phi_1$	0.9663	0.02191	(0.9148; 0.997)	0.0002405	
	$\phi_2$	0.8767	0.03996	(0.7922; 0.9488)	0.000699	
	$\tau_{\eta_1}$	4.242	1.248	(2.264; 7.102)	0.01949	
	$\tau_{\eta_2}$	1.97	0.5564	(1.131; 3.291)	0.008223	
	$\rho_\epsilon$	-0.6388	0.03787	(-0.7089; -0.5604)	0.0003191	
Modelo III	$\mu_1$	-4.23	0.8706	(-5.781 ; -2.453)	0.007228	-3140.96
	$\mu_2$	6.974	0.5982	(-7.89 ; -5.56)	0.004499	
	$\phi_{11}$	0.9817	0.01423	(0.9453 ; 0.9989)	0.0001166	
	$\phi_{21}$	0.02719	0.01679	(0.001848; 0.06484)	0.00007592	
	$\phi_{22}$	0.8977	0.03627	(0.8188; 0.9609)	0.0003291	
	$\tau_{\eta_1}$	5.91	1.582	(3.342; 9.474)	0.02136	
	$\tau_{\eta_2}$	2.573	0.7417	(1.43; 4.314)	0.01038	
	$\rho_\epsilon$	0.6374	0.037777	(0.5596; 0.7071)	0.0003123	

Cuadro 5.2: Medias, desviaciones estándar, intervalos al 95 % de credibilidad, errores Monte Carlo y DIC de los parámetros estimados para cada modelo.

A partir de los resultados presentados, el modelo del que se obtuvo el menor DIC es el modelo II, por lo que este será el modelo seleccionado.

## 5.5. Resultados obtenidos

En las figuras 5.5 y 5.6 se presentan las gráficas de las volatilidades estimadas utilizando el modelo seleccionado, es decir, el modelo II.

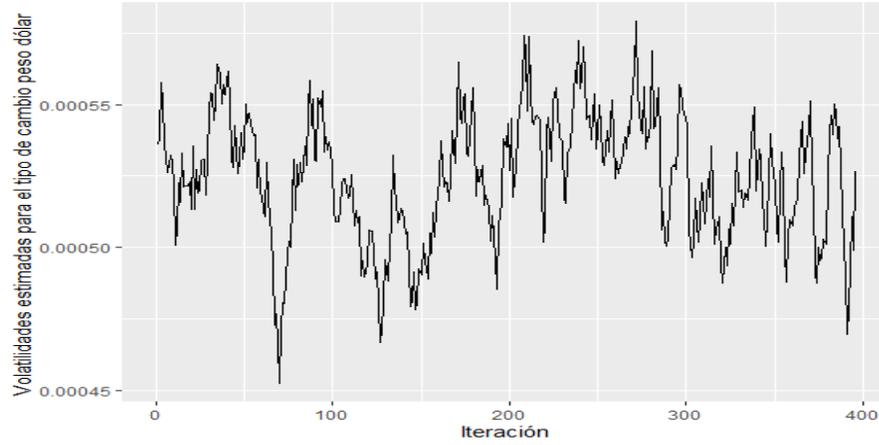


Figura 5.5: Volatilidades estimadas para el tipo de cambio peso-dólar

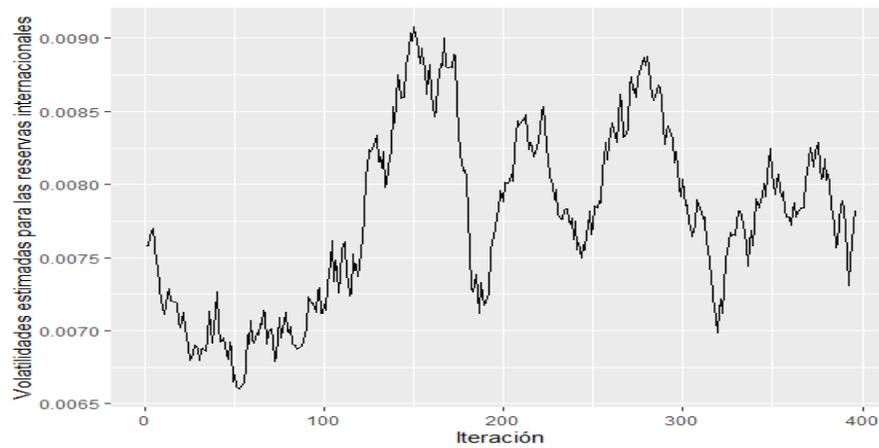


Figura 5.6: Volatilidades estimadas para las reservas internacionales de México

Note que por (4.8) la volatilidad de la  $i$ -ésima series es  $\sigma_i^2(t) = e^{h_i(t)}$ . Adicionalmente, se tiene por el cuadro 5.2 que el modelo seleccionado es el

modelo II. De esta forma se observa que cada  $h_i(t)$ ,  $i = 1, 2$ , está dada por las ecuaciones (4.7).

Las volatilidades fueron estimadas a partir del siguiente algoritmo:

### Algoritmo I

1. Generar los valores para  $h_i(1)$  a partir de sus distribuciones normales  $N(\mu_i, \sigma_i^2)$ , para  $i = 1, 2$ .
2. Para  $t = 2, \dots, N$ , dado  $h_i(t-1)$ ,  $i = 1, 2$ , generar  $h_i(t)$  utilizando su distribución normal  $N(\mu_i + \phi_{ii}[h_i(t-1) - \mu_i], \sigma_{\eta_i}^2)$ .
3. Repetir el procedimiento desde el paso 1 un número  $M$  suficientemente grande de veces.
4. Para cada  $t = 1, \dots, N$  obtener  $\hat{h}_i(t)$  usando:

$$\hat{h}_i(t) = \frac{1}{M} \sum_{k=1}^M h_i^{(k)}(t),$$

donde  $h_i^{(k)}(t)$  es el valor de  $h_i(t)$  generado en la  $k$ -ésima iteración del algoritmo.

5. Utilizando (4.8) obtener  $\sigma_i^2(t)$  con  $\hat{h}_i(t)$ ,  $t = 1, 2, \dots, N$ ,  $i = 1, 2$ .

**Observación 7.** Para las volatilidades estimadas presentadas en las figuras 5.5 y 5.6 se realizaron  $M = 1000$  simulaciones para cada  $t$ .

De la misma forma que se tienen las volatilidades estimadas, es posible obtener los log-retornos estimados y predecir los futuros. Para obtener los log-retornos estimados  $\hat{Y}(t)$ ,  $t = 2, \dots, N$ , basta hacer lo siguiente.

Por la ecuación (4.10) se tiene que:

$$Y_i(t) = e^{\frac{h_i(t)}{2}} \epsilon_i(t),$$

con  $i = 1, 2$ ,  $t = 2, \dots, N$ . Entonces para cada  $\hat{h}_i(t)$  generada utilizando el algoritmo I hacer lo siguiente:

### Algoritmo II

1. Generar  $(\epsilon_1(t), \epsilon_2(t))$  utilizando la distribución normal bivariada con media  $\mathbf{0} = (0, 0)$  y matriz de varianzas y covarianzas dada por la ecuación (4.9), donde  $\rho_\epsilon$  es el valor estimado en el cuadro 5.2.

2. Calcular  $Y_i^{(k)}(t)$  utilizando (4.10), con  $\hat{h}_i(t)$  en lugar de  $h_i(t)$  y utilizando  $(\epsilon_1^{(k)}(t), \epsilon_2^{(k)}(t))$ .
3. Repetir el procedimiento a partir del paso 1 un número  $M$  suficientemente grande de veces.
4. Calcular:

$$\hat{Y}_i(t) = \frac{1}{M} \sum_{k=1}^M Y_i^{(k)}(t),$$

con  $i = 1, 2$ ,  $t = 2, \dots, N$ , donde  $Y_i^{(k)}(t)$  es el  $k$ -ésimo valor obtenido usando el paso 2.

En las figuras siguientes se presentan los log-retornos estimados para ambas series de datos.

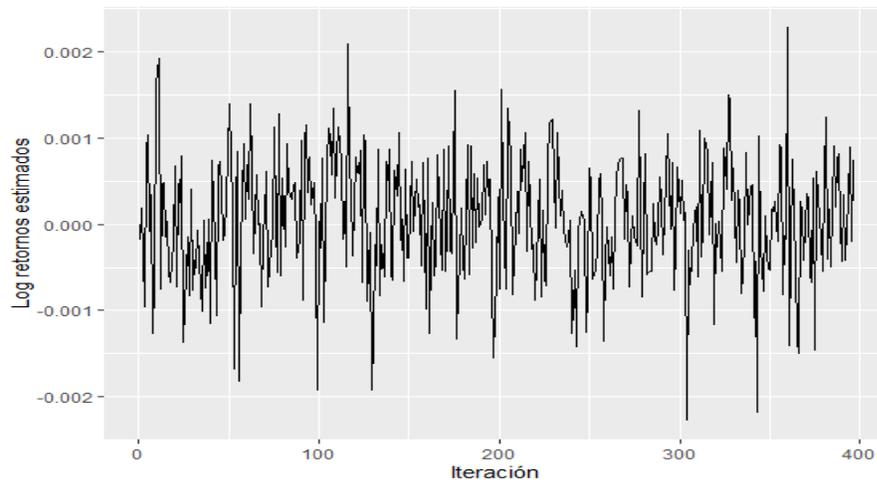


Figura 5.7: Log-retornos estimados para el tipo de cambio peso-dólar

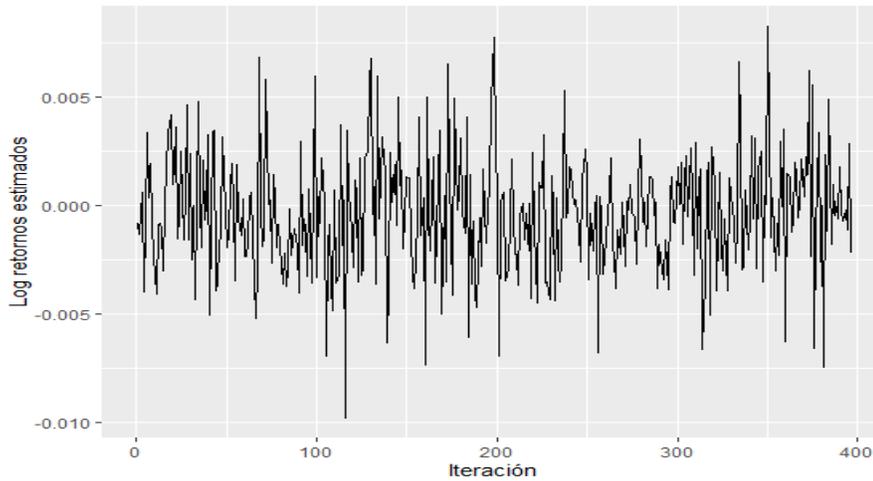


Figura 5.8: Log-retornos estimados para las reservas internacionales de México

## 5.6. Discusión

Como se especificó en la sección 4.3, se supuso a priori que el vector de  $\nu(t) = (\nu_1(t), \nu_2(t))$  era  $\mathbf{0} = (0, 0)$ . A partir del cuadro 5.2 se puede observar que el vector medias a posteriori, es decir, después de incorporar la información proporcionada por los datos observados, fue  $\mu = (-4, 87, -7, 53)$ . Adicionalmente, el valor estimado de  $\rho_\epsilon$  es  $\rho_\epsilon = -0,6388$ , lo que sugiere que existe una correlación negativa entre los log-retornos de la serie de tipo de cambio y la de reservas internacionales.

Las intervenciones en el mercado cambiario han sido una herramienta que las autoridades económicas de varios países han implementado en los últimos años, con el objetivo de procurar su estabilidad cambiaria y reducir la incertidumbre de sus monedas en épocas de inestabilidad.

Reducir el desalineamiento cambiario, limitar la volatilidad del tipo de cambio, administrar las reservas internacionales, y/o garantizar la liquidez en el mercado cambiario en una situación de escasez de moneda extranjera, son algunos de los objetivos que los bancos centrales plantean públicamente al intervenir en el mercado cambiario.

Con base en los resultados obtenidos usando los modelos considerados en este trabajo se puede concluir que, efectivamente, el comportamiento de las reservas internacionales en relación al tipo de cambio presentan una relación inversa de manera creciente.

Es importante notar que las reservas internacionales dependen del tipo de cambio y no al contrario. Cuando el precio del dólar aumenta, el banco central (Banco de México) realiza la venta de dólares con la intención de desacelerar este aumento. Sin embargo, cuando el precio del dólar baja, la moneda mexicana se ve beneficiada y no siempre se realizará una compra de dólares. Dicho comportamiento fue observado inicialmente en los datos, y a partir de los resultados obtenidos en la aplicación de los modelos de volatilidad estocástica se puede concluir que es correcta la interpretación inicial.

Para verificar la efectividad del modelo, se realizó una predicción para los datos y los log-retornos estimados correspondientes al año 2019, esto se puede realizar de la siguiente forma.

Por la ecuación (4.10) tenemos que:

$$Z_i(t) = e^{Y_i(t)} Z_i(t-1), \quad (5.1)$$

donde  $i = 1, 2$ , que es equivalente a:

$$Z_i(t+1) = e^{Y_i(t+1)} Z_i(t). \quad (5.2)$$

Tome  $t = N$  y sea  $t+1 = N+1$ , entonces proceder utilizando el siguiente algoritmo:

### Algoritmo III

1. Para cada  $h_i^{(k)}(t)$  generado usando el algoritmo I genere  $h_i^{(k)}(t+1)$  usando el paso 2 siguiente, para  $k = 1, 2, \dots, 1000$ .
2. Obtener  $h_i(t+1)$  estimado como en el paso 5 del algoritmo I.
3. Para cada  $\hat{h}_i^{(k)}(t+1)$  estimado usando el algoritmo I obtenga  $\hat{Y}_i^{(k)}(t+1)$  a partir del paso 4 del algoritmo II.
4. Para cada  $\hat{Y}_i(t+1)$  obtenido a partir del algoritmo II calcule  $Z_i(t+1)$  usando (5.2).
5. Incrementar el tiempo  $t$  a  $(t+1)$  y repetir el procedimiento a partir del paso 1 para el número de  $t$  que sean necesarios (en este caso para  $t = N+1, N+2, \dots, N+12$ ).

En los cuadros siguientes se presentan los resultados obtenidos para la predicción de los datos durante el año 2019, y se realiza la comparación

con los datos reales, de forma que se pueda apreciar la precisión del modelo seleccionado.

Comparativo para el tipo de cambio

Mes	Predicción	Real
ene-19	19.6827	19.0295
feb-19	19.6866	19.1925
mar-19	19.7015	19.3056
abr-19	19.7175	18.9837
may-19	19.7168	19.6865
jun-19	19.7082	19.1556
jul-19	19.715	18.9956
ago-19	19.7068	20.0821
sep-19	19.7126	19.7250
oct-19	19.7302	19.1846
nov-19	19.7263	19.5006
dic-19	19.7407	18.8555

Cuadro 5.3: Comparativo de los datos reales y las predicciones para el tipo de cambio peso-dólar

Comparativo para las reservas

Mes	Predicción	Real
ene-19	3,423,918,000,000	3,342,686,437,000
feb-19	3,427,235,000,000	3,384,761,320,000
mar-19	3,425,709,000,000	3,423,079,072,000
abr-19	3,425,599,000,000	3,359,856,402,000
may-19	3,423,603,000,000	3,493,553,057,000
jun-19	3,429,685,000,000	3,435,813,388,000
jul-19	3,428,868,000,000	3,620,094,516,000
ago-19	3,426,421,000,000	3,556,320,913,000
sep-19	3,425,773,000,000	3,469,444,774,000
oct-19	3,421,888,000,000	3,469,444,774,000
nov-19	3,431,779,000,000	3,528,003,760,000
dic-19	3,424,285,000,000	3,412,103,062,000

Cuadro 5.4: Comparativo de los datos reales y las predicciones para las reservas internacionales



## Capítulo 6

# Conclusiones

En este trabajo se estudió el comportamiento de dos series de datos financieros, tipo de cambio peso-dólar y reservas internacionales de México, a través de sus volatilidades. Se modelaron conjuntamente sus comportamientos utilizando tres modelos bivariados de volatilidad estocástica, cada uno considerando supuestos diferentes.

Se utilizó la inferencia bayesiana y los métodos Monte Carlo vía cadenas de Markov para estimar los parámetros de estos modelos. Posteriormente se realizó la aplicación a las dos series de datos y se implementaron los tres modelos utilizando el programa Winbugs, con los códigos antes elaborados por Zozolotto (2010).

Se concluyó que el modelo que mejor describe el comportamiento de los datos es el modelo II. A partir de este modelo se observó que sí existe relación entre el tipo de cambio y las reservas internacionales en la economía mexicana. Es importante resaltar que el monto de reservas internacionales depende del nivel en el que se encuentra el tipo de cambio, y no necesariamente el tipo de cambio depende del monto de las reservas.

Utilizando el modelo seleccionado conjuntamente con sus parámetros estimados se pudo estimar la volatilidad de ambas series, así como predecir el comportamiento futuro a partir de los valores de la serie y sus log-retornos. En este caso se consideró un periodo futuro de 12 meses.

Los valores futuros del tipo de cambio peso-dólar y de las reservas internacionales mexicanas también fueron estimados y comparados con los valores reales para los 12 meses del 2019.

Partiendo de la presente tesis se considera ampliar este estudio, evaluando detalladamente la correlación entre los datos a periodos de tiempo determinados. También se plantea evaluar la eficiencia de los modelos de vo-

latilidad estocástica con el enfoque bayesiano y contrastarla con la eficiencia de los mismos modelos utilizando series de tiempo. Así mismo, se propone considerar el caso en el que  $h_1(t)$  está dada por:

$$h_1(t) = \mu_1 + \phi_{11}[h_1(t-1) - \mu_1] + \phi_{21}[h_2(t-1) - \mu_2]$$

para  $t = 1, 2, \dots, N$ .

La volatilidad en este modelo estará dada por:

$$\sigma_i^2(t) = e^{h_i(t) + \omega(t)},$$

con  $i = 1, 2$ ,  $t = 1, \dots, N$  y  $\omega(t)$  otra variable latente independiente e idénticamente distribuida con distribución normal  $N(0, \sigma_\omega^2)$ .

En la versión que se ha descrito arriba, así como las utilizadas en este trabajo se pretende incluir un mayor nivel jerárquico donde se considerarán otros casos en los que las varianzas de las distribuciones normales no sean variables con una distribución gamma o gamma inversa, y sus parámetros tengan comportamientos aleatorios.

# Apéndices



## A.1. Valores generados por el algoritmo MCMC

En este apéndice se presentan los gráficos con los valores de los parámetros generados por los algoritmos MCMC para todos los modelos utilizados en la presente tesis.

### A.1.1 Parámetros presentados en el modelo I

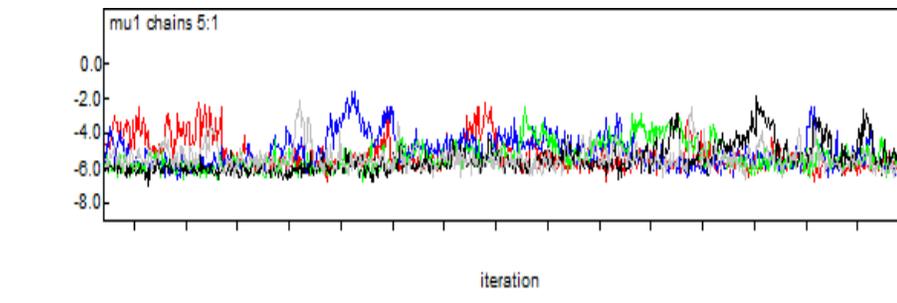


Figura 1: Parámetro  $\mu_1$

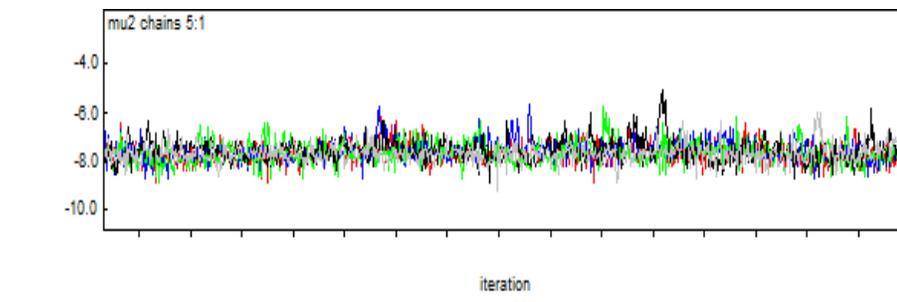
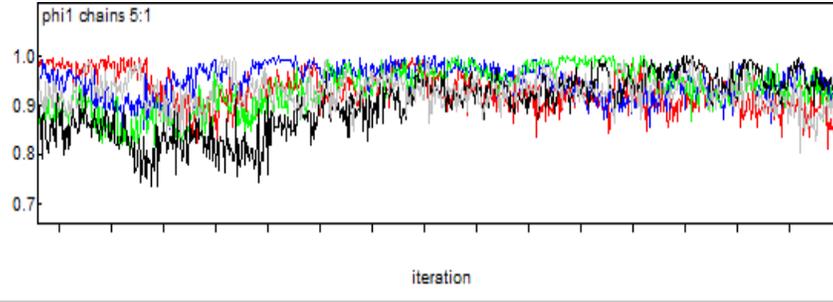
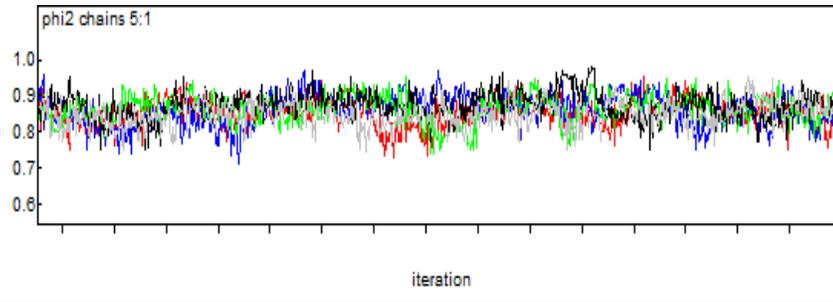
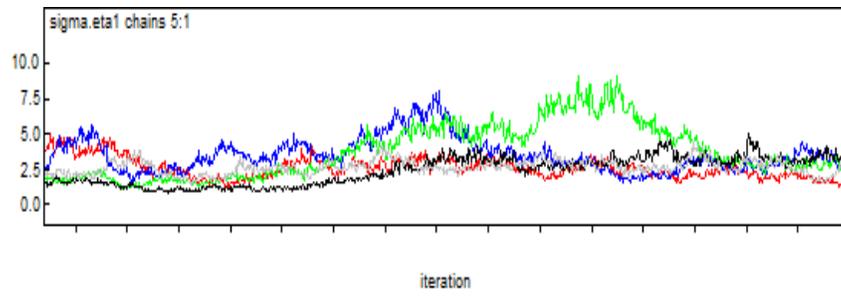
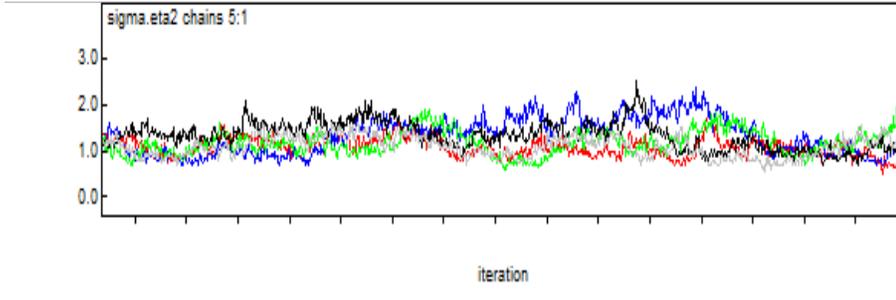
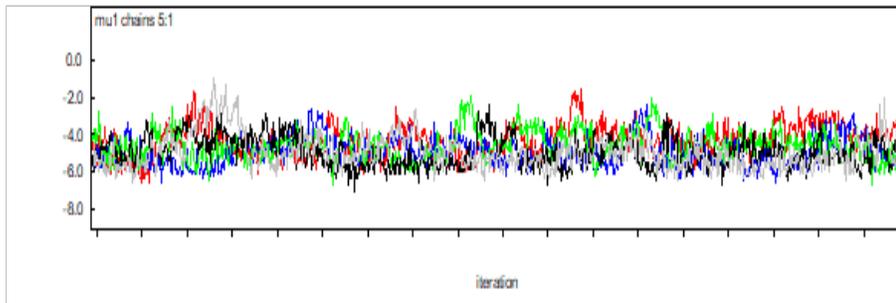
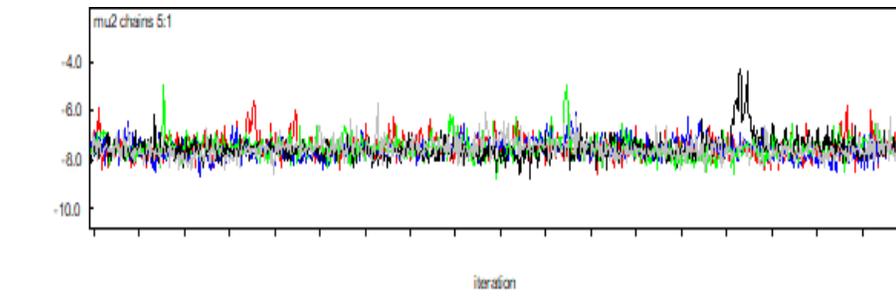


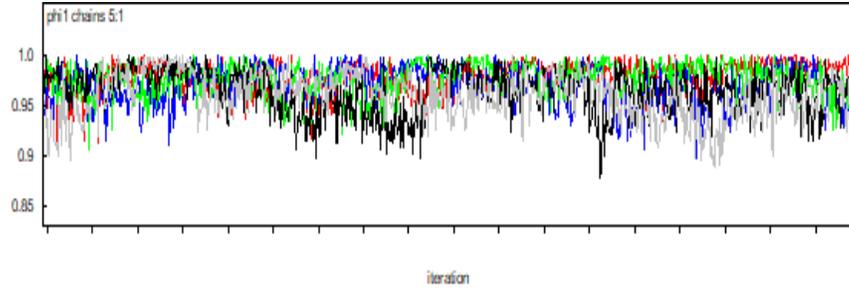
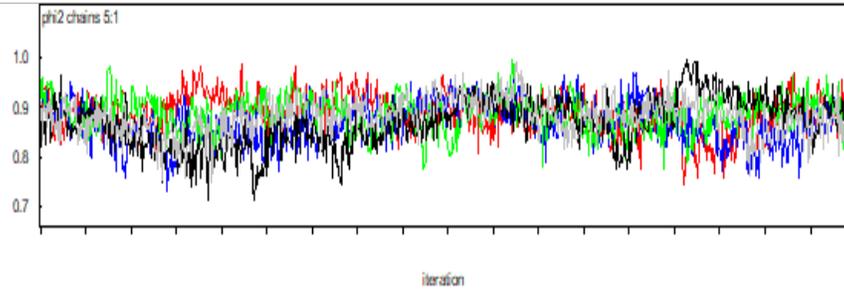
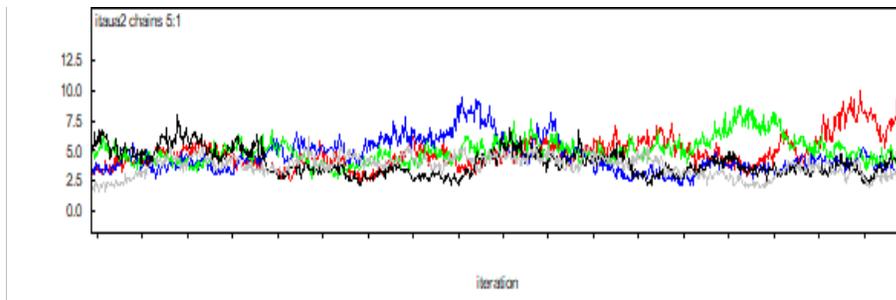
Figura 2: Parámetro  $\mu_2$

Figura 3: Parámetro  $\phi_{11}$ Figura 4: Parámetro  $\phi_{22}$ Figura 5: Parámetro  $\sigma_{\eta_1}$

Figura 6: Parámetro  $\sigma_{\eta_2}$ 

### A.1.2 Parámetros presentados en el Modelo II

Figura 7: Parámetro  $\mu_1$ Figura 8: Parámetro  $\mu_2$

Figura 9: Parámetro  $\phi_{11}$ Figura 10: Parámetro  $\phi_{22}$ Figura 11: Parámetro  $\tau_{\eta_1} = \frac{1}{\sigma_{\eta_1}}$

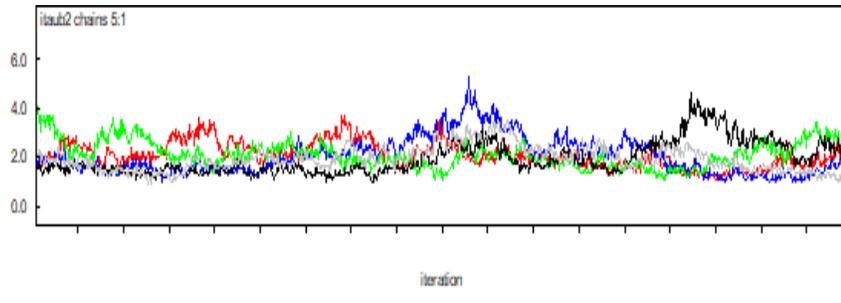


Figura 12: Parámetro  $\tau_{\eta_2} = \frac{1}{\sigma_{\eta_2}}$

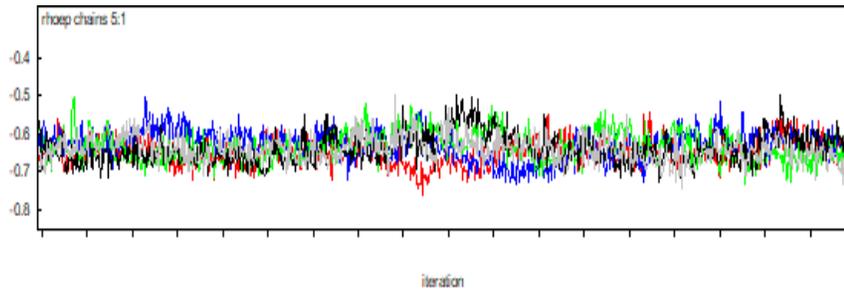


Figura 13: Parámetro  $\rho_\epsilon$

### A.1.3 Parámetros presentados en el Modelo III

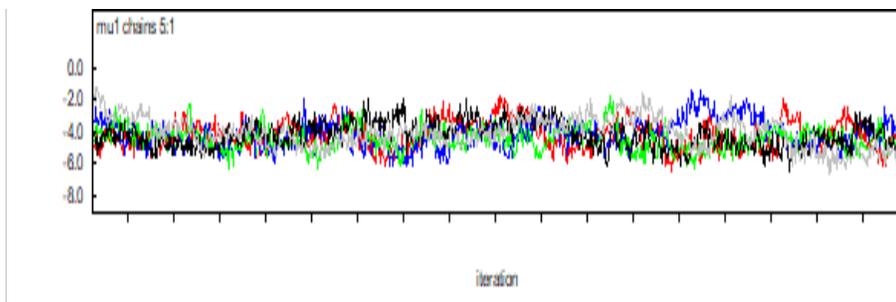
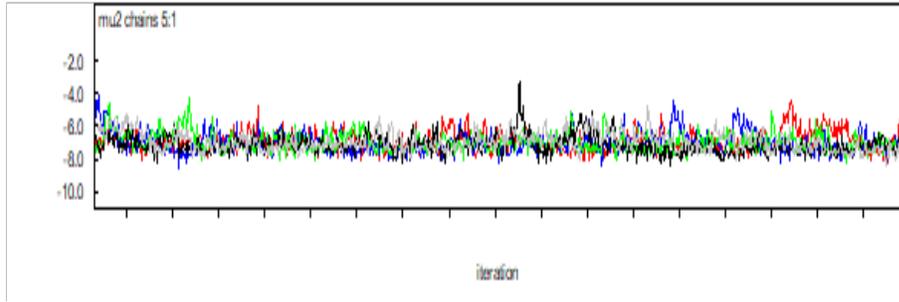
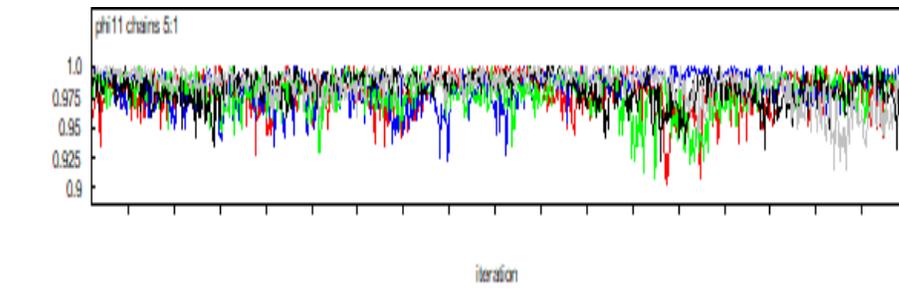
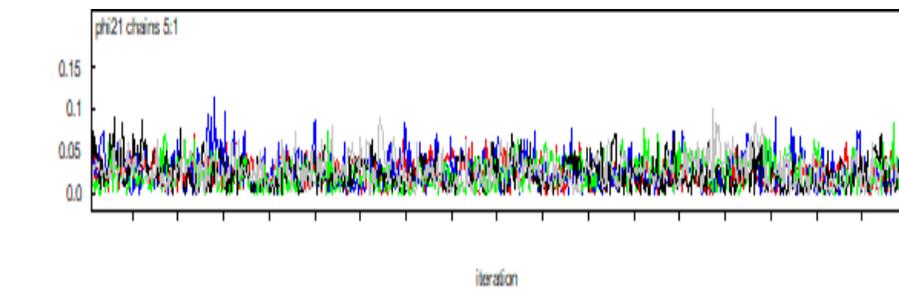


Figura 14: Parámetro  $\mu_1$

Figura 15: Parámetro  $\mu_2$ Figura 16: Parámetro  $\phi_{11}$ Figura 17: Parámetro  $\phi_{21}$

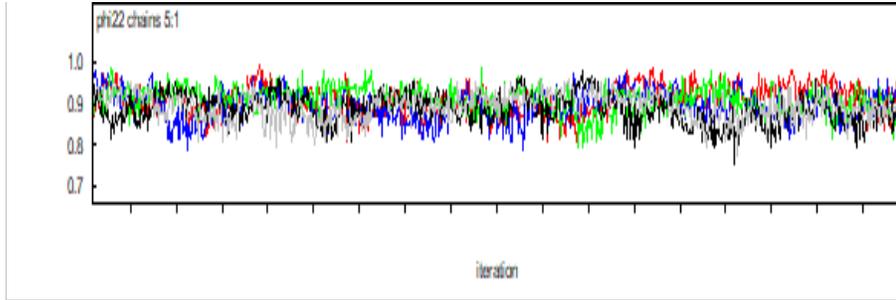


Figura 18: Parámetro  $\phi_{22}$

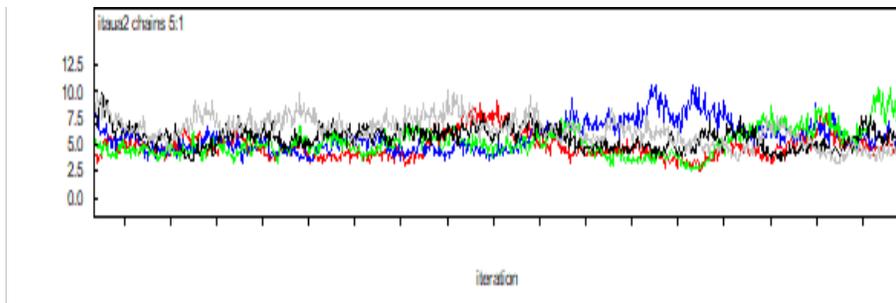


Figura 19: Parámetro  $\tau_{\eta_1} = \frac{1}{\sigma_{\eta_1}}$

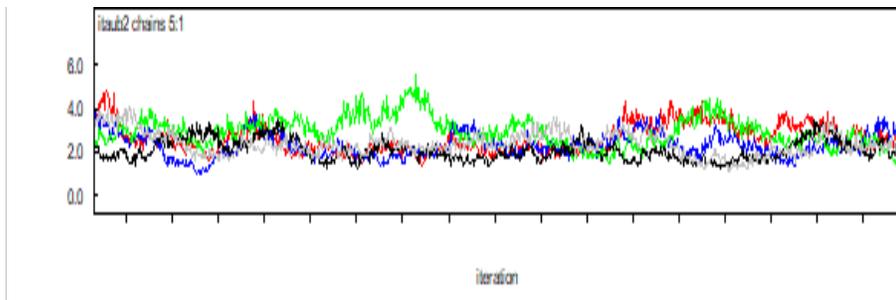


Figura 20: Parámetro  $\tau_{\eta_2} = \frac{1}{\sigma_{\eta_2}}$

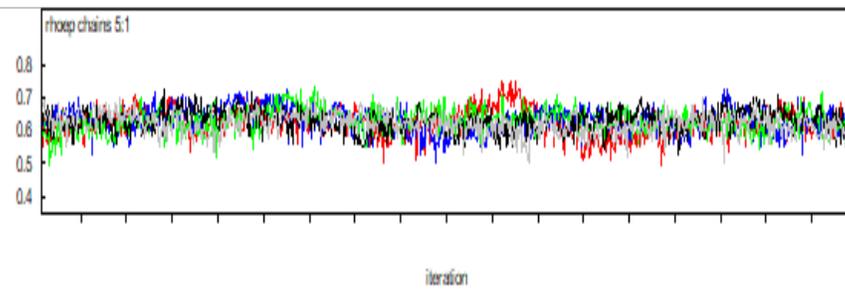


Figura 21: Parámetro  $\rho_\epsilon$

## A.2. Funciones de densidad a posteriori estimadas para los parámetros de cada modelo.

Las funciones de densidad a posteriori para cada uno de los parámetros presentes en los modelos son presentados en las figuras dadas a continuación.

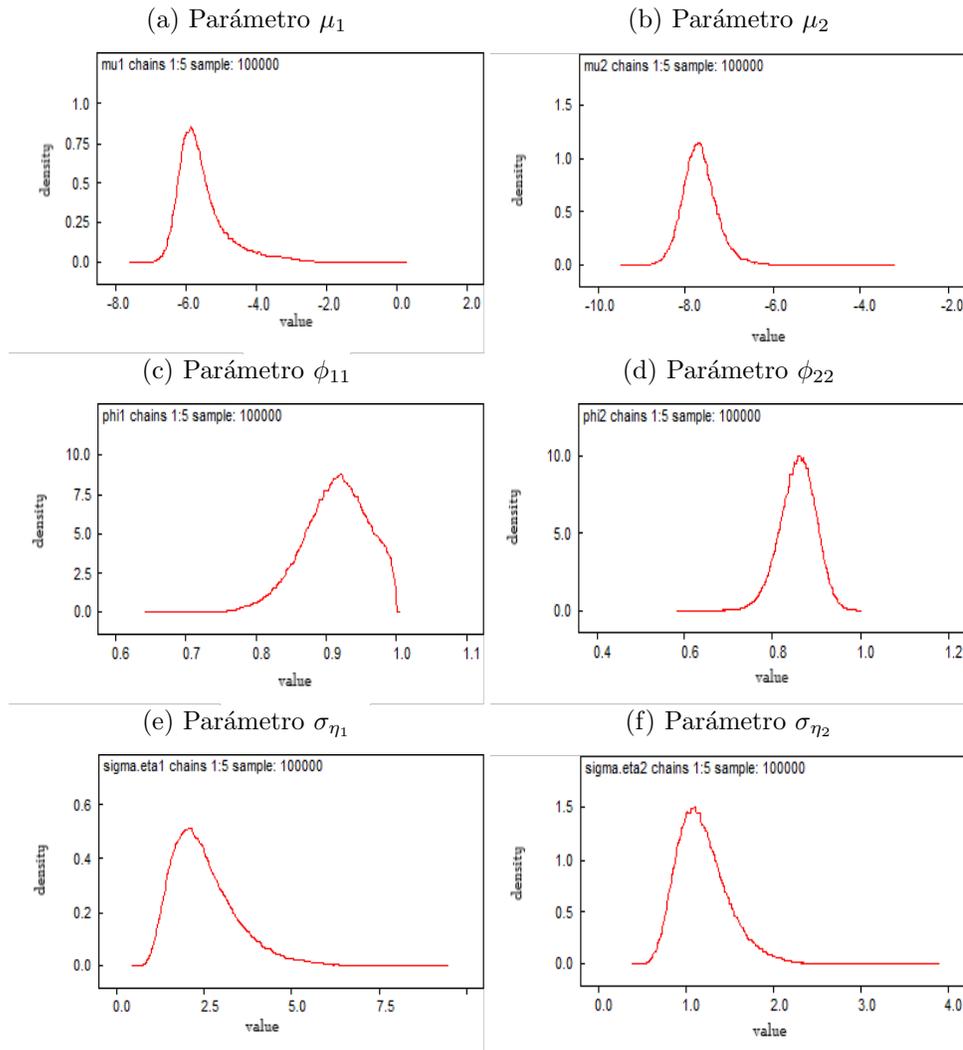


Figura 22: Funciones de densidad a posteriori estimadas para los parámetros en el Modelo I

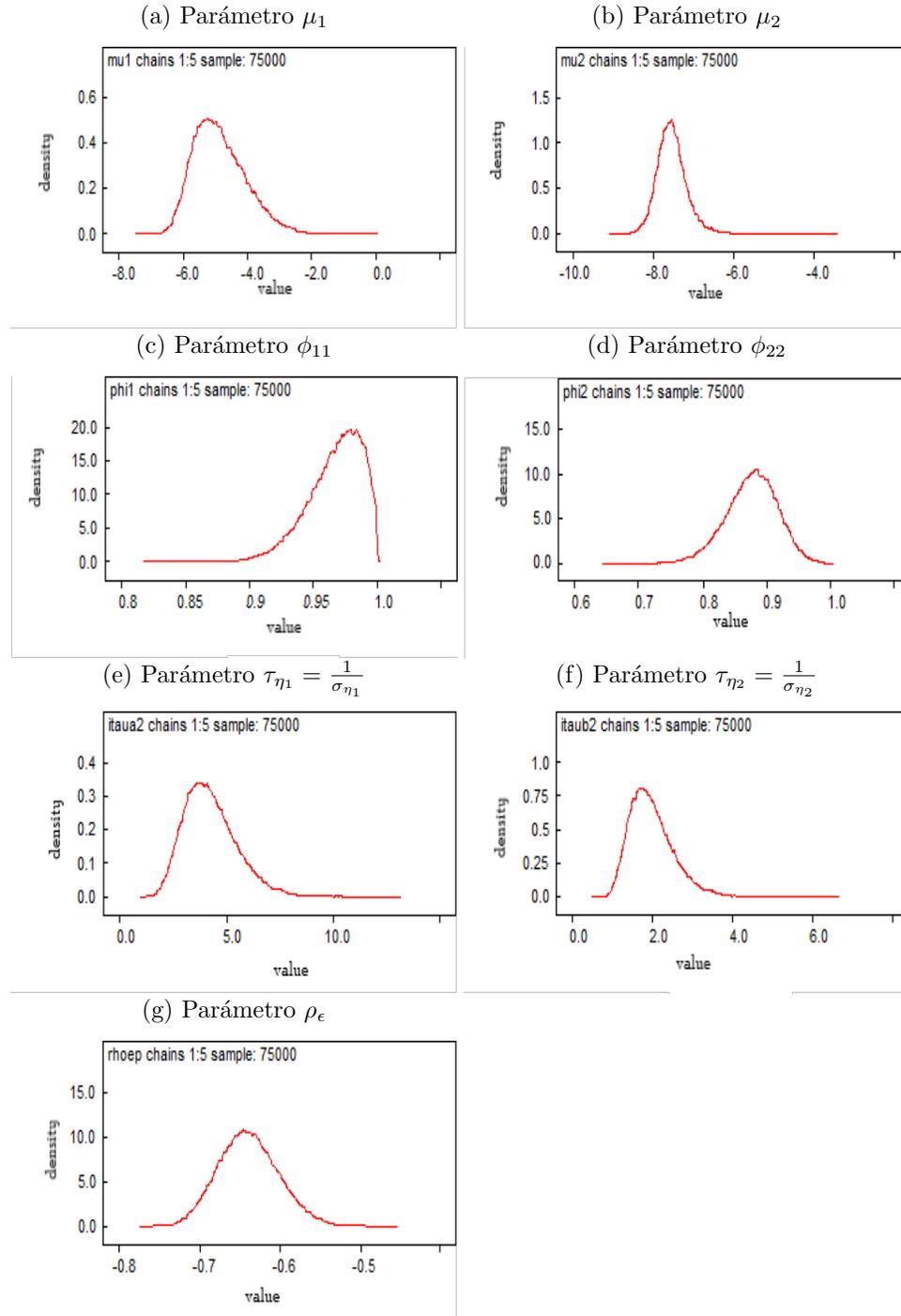


Figura 23: Funciones de densidad a posteriori estimadas para los parámetros en el Modelo II

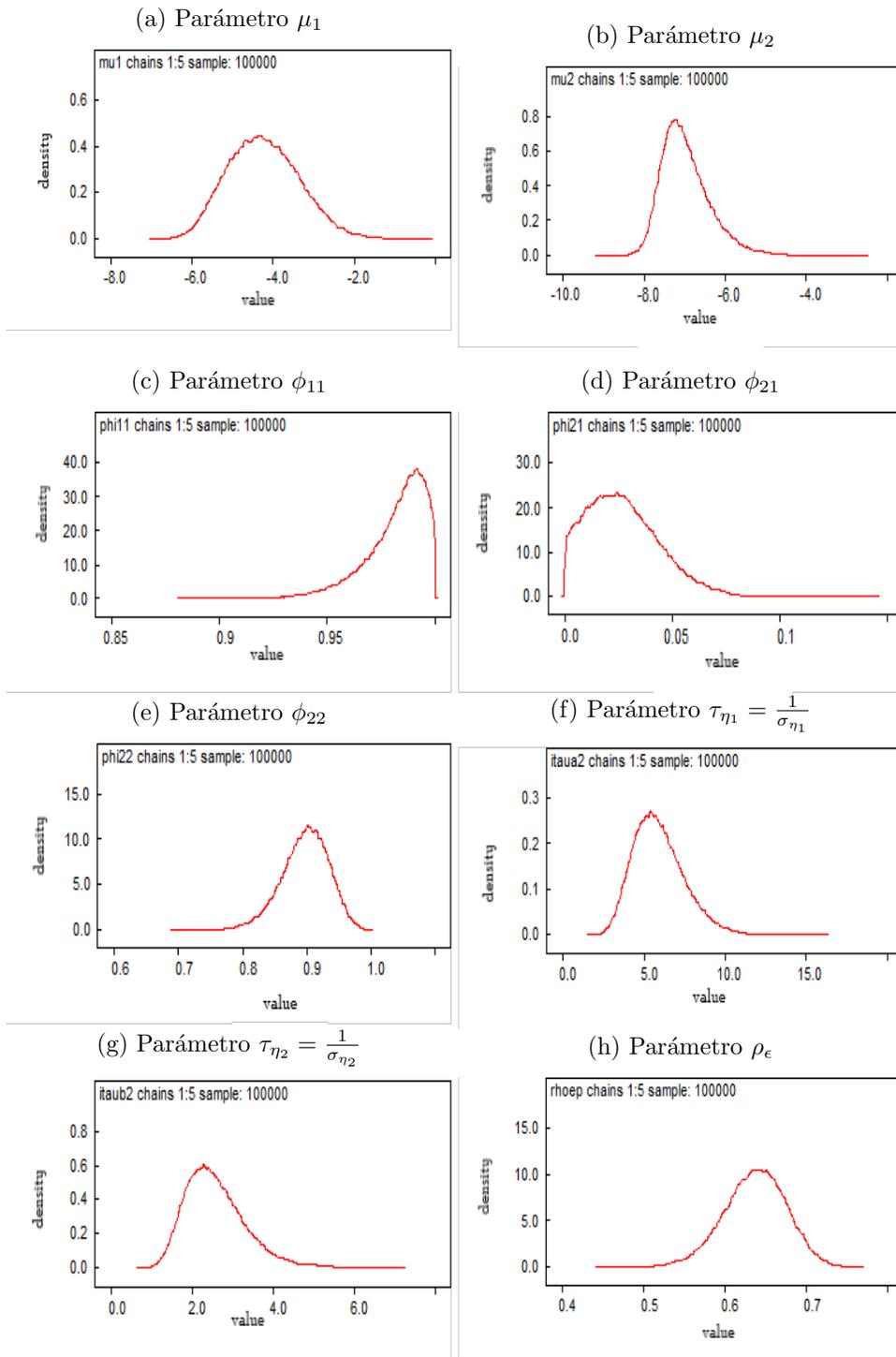


Figura 24: Funciones de densidad a posteriori estimadas para los parámetros en el Modelo III

### A.3. Inclusión de los datos al programa Winbugs

En esta sección se presenta la forma en la que los datos se adaptaron a los códigos de *Winbugs* utilizados en los modelos considerados en este trabajo. Dichos códigos pueden ser encontrados en Zozzolato (2010).

#### A.3.1. Modelo I

Los datos fueron colocados en dos vectores,  $Y1$  para los log-retornos de reservas internacionales y  $Y2$  para los log-retornos del tipo de cambio. Por lo tanto se incluyeron los datos de la siguiente forma:

```
list(N=384,
Y1 = c() ,
Y2 = c())
```

El código completo para este modelo puede ser consultado en Zozolotto (2010), página 94.

#### A.3.2. Modelo II

En este caso los datos fueron reordenados. Se intercalaron los log-retornos de ambas series, de tal forma que la estructura de los datos será una matriz bidimensional con entradas  $(Y_{1,j}, Y_{2,j})$  para  $j = 1, \dots, 384$ . Se agregaron al programa como:

```
list(N=384, Y = structure( .Data = c(Y1_1,Y2_1,Y1_2,Y2_2 ,..., Y1_384,Y2_384) , Dim = c(384,2)))
```

El código completo para este modelo puede ser consultado en Zozolotto (2010), página 95.

#### A.3.3. Modelo III

Para los datos de este modelo se mantuvo la estructura del modelo II. Por lo tanto, se agregaron al programa como:

```
list(N=384, Y = structure( .Data = c(Y1_1,Y2_1,Y1_2,Y2_2 ,..., Y1_384,Y2_384) , Dim = c(384,2)))
```

El código completo del programa para este modelo puede ser consultado en Zozolotto (2010), página 96.

#### A.4. Código implementado para las simulaciones

A continuación se presenta el código de R utilizado para las simulación de las volatilidades estimadas, los log-retornos estimados y para la predicción de los datos correspondientes al año 2019.

```
####Muestra para h,

modelo2bayesiano<-function(mu1,mu2,phi11,phi22,sigmac1,sigmac2,N){
  h11<-rnorm(1,mean=mu1,sd=(sigmac1)^(1/2))
  h22<-rnorm(1,mean=mu2,sd=(sigmac2)^(1/2))
  h1<-c(h11)
  h2<-c(h22)

  ###primero se hará la simulación de h

  if(N>1){
    for(i in 2:N){
      #usando distribución condicional
      hi1<-rnorm(1,mean=mu1+phi11*(h1[i-1]-mu1),sd=(sigmac1)^(1/2))

      hi2<-rnorm(1,mean=mu2+phi22*(h2[i-1]-mu2),sd=(sigmac2)^(1/2))

  ###se agregan h(i) al vector de muestras de h

      h1<-c(h1,hi1)
      h2<-c(h2,hi2)
    }
  }

  ###renglón h1 y con segundo renglón h2

  M<-rbind(h1,h2)
  return(M)

library(ggplot2)
```

62

```
library(scales)
library(MASS)
NMuestra=1000
m1<-modelo2bayesiano(-4.872,-7.534,0.9663,.8767,1/4.242,1/1.97,396)

###MUESTRAS DE H1 Y H2
Muestrah1=m1[1,]
Muestrah2=m1[2,]

for(i in 1:(NMuestra-1)){
  m<-modelo2bayesiano(-4.872,-7.534,0.9663,.8767,1/4.242,1/1.97,396)
  Muestrah1=rbind(Muestrah1,m[1,])
  Muestrah2=rbind(Muestrah2,m[2,])
}

h1_estimada=apply(Muestrah1,2,mean)
h2_estimada=apply(Muestrah2,2,mean)

### VOLATILIDAD ESTIMADA
v1_est=exp(h1_estimada)
v2_est=exp(h2_estimada)

### LOG RETORNOS ESTIMADOS
rho=matrix(c(1,-.6388,-.6388,1),ncol = 2,nrow = 2)
Mlogr1=matrix(0,NMuestra,396)
Mlogr2=matrix(0,NMuestra,396)

for(i in 1:396){
  eps = mvrnorm(NMuestra, c(0,0), rho)
  Mlogr1[,i] = (v1_est[i]^(1/2))*eps[,1]
  Mlogr2[,i] = (v2_est[i]^(1/2))*eps[,2]
}

logr1_est=apply(Mlogr1,2,mean)
logr2_est=apply(Mlogr2,2,mean)

###Predicciones
```

```

Reservas=ReservasReal$ReservaReal
tc=TC_Real$TCReal
Rp=logr1_est [385:396]
Tp=logr2_est [385:396]

PredR=c(Reservas[1]*exp(Rp[1]))
PredTC=c(tc[1]*exp(Tp[1]))

for(i in 1:11){

  R=PredR[i]*exp(Rp[i+1])
  PredR=c(PredR,R)

  a=PredTC[i]*exp(Tp[i+1])
  PredTC=c(PredTC,a)

}

###Gráficas

Te<-c(1:396)
G1<-data.frame(Te,logr1_est)
graph11<-ggplot(G1, aes(x=Te,y=logr1_est)) + geom_line(colour='black')
+labs(x = "Iteración",y="Log retornos estimados")

G2<-data.frame(Te,logr2_est)
graph22<-ggplot(G2, aes(x=Te,y=logr2_est)) + geom_line(colour='black')
+labs(x = "Iteración",y="Log retornos estimados")

GV1<-data.frame(Te,v1_est)
graphV1<-ggplot(GV1, aes(x=Te,y=v1_est)) + geom_line(colour='black')
+labs(x = "Iteración",y="Volatilidades estimadas para las reservas internacionales")

GV2<-data.frame(Te,v2_est)
graphV2<-ggplot(GV2, aes(x=Te,y=v2_est)) + geom_line(colour='black')
+labs(x = "Iteración",y="Volatilidades estimadas para el tipo de cambio peso dólar ")

```



# Bibliografía

- [1] Achcar, J.A., Zozolotto, H.C., Rodrigues, E.R. (2010). Bivariate stochastic volatility models applied to Mexico City ozone pollution data, in G.C. Romano & A.G. Conti (eds.), *Air Quality in the 21st Century*, Nova Publishers. Nueva York, pp. 241-267.
- [2] Adler, G., Tovar C.E. (2014). Intervenciones en el mercado cambiario y su efecto en el tipo de cambio. *Monetaria*, Centro de Estudios Monetarios Latinoamericanos, vol. 0(1), pp. 1-54, enero-jun.
- [3] Bernardo, J.M., Smith, A.F.M. (2007). *Bayesian Theory*. Segunda Edición. Nueva York: Wiley series in probability and statistics: Probability and statistics.
- [4] De Jesus Romo, V. (2013). Implementación del muestreo de Gibbs en R para la estimación de parámetros del modelo de volatilidad estocástica con aplicación a datos de contaminación. Tesis de Maestría. Facultad de Ciencias, Universidad Nacional Autónoma de México, México.
- [5] Gamerman, D., Lopes, H.F (2006). *Markov Chain Monte Carlo: stochastic simulation for Bayesian inference*. Primera Edición. Londres: Chapman & Hall/CRC Press.
- [6] Gelfand, A.E., Smith, A.F.M. (1990). Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association*, pp. 398-409.
- [7] Grimmett, G., Stirzaker, D. (2001). *Probability and Random Processes*. Tercera Edición. EUA: Oxford University Press.
- [8] Karlin, S., Taylor, H.M. (1975). *A first course in stochastic processes*. Segunda Edición. EUA: Academic Press.

- [9] Norris, J.R. (1998). Markov Chains. Primera Edición. Reino Unido: Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press.
- [10] Robert, C.P., Casella, G. (2004). Monte Carlo Statistical Methods, Segunda Edición. Nueva York: Springer-Verlang.
- [11] Ross, S. M. Stochastic processes (1996). Segunda Edición. Michigan: Wiley series in probability and statistics: Probability and statistics.
- [12] Spiegelhalter, D.J., Best, N.G., Carlin, B.P., Van Der Linde, A. (2002). Bayesian measures of model complexity and fit. Journal of the Royal Statistical Society, v.64. Alemania: p.583-639.
- [13] Spiegelhalter, D. J., Thomas, A., Best, N. G., Lunn, D. (1999). Winbugs user manual. Cambridge, Reino Unido: MRC Biostatistics Unit.
- [14] Tovar M., Yslas, R. (2016). Two Models of FX Market Interventions: The Cases of Brazil and Mexico. Banco de Mexico's Working Paper No. 2016-14.
- [15] Zozolotto, H. C. (2010). Aplicação de modelos de volatilidade estocástica em dados de poluição do ar de duas grandes cidades: cidade do México e São Paulo. Dissertação. Tesis de Maestría. Universidad de Sao Paulo, Brasil.