



**UNIVERSIDAD NACIONAL AUTONOMA  
DE MEXICO**

**FACULTAD DE CIENCIAS**

**INFERENCIA BAYESIANA PARA LA VOLATILIDAD EN EL  
MODELO BLACK & SCHOLES**

**T E S I S**

QUE PARA OBTENER EL TITULO DE:

**A C T U A R I O**

P R E S E N T A :

**CARLOS VLADIMIR RODRIGUEZ CABALLERO**



**FACULTAD DE CIENCIAS  
UNAM**

**DIRECTOR DE TESIS: DR. RAMSES HUMBERTO MENA CHAVEZ**

**2005**

m. 343224



Universidad Nacional  
Autónoma de México

Dirección General de Bibliotecas de la UNAM

**Biblioteca Central**



**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.



UNIVERSIDAD NACIONAL  
AUTÓNOMA DE  
MÉXICO

**ACT. MAURICIO AGUILAR GONZÁLEZ**  
**Jefe de la División de Estudios Profesionales de la**  
**Facultad de Ciencias**  
**Presente**

Comunicamos a usted que hemos revisado el trabajo escrito:

Inferencia Bayesiana para la Volatilidad en el Modelo Black & Scholes  
 realizado por Carlos Vladimir Rodríguez Caballero

con número de cuenta 09805922-5 , quien cubrió los créditos de la carrera de:  
 Actuaría

Dicho trabajo cuenta con nuestro voto aprobatorio.

**Atentamente**

Director de Tesis Propietario Dr. Ramsés Humberto Mena Chávez

Propietario Dr. Luis Antonio Rincón Solís

Propietario Mat. Margarita Elvira Chávez Cano

Suplente Act. Jaime Vázquez Alamilla

Suplente M.en C. Hugo Villaseñor Hernández

**Consejo Departamental de Matemáticas**

Act. Jaime Vázquez Alamilla



**FACULTAD DE CIENCIAS**  
**CONSEJO DEPARTAMENTAL**  
**DE**  
**MATEMÁTICAS**

# Dedicatoria

A MIS PADRES. *Porque gracias a su cariño, guía y apoyo he llegado a alcanzar el primer gran anhelo de mi vida. Por todos los incomparables valores que me forjaron para realizarme como un hombre de provecho. Por ser los cimientos más importantes de mi existencia.*

A MI HERMANO. *Por su eterno compañerismo, por abrirme las puertas que me han ayudado, en muy diversos aspectos, a ser el hombre que soy.*

A SARITHA. *Porque jamás podré encontrar otra persona con la cual pueda compartir tantos sueños y metas en la vida, porque este logro mío, es también de ella siendo razón y causa de la inspiración por cada uno de los presentes y futuros triunfos de mi vida.*

A RODRIGO. *Más que un amigo, un hermano. Gracias por toda la amistad incomparable que dió posibilidad a la realización de este trabajo, gracias por tus consejos y tus opiniones sobre la tesis.*

A ZAZIL HA. *Por ser la mejor amiga que alguien pudiera tener, gracias por toda la ayuda que me brindaste durante la carrera, por todas las risas, por aquellas pláticas interminables y por haber compartido conmigo tantos momentos inolvidables.*

A MIS AMIGOS EDGAR DÍAZ, ELIAS FAJARDO, ROMÁN HERNÁNDEZ Y SERGIO LÓPEZ. *En estricto orden alfabético, por su compañía a lo largo de estos 4 años de estudio, porque sin cada uno de ustedes el hecho de que me titule como acturio no sería el mismo y porque sin dudarlo buscaré su apoyo en el futuro para realizar juntos los diversos proyectos que se puedan tener.*

AL DR. RAMSÉS HUMBERTO MENA CHÁVEZ. *Por habernos tenido a Alejandro y a mí la confianza de poder encargarse de los detalles administrativos de la tesis y por sus valiosos comentarios que lograron elevar considerablemente la calidad del trabajo.*

AL DR. LUIS ANTONIO RINCÓN SOLIS, MAT. MARGARITA ELVIRA CHÁVEZ CANO, ACT. JAIME VÁZQUEZ ALAMILLA Y AL M.EN C. HUGO VILLASEÑOR HERNÁNDEZ. *Por cada uno de sus consejos, comentarios y revisiones que me pudieron permitir desarrollar un mejor trabajo.*

A FUNDACIÓN UNAM Y FUNDACIÓN TELMEX. *Por la beca que me otorgaron durante los dos años finales de mi carrera, gracias por estar comprometidos con el futuro del país.*

A LA UNAM, MI ALMA MATER. *Por la maravillosa preparación que me brindaste, por ser sin duda alguna la institución más importante del país, cuya única convicción es la de trabajar día con día en la construcción de un mejor México.*

UNA DEDICATORIA MUY ESPECIAL A  
ALEJANDRO VILLAGRÁN HERNÁNDEZ

*Por la excelente dirección de tesis que ha dado lugar al trabajo actual, por tener el carácter suficiente que permitió saber diferenciar entre una amistad y un trabajo profesional, por hacérmelo entender aún en los peores momentos. Por contagiarme de tu entrega, dedicación y esfuerzo. Por ser más que mi director de tesis, todo un amigo, por tu orientación profesional, por tus innumerables consejos, por hacerme entender las ventajas de la inferencia bayesiana. Por no caer en mis caprichos, ejemplo de ello el nombre de la tesis. Por todos aquellos enojos a larga distancia que no me permitían dormir con tranquilidad, pero que al otro día no había pasado nada y seguíamos trabajando. Gracias por todo lo que me has hecho cambiar, y por presumir que mi tesis de licenciatura la realicé bajo tu dirección.*

*Nunca cambies.*

# Índice general

<b>I. INTRODUCCIÓN</b> . . . . .	<b>1</b>
<b>II. MODELO DE BLACK &amp; SCHOLES</b> . . . . .	<b>3</b>
2.1. Las opciones como derivados financieros . . . . .	3
2.1.1. Definición . . . . .	3
2.1.2. Opciones de compra y venta . . . . .	4
2.1.3. Objetivos . . . . .	6
2.2. Valuación de opciones . . . . .	6
2.2.1. Factores que determinan los valores de las opciones . . . . .	7
2.2.2. Supuestos del modelo de Black & Scholes . . . . .	9
2.2.3. Derivación heurística de la valuación de opciones . . . . .	10
2.3. El modelo de Black & Scholes . . . . .	12
2.3.1. Propiedad lognormal del precio de las acciones . . . . .	12
2.3.2. La distribución de la tasa de retorno . . . . .	13
2.3.3. Volatilidad . . . . .	14
2.4. Prueba del modelo Black & Scholes . . . . .	15
2.4.1. Prueba . . . . .	16
2.4.2. El resultado Black & Scholes . . . . .	17
<b>III. HERRAMIENTAS ESTADÍSTICAS</b> . . . . .	<b>19</b>
3.1. Inferencia bayesiana . . . . .	20
3.1.1. Razonamiento bayesiano . . . . .	20
3.1.2. El papel del análisis bayesiano . . . . .	21
3.1.3. Teorema de Bayes . . . . .	22
3.1.4. Distribución predictiva . . . . .	24
3.1.5. Intervalos de credibilidad . . . . .	24
3.1.6. Un esquema para el análisis bayesiano . . . . .	25
3.1.7. Un ejemplo de inferencia bayesiana . . . . .	25

3.2. Simulación estocástica vía métodos MCMC . . . . .	29
3.2.1. Procesos Estocásticos . . . . .	31
3.2.2. Cadenas de Markov . . . . .	31
3.2.3. Gibbs Sampler . . . . .	33
3.2.4. Metropolis-Hastings . . . . .	35
3.2.5. Diagnóstico de convergencia . . . . .	39
3.3. Series de Tiempo Financieras . . . . .	41
3.3.1. Definición de series de tiempo . . . . .	41
3.3.2. Precio . . . . .	42
3.3.3. Retornos . . . . .	42
3.3.4. Volatilidad . . . . .	43
3.3.5. Modelos de series de tiempo . . . . .	45
3.3.6. Modelos de series de tiempo para la media . . . . .	46
3.3.7. Modelos de series de tiempo para la volatilidad . . . . .	49
3.3.8. Proceso ARCH y el modelo Black & Scholes . . . . .	53
<b>IV. INFERENCIA BAYESIANA PARA LA VOLATILIDAD EN EL MODELO BLACK &amp; SCHOLES . . . . .</b>	<b>55</b>
4.1. Alternativas realizadas y propuesta actual al modelo de Black & Scholes	56
4.1.1. Alternativas realizadas . . . . .	56
4.1.2. Propuesta actual . . . . .	57
4.2. Elementos para el desarrollo de la inferencia bayesiana sobre el modelo ARCH . . . . .	58
4.2.1. Elementos para inferencia bayesiana . . . . .	59
4.2.2. Elementos para MCMC . . . . .	61
4.3. Desarrollo MCMC . . . . .	64
4.3.1. Implementación del Metropolis-Hastings de propuesta independiente con retornos simulados . . . . .	66
4.3.2. Implementación del Metropolis-Hastings de propuesta independiente con retornos reales . . . . .	70
4.3.3. Interpretación de resultados . . . . .	77

4.3.4. Out-of-Sample . . . . .	82
<b>V. CONCLUSIONES . . . . .</b>	<b>88</b>
<b>REFERENCIAS . . . . .</b>	<b>90</b>



# Capítulo I

## INTRODUCCIÓN

Los mercados de derivados son una herramienta importante para la planeación y adecuada administración de riesgos. Sus orígenes datan del siglo XVII en lo que se denomina *Mercados Complementarios* diseñados para cubrir riesgos. Desde entonces el significado de riesgo juega un rol muy importante en el mundo de los negocios. La importancia de estudiar de manera adecuada el riesgo no significa adivinar el futuro, sino simplemente cubrirse ante riesgos que sabemos que de ocurrir, causarían una grave afectación en la empresa.

En la actualidad los derivados son negociados en mercados organizados (Bolsas) y en mercados extrabursátiles, llamados *over-the-counter*, existiendo distintos tipos de instrumentos financieros, entre los que destacan las opciones, los *forwards*, los futuros, los *swaps* y los *warrants*. En México los orígenes de este tipo de instrumentos financieros se encuentran en los Petrobonos (1977-1988), después en las Obligaciones convertibles en acciones (emitidos por los Bancos), seguido de los Títulos Opcionales o *Warrants* (1993) y finalizando con la creación del Mercado Mexicano de Derivados "*MexDer*" (diciembre de 1998). A partir de este año se empieza el trabajo con los derivados de manera organizada en bolsa de valores con el objetivo de ofrecer mecanismos de cobertura sobre las principales variables económicas que afectan a la empresa mexicana. Aunque el mercado de derivados está en pleno proceso de madurez en el país y el mundo, la utilización de las opciones como mecanismo de cobertura de riesgos financieros resulta cada vez más imprescindible.

Cuando se habla de opciones y su valuación es difícil no mencionar, en algún momento, el modelo propuesto por Fisher Black y Myron Scholes. Este modelo, a pesar de sus limitaciones, es simultáneamente uno de los más usados en la práctica financiera y uno de los pilares en la construcción de la teoría moderna de las finanzas.

Durante los años noventa se ha visto la unión simbiótica de las matemáticas, las finanzas, el desarrollo computacional y la economía global. En los mercados financieros se realizan operaciones por 2 billones de dólares diarios y son frecuentes los complejos

derivados financieros tales como las propias opciones.

Desde la aparición en 1973 de la fórmula Black & Scholes, la comunidad financiera ha adoptado un abundante conjunto de herramientas y modelos matemáticos en permanente desarrollo, gran parte de ellos se basan bajo supuestos rigurosos que han sido debatidos en los últimos años, sin embargo solo ha sido en contadas ocasiones cuando se ha intentado trabajar en ellos con el propósito de encontrar modelos más adaptables a la realidad. Aunque existen muchos supuestos que se han estudiado una y otra vez en la última década, tal vez el mayor de ellos bajo revisión es la condición de una varianza constante.

Después del desarrollo del modelo ARCH por Robert Engle en 1982, la modelación econométrica centró su atención en el problema de la heteroscedasticidad, y a partir de ello en la época actual se ha intentado trabajar con la volatilidad en modelos de valuación de opciones y demás modelos de derivados financieros, comandados siempre por el modelo de Black & Scholes como pilar de todos ellos.

En fechas recientes la simulación estocástica se ha vuelto parte esencial en los desarrollos de los modelos financieros. La palabra simulación se refiere al tratamiento de un problema real a través de la reproducción controlada en un ambiente experimental. Dicho ambiente es frecuentemente proporcionado por equipos computacionales. Las técnicas de simulación estocástica, como lo son las Monte Carlo, tienen características que explican sus éxitos recientes en la inferencia estadística y su implementación en modelos financieros.

Por otro lado la inferencia bayesiana en los modelos ARCH ha sido implementada cada vez con mayor continuidad usando algoritmos de simulación estocástica, comenzando con el Gibbs Sampler y más recientemente con el algoritmo del Metropolis-Hastings.

Ahora bien la línea de investigación en el modelo Black & Scholes y a su vez del mercado de derivados se centran en dos puntos: la volatilidad en el mercado y la distribución para el precio de las opciones. En la presente tesis se buscan ambos puntos con el objetivo de proponer un mejor manejo de la volatilidad para el modelo y la exposición de una distribución posterior para el precio de una opción de compra o venta.

## Capítulo II

### MODELO DE BLACK & SCHOLES

#### *2.1. Las opciones como derivados financieros*

Un producto financiero derivado es un instrumento financiero cuyo valor depende del valor de otros, ver Stampfli (2004). Es decir, el valor podría derivarse en forma indirecta del valor de otro instrumento intercambiado. En este caso, el precio futuro estará siempre ligado al precio del otro valor en una fecha futura. A este tipo de instrumentos financieros se les denomina **derivado financiero**, cuyo valor se denomina **valor o activo subyacente**.

En años recientes los futuros y las opciones se han convertido en mercados muy importantes en el mundo de las finanzas y de las inversiones. Se ha alcanzado el punto donde es esencial que todos los profesionales en finanzas entiendan cómo es que trabajan estos mercados, cómo pueden ser usados, qué determina el precio de estos instrumentos y su forma de valuación.

El objetivo de este capítulo es familiarizarse con la terminología básica de la valuación de opciones, seguido de una presentación introductoria del modelo Black & Scholes y finalizando con la obtención formal del mismo.

##### **2.1.1. Definición**

Las opciones fueron por primera vez comercializadas en un mercado organizado en 1973. Desde entonces ha habido un crecimiento importante en los mercados de opciones. Las opciones son ahora comercializadas en muchas bolsas alrededor del mundo. Enormes volúmenes de opciones son también comercializadas *over-the-counter* por bancos y otras instituciones financieras.

Una opción es un contrato que le proporciona a su poseedor el derecho, más no la obligación, de comprar o vender algún activo a un precio fijo en una fecha determinada o antes de ella. Haciendo hincapié en la definición se debe observar que las opciones son un tipo único de contrato financiero porque le proporcionan al comprador el derecho, pero no la obligación, de hacer algo; es decir, el comprador usa la opción tan sólo si ello representa una alternativa conveniente; de lo contrario, la opción puede ser desechada. Existe un vocabulario especial asociado con este tipo de

producto financiero derivado, para revisar más detalles, ver Hull (2000).

Por el derecho que otorga la opción al comprador de la misma, existen dos tipos:

- Opciones de compra (call option)
- Opciones de venta (put option)

### 2.1.2. Opciones de compra y venta

El tipo más común de opción recibe el nombre de *opción de compra*, tal y como se define en Ross (1999), este instrumento financiero le proporciona a su propietario el derecho de comprar un cierto activo a un precio fijo durante un período determinado. Las condiciones de la opción de compra son :

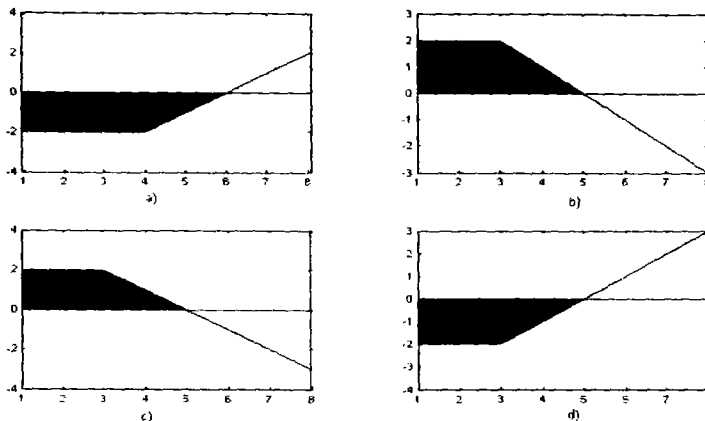
- El comprador de la opción paga al vendedor una comisión llamada prima.
- En la fecha de vencimiento, el tenedor de este contrato podría pagarle al emisor del mismo el precio de ejercicio.
- Si el emisor del contrato recibe el precio de ejercicio del tenedor, el emisor tiene que entregar una acción al tenedor en la fecha de vencimiento.

Por su parte, una *opción de venta* le proporciona al tenedor, el derecho a vender las acciones a un precio de ejercicio fijo hasta una fecha predeterminada. Dicho de otra forma se le conoce como opción de venta a la posibilidad de comprar una oportunidad para vender una acción en el futuro a un precio garantizado, incluso si no se es propietario de acción alguna. Las condiciones de la misma son:

- El comprador de la opción paga al vendedor una comisión llamada prima.
- En la fecha de vencimiento, el tenedor de este contrato puede darle al emisor una acción o, en forma equivalente, el precio de mercado de una acción.
- Si el emisor del contrato recibe del tenedor la acción o su precio, el emisor tiene que pagar la comisión de ejercicio al tenedor en la fecha de vencimiento.

En aspectos financieros es importante conocer un perfil de pérdidas y ganancias ya que es éste el que permite conocer y comprender la evolución que tenga un instrumento financiero.

El perfil de pérdidas y ganancias para una opción de compra y una opción de venta para el inversionista que mantiene una posición larga y una corta se presenta en la figura (1).



**Figura 1:** Perfiles de rendimiento. a) Call largo, b) Call corto, c) Put largo, d) Put corto. El área en azul representa la región comprendida por la prima.

A partir de la figura (1), se resume la información de la siguiente manera:

- Call largo. *Derecho de compra de un activo a un precio fijo en un plazo determinado.* Se debe de utilizar en una tendencia alcista, la ganancia se incrementará tanto como el precio del activo suba. La pérdida será la prima.
- Call corto. *Obligación de venta de un activo a un precio fijo en un plazo determinado.* Se debe de utilizar en una tendencia bajista o paralela, la ganancia será la prima y la pérdida se incrementará tanto como el precio del activo suba.
- Put largo. *Derecho de venta de un activo a un precio fijo en un plazo determinado.* Se debe de utilizar en una tendencia bajista, la ganancia se incrementará tanto como el precio del activo baje y la pérdida será la prima.
- Put corto. *Obligación de compra de un activo a un precio fijo en un plazo determinado.* Se debe de utilizar en una tendencia alcista, la ganancia será la prima y la pérdida se incrementará tanto como el precio del activo baje.

De manera similar, las opciones se pueden clasificar también de acuerdo al tiempo en que se puede ejercer el derecho que ellas otorgan en:

- Opciones europeas

- Opciones americanas

Concorde con Diaz (2002), las opciones europeas son aquéllas que sólo pueden ser ejercidas en la fecha de vencimiento; mientras que las opciones americanas son aquéllas que se pueden ejercer durante la vida de la opción, es decir, en cualquier momento antes de la expiración.

### 2.1.3. Objetivos

Es también importante identificar los objetivos para los cuales se utiliza este tipo de productos derivados. Los objetivos de las opciones se pueden agrupar generalmente en dos categorías de acuerdo al nivel agregado. Primero, los objetivos a nivel microeconómico y, segundo, al macroeconómico. Una opción tiene básicamente dos objetivos a nivel microeconómico:

- Es un producto financiero con el cual un inversionista puede protegerse del riesgo.
- Su utilización podría ser usado por los inversionistas simplemente para invertir o especular.

El término *especular* no tiene aquí una connotación negativa. De hecho, puede ser tomado como una actividad totalmente válida y hasta sana, en el sentido de que provee liquidez a los mercados.

A nivel macroeconómico se encuentran los siguientes objetivos:

- Formación más eficiente de precios de los valores subyacentes.
- Mejorar los niveles de liquidez en el mercado.
- Ampliar las oportunidades de arbitraje.
- Permitir perfiles de riesgo y rendimientos controlables.

## 2.2. *Valuación de opciones*

Se ha explicado de manera general el significado de una opción, sus características más relevantes y la forma en que operan tanto en un mercado organizado como en el *over-the-counter*. Sin embargo es necesario, aparte de conocer todo lo anterior, comprender los modelos existentes que hacen posible la valuación de opciones. Al

hablar de valuación de opciones se debe mencionar el modelo propuesto por Fisher Black, Myron Scholes y Robert Merton a principios de los 70's. De hecho, y a pesar de las limitaciones del modelo, éste ha tenido una enorme influencia en la forma en que se comercializan las opciones en el mercado financiero, siendo el pivote más importante del crecimiento y los éxitos de la ingeniería financiera como pilar de la teoría moderna de las finanzas en los ochentas y noventas. Sin embargo no fue sino hasta 1997 cuando el modelo rindió frutos, la Fundación Nobel otorgó el Premio Nobel de Economía a Myron Scholes y Robert Merton, también se otorgó un respetuoso recuerdo al ya entonces acaecido Fisher Black.

La finalidad de esta sección es brindar una idea clara del modelo y sus implicaciones técnicas, por el momento dando solamente explicaciones intuitivas. El análisis de Black & Scholes se centra en valorar opciones cuyo subyacente no paga dividendo. Se realizan los siguientes supuestos sobre el subyacente:

- Se mueve suave y continuamente.
- Tiene una tasa de retorno instantánea  $m$ .

Para verificar los supuestos así como sus hipótesis, ver Sabau (1997) .

El modelo más importante en la valuación de opciones es el de Black & Scholes. Primero se explicarán los factores y los supuestos bajo los cuales está definido el modelo y después se derivará de manera sumamente informal el mismo, de tal manera que la fórmula tenga sentido y pueda ser interpretada con facilidad. A esta derivación se le denominará *derivación heurística*. Posteriormente, se presentará un análisis más detallado. Al final del capítulo se presenta la obtención formal de la fórmula de Black & Scholes.

### **2.2.1. Factores que determinan los valores de las opciones**

Aquí se hace referencia solamente a las opciones americanas porque son las que se negocian en el mundo real. Las diferencias de las opciones europeas en comparación con las americanas pueden verse en Ross (1999) y Hull (2000).

Los factores que determinan los valores de una opción de compra pueden clasificarse con base en dos conjuntos.

El primero contiene las características de un contrato de opciones; las dos características son el precio de expiración y la fecha de ejercicio.

El segundo conjunto de factores que afecta el precio de la opción de compra está relacionado con las características de las acciones y del mercado.

**Precio de ejercicio (*strike price*).** Debe de entenderse como el precio al cual el tenedor de una opción puede comprar o vender la acción de base. Mientras más alto sea el precio de ejercicio, más bajo será el valor de una opción de compra.

**Fecha de expiración.** El valor de una opción americana debe ser por lo menos tan grande como el valor de otra opción que fuera idéntica pero con un plazo de expiración más corto. A diferencia de esto, las opciones europeas no necesitan tener esta relación por la forma en que están definidas.

**Precio de las acciones.** Mientras más alto sea el precio de una acción, más valiosa será la opción de compra

**Variabilidad del activo subyacente.** Mientras mayor sea la variabilidad del activo subyacente, más valiosa será la opción de compra.

**La tasa de interés.** Los precios de las opciones de compra también están en función del nivel de las tasa de interés, de hecho, el valor de una opción de compra se encuentra positivamente relacionado con las tasas de interés.

Una vez examinado a muy grandes rasgos los factores que determinan los valores de una opción de compra, es sencillo examinar cuales son los factores que determinan el valor de una opción de venta, de hecho el comportamiento de los tres factores sobre el valor de una opción de compra es completamente opuesto a la de una opción de venta.

A grandes rasgos se puede resumir lo siguiente.

- 1 El precio de mercado de la opción disminuye a medida que aumenta el precio de la acción.
- 2 El valor de mercado de una opción de venta con un precio de ejercicio alto es mayor que el valor de una opción de venta que fuera idéntica excepto en que tuviera un precio de ejercicio bajo.



3 Una tasa de interés alta afecta de manera adversa el valor de una opción de venta.

El efecto de los otros dos factores (volatilidad y el plazo para la fecha de ejercicio) es el mismo en el valor de una opción de compra o venta.

A manera de resumen en el cuadro (1) se reúnen los cinco factores más importantes que influyen en el precio de una opción (compra o venta) americanas.

	Opción de compra	Opción de venta
Valor del activo subyacente	+	-
Precio de ejercicio	-	+
Volatilidad de la acción	+	+
Tasa de interés	+	-
Plazo para la fecha de ejercicio	+	+

**Cuadro 1:** Factores que afectan los valores de las opciones americanas

Los signos(+,-), en el cuadro (1), indican el efecto de las variables sobre el valor de la opción.

### 2.2.2. Supuestos del modelo de Black & Scholes

Se ha explicado que el valor de una opción de compra es una función de cinco variables:

- 1 El precio actual del valor subyacente.
- 2 El precio de ejercicio.
- 3 El plazo hasta la fecha de expiración.
- 4 La variabilidad del activo de base.
- 5 La tasa de interés libre de riesgo.

Conociendo ya esto, se proseguirá a presentar el modelo de Black & Scholes, el cual justamente lo que hace es ajustar estas cinco variables para poder calcular el valor de compra de una opción. Sin embargo antes de presentar la derivación heurística y la obtención formal de dicho modelo es necesario conocer los supuestos bajo los cuales está planteado.

- Se mueve suave y continuamente.
- Tiene una tasa de retorno instantánea  $m$ .
- El precio de las acciones sigue un movimiento basado en un crecimiento constante con perturbaciones aleatorias frecuentes.
- El subyacente no paga dividendos.
- La volatilidad se supone conocida y constante durante la vida de la opción.
- La tasa de interés libre de riesgo es constante durante la vida de la opción.
- Un inversionista al vender una acción u opción en corto tendrá disponibles todos los recursos producto de la venta.
- No hay costos de transacción para las acciones, tampoco para las opciones.
- Las transacciones que realice un inversionista no afectan la tasa de impuestos que éste va a pagar.
- La opción es de tipo europeo.

Acorde con la utilización del modelo en el ejercicio profesional, no todos los supuestos tienen la misma importancia en la valuación de opciones. Muchos de ellos aceptan ajustes sin alterar en gran medida el valor de la opción, otros no. De hecho, hay quienes aseguran que solo con que algunos supuestos se cumplan el mercado se comportará de una manera similar al modelo. Su validez e importancia se estudiará más adelante, no obstante se debe mencionar que el modelo supone una volatilidad constante en el tiempo, y esto no es del todo cierto, de hecho resulta ser tan erróneo que llega a cuestionar la veracidad de la utilización del modelo. A partir de este hecho, lo que se busca en el presente trabajo es cuestionar que si bien el modelo Black & Scholes supone una volatilidad constante, esta suposición no es del todo correcta y que es entonces necesario modelar la volatilidad como un parámetro que evoluciona a través del tiempo.

Al final del capítulo se abundará más en este punto.

### **2.2.3. Derivación heurística de la valuación de opciones**

El cenit de todos los procedimientos de valuación de opciones se alcanza cuando se llega a la fórmula de Black & Scholes. Este modelo matemático está diseñado para

calcular el precio de una opción europea ya sea de compra o venta.

La opción de compra se calcula de la siguiente manera

$$c = S_0 \Phi(d_1) - X e^{-rT} \Phi(d_2) \quad (1)$$

Mientras que la opción de venta se calcula

$$p = X e^{-rT} \Phi(-d_2) - S_0 \Phi(-d_1) \quad (2)$$

Donde  $d_1 = \frac{\ln(S_0/X) + (r + \sigma^2/2)T}{\sigma\sqrt{T}}$  y  $d_2 = \frac{\ln(S_0/X) + (r - \sigma^2/2)T}{\sigma\sqrt{T}} = d_1 - \sigma\sqrt{T}$  y  $\Phi(x)$  es la función de distribución acumulativa de probabilidad de una variable aleatoria que está normalmente distribuida con media cero y varianza uno.

Los demás parámetros son:

- $S_0$  = es el precio actual de las acciones
- $X$  = precio de ejercicio de una opción de compra.
- $r$  = tasa de rendimiento continua y libre de riesgo anualizada.
- $\sigma^2$  = Varianza (por año) del rendimiento continuo sobre la acción.
- $T$  = tiempo (en años) para la fecha de expiración.

La derivación heurística se hará sobre una opción de compra y puede ser fácilmente extendida sobre una opción de venta. Para ver más detalles acerca de la derivación heurística del modelo Black & Scholes, ver Jarrw (1983).

Esta derivación concluye que el valor de una opción de compra, cuyo subyacente lo constituye una acción, es simplemente el valor presente de la posible cantidad *dentro del dinero* en la fecha del vencimiento. En términos financieros se dice que una opción está dentro del dinero si el precio está sobre el strike en un call o debajo de éste en un put.

A continuación se desarrolla esta forma de valorar la opción. Lo primero que se debe de hacer es observar que el valor de una opción de compra en la fecha de vencimiento es:

$$C_1 = \begin{cases} S_1 - X & \text{si } C \text{ se encuentra dentro del dinero} \\ 0 & \text{si } C \text{ se encuentra fuera del dinero} \end{cases} \quad (3)$$

Este valor necesita ser descontado para obtener el valor presente. Por ello se puede decir, hasta el momento, que el valor de una opción de compra es la mayor cantidad

entre el valor presente de la cantidad dentro del dinero en el vencimiento y cero, ya que el valor de una opción nunca será negativa.

Matemáticamente la relación anterior queda expresada de la siguiente manera:

$$C = e^{-rT} \text{Max}[S_1 - X, 0] \quad (4)$$

Cabe señalar que se ignora cuál es el precio de la opción en el período 1 por lo que aparentemente ya no es posible continuar con la derivación. No obstante, con una buena estimación de cuál será dicho precio parece no haber un gran problema. Más específicamente, siendo posible estimar algunos precios de la acción al día de expiración de la opción así como su probabilidad de ocurrencia, la fórmula (4) cobra sentido y se convierte en un proceso equivalente a definir un rango potencial que cubra los posibles precios de las acciones al día de la expiración de la opción, calculando el valor intrínseco con cada uno de los posibles precios estimados al definir el rango ya mencionado, después será necesario ponderar cada valor intrínseco positivo por su respectiva probabilidad de ocurrencia. Después de este pequeño procedimiento se deben sumar todos los valores encontrados y expresarlo finalmente en valor presente. Por consiguiente, este proceso dice que el valor de una opción de compra es tan solo el valor presente de la suma de los posibles valores intrínsecos positivos ponderados cada uno por su probabilidad de ocurrencia. Este procedimiento es exactamente el que el modelo de Black & Scholes sigue para valuar las opciones de compra.

### 2.3. El modelo de Black & Scholes

Lo que se hace en esta sección es analizar completamente la derivación del modelo de Black & Scholes.

#### 2.3.1. Propiedad lognormal del precio de las acciones

Una variable aleatoria  $Y$  se distribuye lognormal con parámetros  $\mu$  y  $\sigma$  si  $\ln(Y)$  es una variable aleatoria normal con media  $\mu$  y varianza  $\sigma^2$ . Esto es,  $Y$  es lognormal si puede ser expresado como  $Y = e^X$ , donde  $X$  es una variable aleatoria normal.

Sea  $S(t)$  el precio de la acción en el tiempo  $t$ . El modelo para  $S$  está dado por la ecuación diferencial estocástica, para más detalles, ver Hull (2000).

$$dS = \mu S dt + \sigma S dz \quad (5)$$

donde  $\mu$  y  $\sigma$  son constantes y  $z$  es un proceso de Wiener estándar. factorizando  $S$  se obtiene

$$\frac{dS}{S} = \mu dt + \sigma dz$$

ahora bien  $\frac{dS}{S} = d\ln(S)$  y tomando en cuenta que  $S(t)$  se distribuye lognormal,

$$E[\ln(S(t))] = \mu - \frac{1}{2}\sigma^2$$

$$\sigma[\ln(S(t))] = \sigma$$

Por tanto,

$$d \ln(S) = \left( \mu - \frac{1}{2}\sigma^2 \right) dt + \sigma dz \quad (6)$$

de donde se observa que la variable  $\ln(S)$  sigue un proceso de Wiener generalizado. Por lo tanto, el cambio en  $\ln(S)$  entre el tiempo 0 y  $T$  está normalmente distribuido tal que

$$\ln(S_T) - \ln(S_0) \sim N \left[ \left( \mu - \frac{\sigma^2}{2} \right) T, \sigma\sqrt{T} \right]$$

Se sigue que

$$\ln \left( \frac{S_T}{S_0} \right) \sim N \left[ \left( \mu - \frac{\sigma^2}{2} \right) T, \sigma\sqrt{T} \right] \quad (7)$$

y finalmente

$$\ln(S_T) \sim N \left[ \ln(S_0) + \left( \mu - \frac{\sigma^2}{2} \right) T, \sigma\sqrt{T} \right] \quad (8)$$

donde  $S_T$  es el precio de la acción a un tiempo futuro  $T$ ,  $S_0$  es el precio del mismo al tiempo 0, y  $N(\mu, \sigma)$  denota una distribución normal con media  $\mu$  y desviación estándar  $\sigma$ . La ecuación (8) muestra la condición de la distribución lognormal que se había mencionado al comienzo de esta sección.

### 2.3.2. La distribución de la tasa de retorno

La propiedad de que los precios accionarios suelen distribuirse de manera lognormal, puede ser usada para proveer información acerca de la función de distribución para la tasa de retorno ganada en una acción entre el tiempo 0 y  $T$ . Es importante mencionar que la tasa de retorno se compone de manera continua y anual. Se define a dicha tasa entre el tiempo 0 y  $T$  como  $\eta$ . Entonces

$$S_T = S_0 e^{\eta T} \quad (9)$$

y

$$\eta = \frac{1}{T} \ln \frac{S_T}{S_0} \quad (10)$$

De la ecuación (7) se sigue que

$$\eta \sim N \left( \mu - \frac{\sigma^2}{2}, \frac{\sigma}{\sqrt{T}} \right) \quad (11)$$

Entonces la tasa de retorno anual compuesta continuamente está normalmente distribuida con media  $\mu - \frac{\sigma^2}{2}$  y desviación estándar  $\frac{\sigma}{\sqrt{T}}$

### 2.3.3. Volatilidad

El modelo de Black & Scholes supone una volatilidad constante tal y como se ha dicho ya con anterioridad. Esto supone que a partir de la ecuación (11), la volatilidad del precio de una acción puede estar definida como la desviación estándar de un retorno, también de la ecuación (8) se muestra que la volatilidad es la desviación estándar del logaritmo natural del precio de la acción al final del año.

A pesar de que a lo largo del presente trabajo se supone que la volatilidad no es constante a lo largo del tiempo, se presenta la manera empírica de obtenerla, según el modelo de Black & Scholes. Se define:

$n + 1 =$  número de observaciones.

$S_i =$  precio de la acción al final del  $i$ -ésimo intervalo ( $i = 0, 1, \dots, n$ ).

$t =$  longitud del intervalo de tiempo en años

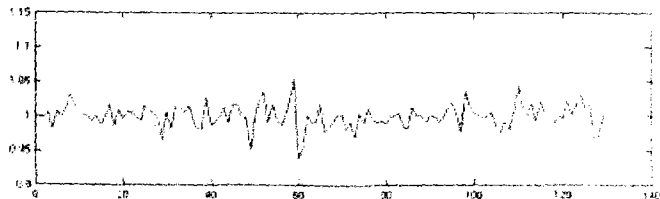
$$\text{y sea } u_i = \ln \left( \frac{S_i}{S_{i-1}} \right) \text{ para } i = 1, 2, \dots, n$$

Ya que  $S_i = S_{i-1}e^{u_i}$ ,  $u_i$  es la tasa de retorno en el  $i$ -ésimo intervalo. La tasa no está anualizada. El estimador usual,  $s$ , de la desviación estándar de la  $u_i$ 's está dado por  $s = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (u_i - \bar{u})^2}$  donde  $\bar{u}$  es la media de la  $u_i$ 's. De la ecuación (7), la desviación estándar de la  $u_i$ ,  $s$  es  $\sigma\sqrt{\tau}$ , la variable  $s$  es por lo tanto un estimador de  $\sigma\sqrt{\tau}$ . Entonces  $\sigma$  por si mismo puede ser estimado como  $\sigma^*$ , donde  $\sigma^* = \frac{s}{\sqrt{\tau}}$

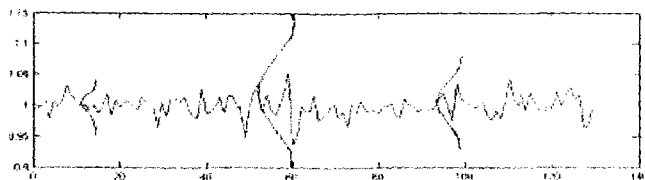
En la sección (1.2.2) ya se hizo hincapié en que la volatilidad de los rendimientos de una acción no es constante, no es tan fácil entender o convencerse de esto. Regresando al problema original, acerca de ésta conjetura, la forma más sencilla de poder entender porque el supuesto de que la volatilidad del retorno sobre el precio de la acción sea constante a lo largo del tiempo no tiene exactamente la validez que se requiere, es precisamente observar la figura (2).

De esta figura se debe entender que cada punto a lo largo del tiempo para los retornos en específico de esta acción es la realización de un proceso estocástico, de esta manera y abusando de la figura anterior, se puede entonces comprender lo siguiente.

Esto es, que en cada instante del tiempo, siendo la realización de un proceso estocástico diferente, entonces se encuentran distribuciones (en este caso de la distribución normal) con valores diferentes en sus parámetros; es decir, que si bien en



**Figura 2:** Retornos del 24/11/03 al 23/04/04 de ARA\*



**Figura 3:** Ejemplificación de la realización de un proceso estocástico

cada instante del tiempo la realización del proceso estocástico dá como resultado varias distribuciones gaussianas, éstas no necesariamente deben tener los mismos valores en sus parámetros. Lo cual a partir de la figura (3) es claro observarlo, ya que la distribución normal dibujada en la parte de en medio, su varianza debe ser mayor que cualquiera de las otras dos, estos es porque en la realización del punto mencionado su variabilidad es mayor, como se observa en la misma figura. Así es que la variable aleatoria del retorno del precio nominal de la acción no puede estar idénticamente distribuida. El análisis anterior podría encontrarse algo confuso ya que se requiere entender el significado de lo que sería una realización de un proceso estocástico, sobretodo saber reconocer que una simple realización podría cambiar el valor de los parámetros de la distribución o la distribución misma; finalmente esto es lo que hace considerar de nuevo el supuesto de que la volatilidad sea o no constante.

## 2.4. Prueba del modelo Black & Scholes

Una prueba sumamente formal de la ecuación diferencial estocástica de Black & Scholes puede ser encontrada en Luenberger (1999). En esta sección solo se presenta la obtención formal del modelo Black & Scholes.

Ya se ha presentado dicho modelo en la ecuación (1).

El resultado que se quiere probar es que si  $V$  está lognormalmente distribuida y la

desviación estándar de  $\ln(V)$  es  $s$  entonces

$$E[\max(V - X, 0)] = E(V)N(d_1) - XN(d_2) \quad (12)$$

donde

$$d_1 = \frac{\ln[E(V)/X] + s^2/2}{s}$$

$$d_2 = \frac{\ln[E(V)/X] - s^2/2}{s}$$

y  $E$  denota el valor esperado.

### 2.4.1. Prueba

Se define  $g(V)$  como la función de densidad de probabilidad de  $V$ . Entonces

$$E[\max(V - X, 0)] = \int_X^\infty (V - X) g(V) dV \quad (13)$$

La variable  $\ln(V)$  está normalmente distribuida con desviación estándar  $s$ . A partir de las propiedades de la distribución lognormal la media de  $\ln(V)$  es  $m$  donde

$$m = \ln[E(V)] - \frac{s^2}{2} \quad (14)$$

Se define una nueva variable

$$Q = \frac{\ln(V) - m}{s} \quad (15)$$

Esta variable está normalmente distribuida con media 0 y desviación estándar 1. Se denota la función de densidad para  $Q$  por  $h(Q)$  tal que  $h(Q) = \frac{1}{\sqrt{2\pi}} e^{-\frac{Q^2}{2}}$ . Ahora, usando la ecuación (14) para convertir la expresión en el lado derecho de la ecuación (13) desde una integral sobre  $V$  para una integral sobre  $Q$  se obtiene

$$\widehat{E}[\max(V - X, 0)] = \int_{(\ln(X)-m)/2}^\infty (e^{Qs+m} - X) h(Q) dQ$$

o bien

$$\widehat{E}[\max(V - X, 0)] = \int_{(\ln(X)-m)/2}^\infty (e^{Qs+m}) h(Q) dQ$$

$$- X \int_{(\ln(X)-m)/2}^\infty h(Q) dQ \quad (16)$$

Ahora

$$e^{Qs+m} h(Q) = \frac{1}{\sqrt{2\pi}} e^{(-Q^2 + 2Qs + 2m)/2}$$



$$\begin{aligned}
&= \frac{1}{\sqrt{2\pi}} e^{[-(Q-s)^2 + 2m + s^2]/2} \\
&= \frac{e^{m+s^2/2}}{\sqrt{2\pi}} e^{[-(Q-s)^2]/2} \\
&= e^{m+s^2/2} h(Q-s)
\end{aligned}$$

Esto significa que la ecuación (16) llega a ser

$$\begin{aligned}
\widehat{E}[\max(V-X, 0)] &= e^{m+s^2/2} \int_{(\ln(X)-m)/s}^{\infty} h(Q-s) dQ \\
&\quad - X \int_{(\ln(X)-m)/s}^{\infty} h(Q) dQ
\end{aligned} \tag{17}$$

Si se define  $N(x)$  como la probabilidad que una variable con media de 0 y desviación estándar de 1 sea menor que  $x$ , la primera integral en la ecuación (17) es

$$1 - N \left[ \frac{\ln(X) - m}{s} - s \right]$$

o

$$N \left[ \frac{-\ln(X) + m}{s} + s \right]$$

Sustituyendo por  $m$  a partir de la ecuación (14) se tiene que

$$N \left[ \frac{\ln[E(V)/X] + s^2/2}{s} \right] = N(d_1)$$

De manera similar la segunda integral en la ecuación (16) es  $N(d_2)$ , entonces la ecuación (17) es

$$\widehat{E}[\max(V-X, 0)] = e^{m+s^2/2} N(d_1) - X N(d_2)$$

Entonces, sustituyendo por  $m$  a partir de la ecuación (14) se sigue lo que se quería demostrar.

#### 2.4.2. El resultado Black & Scholes

Ahora se considera un opción de compra de una acción que no paga dividendo en el tiempo  $T$ . El precio de ejercicio es  $X$ , la tasa de libre de riesgo es  $r$ , el valor actual es  $S_0$ , y la volatilidad es  $s$ . Como se muestra en la ecuación siguiente

$$c = e^{-rT} \widehat{E}[\max(S_T - X, 0)] \tag{18}$$

donde  $S_T$  es el precio del activo al tiempo  $T$  y  $\widehat{E}$  denota la esperanza en un mundo de riesgo neutral. Bajo el proceso estocástico asumido por Black & Scholes (5),  $S_T$  es

lognormal. También a partir de la ecuación (8) y  $E(S_T) = S_0 e^{\mu T}$  (valor esperado de  $S_T$ ),  $\widehat{E}(S_T) = S_0 e^{rT}$  y la desviación estándar de  $\ln(S_T)$  es  $\sigma\sqrt{T}$ .

A partir de la fórmula (18) que se acaba de probar implica

$$c = S_0 N(d_1) - X e^{-rT} N(d_2)$$

donde

$$d_1 = \frac{\ln[S_0/X] + (r + \sigma^2/2) T}{\sigma \sqrt{T}}$$
$$d_2 = \frac{\ln[S_0/X] + (r - \sigma^2/2) T}{\sigma \sqrt{T}} = d_1 - \sigma\sqrt{T}$$

Que era la ecuación (1). Justamente el modelo Black & Scholes para calcular el precio de una opción de compra.

En la última ecuación se debe resaltar  $\sigma$  para indicar justamente que se supone constante y de esta manera cubrir completamente la presentación del modelo de Black & Scholes, con el objetivo de que sirva de apoyo teórico para el tercer capítulo en donde se trabajará profundamente con este modelo.

## Capítulo III

# HERRAMIENTAS ESTADÍSTICAS

Este capítulo tiene como objetivo presentar las herramientas estadísticas que se utilizarán para encontrar la distribución posterior del precio de una opción. No se busca presentar toda la teoría existente ya en libros de texto porque no es la finalidad del capítulo, sin embargo se marcarán bibliografías para que aquel lector que quisiese ahondar en la teoría pueda hacerlo.

De esta manera este segundo capítulo se dividirá en tres secciones; la primera de éstas se refiere a la inferencia bayesiana, la segunda a la simulación estocástica y la tercera a las series de tiempo financieras.

Siendo pilar en la tesis la inferencia bayesiana, es necesario que dentro de la primera parte de este capítulo se le dedique un estudio a sus principales resultados. También se busca dar a entender con claridad cuál es la razón de utilizar este enfoque estadístico a lo largo del trabajo.

En la econometría y la estadística existen una serie de modelos que son cada vez más utilizados en áreas como la financiera. En la tercera parte del capítulo se presentarán los modelos más importantes de series de tiempo financieras conjuntamente con una breve explicación de las variables más importantes analizadas por estos tipos de modelos.

A lo largo del trabajo se utilizará un modelo de series de tiempo con inferencia bayesiana con la idea de poder muestrear distribuciones de probabilidad que si bien pudiera resultar imposible o muy complicado realizarlo mediante métodos analíticos, se puede superar esta adversidad utilizando esquemas de simulación estocástica, los llamados métodos MCMC (*Monte Carlo Markov Chain*)

Ya que la utilización de la simulación estocástica en la tesis resulta indispensable, en la segunda sección se presenta una breve explicación de los métodos MCMC, haciendo énfasis en los algoritmos *Gibbs Sampler* y *Metropolis-Hastings*, este último será el que permita muestrear las distribuciones posteriores de interés.

## 3.1. Inferencia bayesiana

### 3.1.1. Razonamiento bayesiano

En contraste con el enfoque clásico, Bayes invirtió el razonamiento común de la estadística y orientó su atención en el problema de inferir las probabilidades que  $P(A)$  toma en diversos valores, dado lo que ha sido observado en un muestreo realizado. Se le ha denominado a este concepto como probabilidad inversa debido justamente a su inversión con respecto del enfoque clásico.

De una manera más formal, en el enfoque clásico la inferencia estadística está idealizada para dirigir la atención a un conjunto de vector de datos hipotéticos  $y_1, y_2, \dots, y_j, \dots$  los cuales podrían haber sido generados por el modelo probabilístico  $p(y | \theta_0, \sigma_0^2)$  de una función de distribución como pudiera ser  $p(y | \theta_0, \sigma_0^2) \propto \sigma_0^{-n} \exp \left[ -\frac{1}{2\sigma_0^2} \sum_{t=1}^n (y_t - \theta_0)^2 \right]$ ,  $-\infty < y_t < \infty$  donde  $\theta_0$  y  $\sigma_0^2$  son hipotéticamente los valores ciertos de  $\theta$  y  $\sigma^2$ . Anadido a esto se seleccionan los estimadores  $\hat{\theta}(y)$  y  $\hat{\sigma}^2(y)$  quienes son funciones del vector de datos. Por cada vector de datos hipotéticos  $y_j$  son calculados los valores de  $\hat{\theta}(y_j)$  y  $\hat{\sigma}^2(y_j)$  y por tanto los conjuntos referentes son generados por  $\hat{\theta}(y)$  y  $\hat{\sigma}^2(y)$ . Las inferencias entonces son realizadas para comparar los valores de  $\hat{\theta}(y)$  y  $\hat{\sigma}^2(y)$  en realidad observadas con su distribución muestral generado por los conjuntos referentes. Prosiguiendo en realizar intervalos de confianza y pruebas de hipótesis para verificar la veracidad de los estimadores.

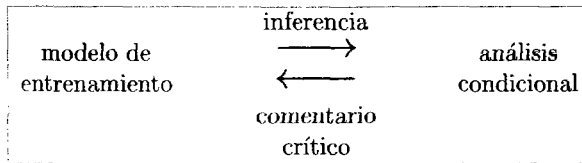
En contraste con este enfoque, la inferencia bayesiana introduce como parte del modelo una distribución **a priori**  $p(\theta, \sigma^2)$ . Ésta es supuesta para expresar un estado de conocimiento o ignorancia acerca de  $\theta$  y  $\sigma^2$  antes de que los datos sean obtenidos. Dado la distribución **a priori**, el modelo probabilístico  $p(y | \theta, \sigma^2)$  y los datos  $y$ , es ahora posible calcular la distribución de probabilidad  $p(\theta, \sigma^2 | y)$  de  $\theta$  y  $\sigma^2$ , dados los datos  $y$ . A esta distribución se le llama distribución **posterior** de  $\theta$  y  $\sigma^2$ . A partir de ésta se realizan todas las inferencias acerca de los parámetros.

El análisis bayesiano en la investigación científica toma una jerarquía significativa ya que como nunca se está seguro de que un modelo propuesto sea completamente apropiado, entonces se debe de proceder de tal manera que las partes inadecuadas del mismo puedan ser tomadas en cuenta y sus implicaciones puedan ser consideradas para que pueda seguir funcionando y ajustándose cada vez mejor. Para hacer esto se debe considerar un análisis estadístico como lo muestra el cuadro (2).

Este proceso usualmente comienza por un modelo que ya haya sido aceptado y que se está tentativamente entrenando. El trabajo multidisciplinario entre un investigador

científico y un estadístico debe de enfocarse en la elección apropiada de los parámetros que explican mejor al fenómeno para ser seguidos de la inferencia estadística acerca de éstos parámetros condicionales que tiene como fin la corrección del primer modelo tentativo. Éstas inferencias llevan el nombre de análisis condicional. Después de este proceso iterativo, si el modelo es correcto entonces proveerá todo lo que hay que saber acerca del problema bajo estudio dado el conjunto de datos disponibles.

Para una discusión más profunda, ver Box & Jenkins (1970).



**Cuadro 2:** Análisis estadístico de datos como un proceso iterativo de construcción de un modelo

### 3.1.2. El papel del análisis bayesiano

Las aplicaciones del teorema de Bayes son ejemplos de inferencia estadística. Aunque la inferencia es solamente una parte del análisis estadístico, el cual en turno es solamente una parte del diseño y del análisis, usada en la investigación iterativa, es una parte importante.

Se ha considerado en diferentes textos que las derivaciones del teorema de Bayes están apropiadamente relacionadas con el papel que se juega en la investigación científica debido a que:

- 1 Realiza una suposición precisa que se introduce en la parte izquierda del cuadro (2), y después por medio de un *filtro* se realizan las inferencias en la parte derecha de la misma figura.
- 2 A partir del punto anterior, se sigue que, dado el modelo, el análisis bayesiano hace uso automáticamente de toda la información a partir de los datos.
- 3 De hecho, de manera más profunda se sigue que las inferencias que no terminan siendo aceptadas deben de venir de suposiciones inapropiadas y no de suposiciones inadecuadas del sistema inferencial. Es decir que el análisis bayesiano está siempre expuesto al proceso de crítica del modelo y esto significa una retroalimentación como en el cuadro (2).

- 4 Una parte esencial en la inferencia bayesiana es que no surgen problemas como los que se encuentran en la teoría del muestreo, es decir, en la determinación de cuan grande tiene que ser el tamaño de la muestra para una buena elección de estimadores e intervalos de confianza.
- 5 El punto más importante es que la inferencia bayesiana provee una forma satisfactoria de introducir apropiadamente y mantenerse al tanto de las suposiciones acerca del conocimiento o ignorancia *a priori*.

### 3.1.3. Teorema de Bayes

Supóngase que  $y = \{y_1, \dots, y_n\}$  es un vector de  $n$  observaciones cuya distribución de probabilidad  $p(y | \theta)$  depende de los valores de  $k$  parámetros  $\theta = \{\theta_1, \dots, \theta_k\}$ . Supóngase también que  $\theta$  tiene por si mismo una función distribución  $p(\theta)$ . Entonces,

$$p(y|\theta) p(\theta) = p(y, \theta) = p(\theta|y) p(y) \quad (19)$$

Dado los datos observados  $y$ , la distribución condicional de  $\theta$  es

$$p(\theta|y) = \frac{p(y|\theta) p(\theta)}{p(y)} \quad (20)$$

Se puede escribir alternativamente la ecuación (20) como

$$p(\theta|y) \propto p(y|\theta) p(\theta)$$

o

$$p(\theta|y) = c p(y|\theta) p(\theta) \quad (21)$$

En la ecuación (21),  $p(\theta)$  es llamada distribución *a priori* de  $\theta$ , similarmente a  $p(\theta|y)$  se le llama distribución posterior de  $\theta$  dado  $y$ . La constante  $c$  es utilizada para que la distribución integre (o sume en caso de ser discreto) uno. Al utilizar  $p(y | \theta)$  se debe pensar como una función de  $\theta$  que se le denomina *función de verosimilitud*. De manera formal la función de verosimilitud queda expresada como  $\ell(\theta|y) = p(y|\theta)$ . La función de verosimilitud juega un papel muy importante en la inferencia bayesiana ya que es la función a través de la cual los datos  $y$  modifican el conocimiento *a priori* de  $\theta$ .

Con esta definición,  $p(\theta)$  como la distribución *a priori* para  $\theta$ , y  $p(\theta|y)$  como la distribución posterior para  $\theta$  dado  $y$  se puede expresar el teorema de Bayes como

**Distribución posterior  $\propto$  Distribución *a priori*  $\times$  Verosimilitud**

Historicamente, la elección de una densidad a priori para caracterizar una situación donde se conoce poco (en algunos casos, nada) de dicha densidad ha sido realmente extensa, y aún así, todavía es analizada y refutada. De hecho ha sido punto central de la discusión actual acerca del funcionamiento correcto de la inferencia bayesiana. En la inferencia bayesiana cuando se carece de todo conocimiento acerca de la distribución a priori y es supuesto que ésta se distribuye uniforme se le conoce bajo el nombre de **distribución a priori no informativa**.

Otra forma de afrontar el reto de suponer una distribución a priori, es mediante lo que se llama **análisis conjugado**. Por familia conjugada de distribuciones a priori se debe entender una familia de distribuciones a priori que, cuando son combinadas con la función de verosimilitud mediante el teorema de Bayes, resulta una distribución posterior que es de la misma familia paramétrica de distribuciones que la distribución a priori.

Es importante tener en mente que uno nunca se encuentra en un estado de completa ignorancia, además, la afirmación de un pequeño conocimiento a priori puede solamente tener un significado relacionado con la información proveniente de un experimento.

Ahora bien, una distribución a priori es supuesta para representar un conocimiento acerca de los parámetros antes de que los resultados de un proyecto experimental sean conocidos. Entonces, la principal cuestión es exactamente cómo seleccionar una distribución a priori la cual provea de información que tenga relación con el experimento deseado.

Este punto es crucial en la inferencia bayesiana ya que una mala elección de una distribución a priori desembocará en una distribución posterior errónea.

Hasta este momento se ha descrito la función posterior de distribución de un muestreo inicial de observaciones  $\{y_1, \dots, y_n\}$  que da el teorema de Bayes en la fórmula (21) sin embargo también es posible encontrar una distribución posterior utilizando una mayor cantidad de información.

Supóngase que se tiene una muestra inicial de observaciones  $y_1$ , y como ya se ha visto la fórmula de Bayes da  $p(\theta | y_1) \propto p(\theta) \ell(\theta | y_1)$ . Entonces ahora se debe suponer que se cuenta con una segunda muestra de observaciones  $y_2$  independientemente distribuidos de la primera muestra, entonces se tiene que

$$p(\theta | y_2, y_1) \propto p(\theta) \ell(\theta | y_1) \ell(\theta | y_2) \propto p(\theta | y_1) \ell(\theta | y_2) \quad (22)$$

La ecuación (22) está expresada precisamente de la forma de la ecuación (21) excepto

que  $p(\theta|y_1)$ , la distribución posterior de  $\theta$  dado  $y_1$ , juega el papel de la distribución a priori para la segunda muestra. Obviamente este proceso puede ser repetido cualquier número de veces. En particular si se tiene  $n$  observaciones independientes, la distribución posterior puede ser recalculada después de cada observación nueva, así que en la  $m$ -ésima entrada la verosimilitud asociada con  $m$  observaciones está combinada con la distribución posterior de  $\theta$  después de  $m - 1$  observaciones para dar una nueva distribución posterior

$$p(\theta|y_1, \dots, y_m) \propto p(\theta|y_1, \dots, y_{m-1}) \ell(\theta|y_m), \quad m = 2, \dots, n \quad (23)$$

Lo importante que se tiene que entender con el resultado dado en la fórmula (23) es que de manera general el teorema de Bayes describe el proceso de aprendizaje a partir de la experiencia, y muestra cómo el conocimiento acerca de la naturaleza del estado representado por  $\theta$  es continuamente modificado cuando se disponen de nuevos datos.

### 3.1.4. Distribución predictiva

Se define como distribución predictiva a la distribución marginal

$$p(X) = \int p(X|\theta) p(\theta) d\theta \quad (24)$$

El uso más importante de esta distribución es que permite verificar las suposiciones subyacentes.

### 3.1.5. Intervalos de credibilidad

En la inferencia bayesiana, los intervalos de credibilidad son la contraparte del concepto de los intervalos de confianza en el análisis estadístico clásico y se definen como sigue:

Un intervalo de credibilidad al  $100(1 - \alpha)\%$  para  $\theta$  es un subconjunto  $C$  de  $\Theta$  tal que

$$1 - \alpha \leq P(C|x) = \int_C dF^{\pi(\theta|x)} d\theta = \begin{cases} \int_C \pi(\theta|x) d\theta & \text{caso continuo} \\ \sum_{\theta \in C} \pi(\theta|x) & \text{caso discreto} \end{cases} \quad (25)$$

A partir de que la distribución posterior  $\pi$  es una probabilidad actual en  $\Theta$ , uno puede



hablar significativamente de la probabilidad de que  $\theta$  esté en  $C$ . Esto es justamente el contraste con los intervalos de confianza clásicos, los cuales pueden solamente ser interpretados en términos de probabilidad de cobertura, es decir, la probabilidad que una  $X$  aleatoria esté en tal intervalo de confianza  $C(X)$  que contiene a  $\theta$ .

Para estudiar las propiedades de los intervalos de credibilidad y para ver una discusión entre intervalos de credibilidad y de confianza, se recomienda ver Berger (1988) y Lee (1989).

### 3.1.6. Un esquema para el análisis bayesiano

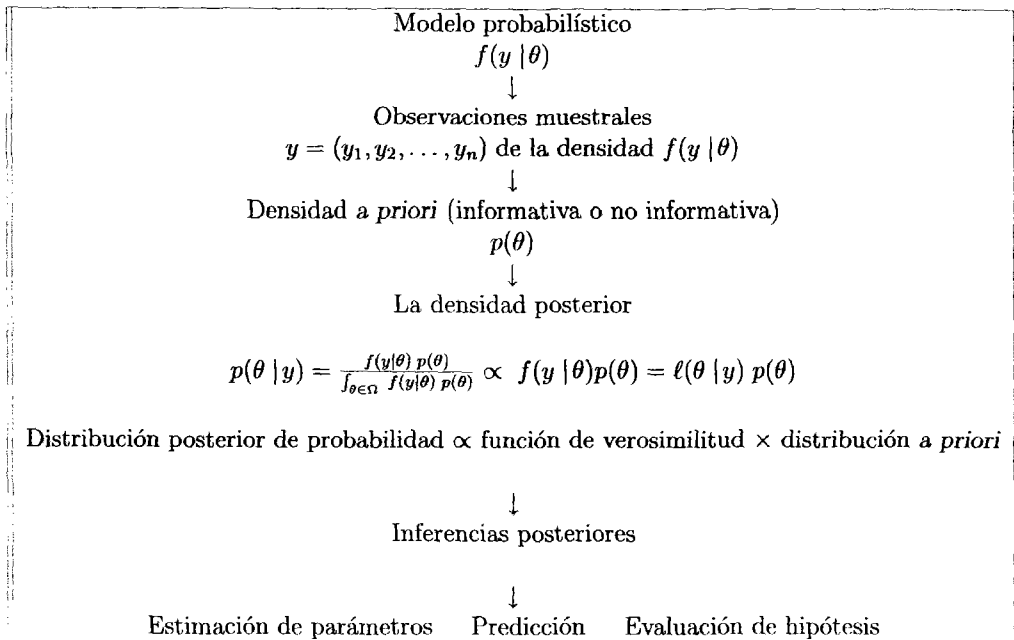
Como se ha ido explicando desde la primera sección de este capítulo, los componentes principales de la inferencia bayesiana consisten de los datos muestrales, la densidad a priori y posterior de los parámetros y la distribución predictiva de las observaciones externas a la muestra. Ahora, viendo estos conceptos bajo un punto de vista de datos y vectores de parámetros pueden ser escritos de la siguiente forma:

- 1 Información muestral  $y = (y_1, y_2, \dots, y_n)$  quien tiene una función de densidad de probabilidad conjunta  $f(y | \theta)$  y una función de verosimilitud asociada  $\ell(\theta | y), \theta \in \Theta$
- 2 Información a priori en forma de una densidad de probabilidad a priori  $p(\theta), \theta \in \Omega$  para el parámetro  $\theta$  en el modelo de probabilidad  $f(y | \theta)$ , y
- 3 La función de verosimilitud  $\ell(\theta | y)$  y la densidad a priori  $p(\theta)$  combinada por la ecuación (21) para producir la densidad posterior de  $\theta$ .

De la densidad posterior de  $\theta$ , la cual es el cimientto de la inferencia bayesiana, una base condicional *post-datos* para inferir acerca de  $\theta$  son resumidos en la forma de una distribución conjunta de probabilidad. A partir de los componentes principales, el enfoque bayesiano puede ser ilustrado como en el cuadro (3).

### 3.1.7. Un ejemplo de inferencia bayesiana

Lo que se pretende a continuación es presentar a manera de ejemplo un uso del análisis bayesiano aplicado a encontrar la distribución posterior de un modelo de regresión bajo los supuestos que se mencionan en la sección siguiente. La razón de utilizar este arquetipo es porque se acerca un poco más al uso de la inferencia bayesiana en el trabajo presente, el cual y como ya se ha mencionado con anterioridad es con base a



**Cuadro 3:** Diagrama para llevar a cabo un análisis bayesiano aplicado

encontrar la distribución de probabilidad posterior del precio de una opción mediante el modelo Black & Scholes, y antes de esto proponer una distribución posterior para la volatilidad utilizando inferencia bayesiana en un modelo de serie de tiempo.

*3.1.7.1. Regresión lineal bayesiana bajo un supuesto de normalidad y una distribución a priori no informativa*

Un modelo de regresión lineal múltiple queda determinado por la ecuación siguiente

$$Y = x \beta + \epsilon \tag{26}$$

donde

$$Y \sim N(x\beta, \sigma^2 I_n) \tag{27}$$

y

$$\epsilon \sim N(0, \sigma^2 I_n) \tag{28}$$

Usualmente se hace la suposición de que los parámetros  $\beta$  y  $\sigma$  son constantes fijas desconocidas.

Dado el modelo de regresión, se denota la función de densidad conjunta que abarca la muestra de observaciones  $y$  así como los valores de  $x$ ,  $\beta$  y  $\sigma$  por  $f(y, x, \beta, \sigma)$ . La función de verosimilitud correspondiente es entonces

$$\ell(\beta, \sigma | y, x) = (2\pi\sigma^2)^{-n/2} \exp \left[ \frac{-(y - x\beta)'(y - x\beta)}{2\sigma^2} \right] \quad (29)$$

Ahora debe considerarse que la información a priori es vaga; es decir no informativa. Para representar esta información vaga en los valores de  $\beta$  y  $\sigma$  en un análisis bayesiano se caracteriza la aleatoriedad de los vectores  $(\mathbf{B}, \Sigma)$  para especificar su distribución de probabilidad como

$$(\mathbf{B}, \Sigma) \sim p(\beta, \sigma) \equiv p(\beta)p(\sigma) \propto \frac{1}{\sigma}, \sigma \in (0, \infty) \text{ y } \beta \in \mathbb{R}^k \quad (30)$$

donde

$$p(\beta) \propto c \text{ y } p(\sigma) \propto \frac{1}{\sigma} \quad (31)$$

Teniendo ahora la función de verosimilitud y la información a priori se puede proceder en definir la distribución posterior de los parámetros en el modelo de regresión lineal (26), (27) y (28). Primero, dada la suposición de normalidad (28) y (30), la función de verosimilitud para los parámetros puede estar representada por

$$\begin{aligned} \ell(\beta, \sigma | y, x) &\propto \frac{1}{\sigma^n} \exp \left[ -\frac{1}{2\sigma^2} (y - x\beta)'(y - x\beta) \right] \\ &\propto \frac{1}{\sigma^n} \exp \left[ -\frac{1}{2\sigma^2} \left[ (n - k) \hat{\sigma}^2 + (\beta - \hat{b})' x' x (\beta - \hat{b}) \right] \right] \end{aligned} \quad (32)$$

donde  $\hat{\sigma}^2 = \frac{(y - x\hat{b})'(y - x\hat{b})}{(n - k)}$  y  $\hat{b} = (x'x)^{-1} x'y$ . Después de la combinación de la función de densidad a priori (30) y la función de verosimilitud (32) y usando el teorema de Bayes, la función de distribución conjunta para  $\beta$  y  $\Sigma$  queda definida como

$$p(\beta, \sigma | y, x) \propto \frac{1}{\sigma^{n+1}} \exp \left[ -\frac{1}{2\sigma^2} \left[ (n - k) \hat{\sigma}^2 + (\beta - \hat{b})' x' x (\beta - \hat{b}) \right] \right] \quad (33)$$

De la fórmula anterior y utilizando la definición de la función de densidad condicional, se sigue que la función de densidad posterior para  $\beta$ , dado  $\sigma$ , es una función de

distribución normal multivariada  $k$ -dimensional con media  $\mathbf{B}$  y covarianza  $\sigma^2 (\mathbf{x}' \mathbf{x})^{-1}$

$$p(\beta | \sigma, y, x) \propto \exp \left[ -\frac{(\beta - \hat{b})' x' x (\beta - \hat{b})}{2\sigma^2} \right] \quad (34)$$

Como ejemplo a lo anterior considérese un modelo estadístico simple como  $Y_i = \theta + \epsilon_i$ , donde  $\epsilon_i \sim iid N(0, 1)$  para  $i = 1, 2, \dots, n$ . Asumase también que la información a priori para  $\theta$  es no informativa y se usará el hecho, como ya se ejemplificó en la ecuación(31),  $p(\theta) \propto c$ . La función de verosimilitud está dada por

$$\ell(\theta | y) = (2\pi)^{-n/2} \exp \left\{ -\frac{1}{2} [(n-1) \hat{\sigma}^2 + n(\theta - \bar{y})^2] \right\}$$

donde  $\bar{y} = n^{-1} \sum_{i=1}^n y_i$  y  $\hat{\sigma}^2 = \sum_{i=1}^n (y_i - \bar{y})^2 / (n-1)$ .

Entonces la distribución posterior para  $\theta$  es proporcional a

$$p(\theta | y) \propto \exp(-n(\theta - \bar{y})^2 / 2)$$

La cual tiene la forma de una distribución normal con media posterior  $\bar{y}$ .

La figura (4) muestra una simulación de la distribución posterior para  $Y_i = \theta + \epsilon_i$ , donde  $\epsilon_i \sim iid N(0, 1)$ .

A manera de ejemplo se presenta en la figura (5) una comparación entre una distribución posterior del modelo estadístico  $y = \beta_1 + \beta_2 x_2 + \beta_3 x_3 + \epsilon_i$ , donde  $\epsilon_i \sim iid N(0, \sigma^2)$  para  $i = 1, \dots, n$  proveniente de una distribución a priori no informativa y otra proveniente de una distribución a priori informativa con respecto a  $\beta$  y  $\sigma$ .

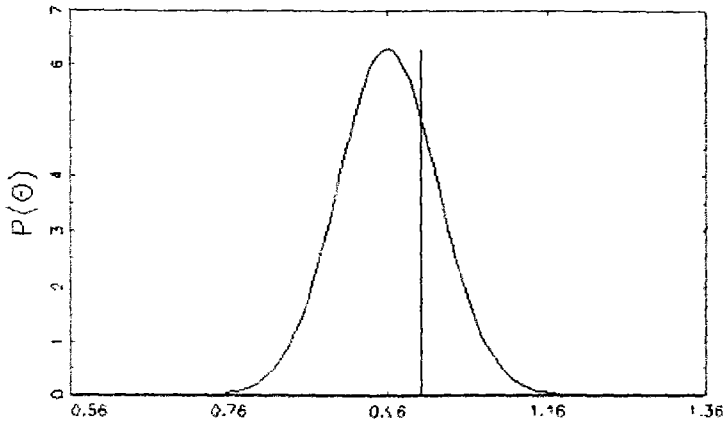
Se debe mencionar que el kernel de la distribución a priori para  $\beta$  es una normal multivariada con vector media posterior  $\mu$  y matriz de covarianzas  $\sigma^2 \psi$ , mientras que el kernel de la distribución a priori para  $\sigma$  es la raíz cuadrada invertida de una gamma; es decir  $Z^{-1/2}$  donde  $Z$  tiene una distribución Gamma.

Estas distribuciones a priori fueron obtenidas mediante

$$p(\beta, \sigma) = p(\beta | \sigma) p(\sigma)$$

y ésta a su vez es obtenida a partir de la distribución a priori conjunta

$$p(\beta, \sigma) \propto \sigma^{-m} \exp \left\{ -\frac{1}{2\sigma^2} \left[ \eta + (\beta - \mu)' \psi^{-1} (\beta - \mu) \right] \right\}$$



**Figura 4:** Distribución de probabilidad posterior de  $\theta$ . Tamaño de la muestra = 250. La línea vertical indica el valor real de  $\theta$ , el cual es 1

donde  $\eta > 0$  y  $\psi$  es simétrica definida positiva.

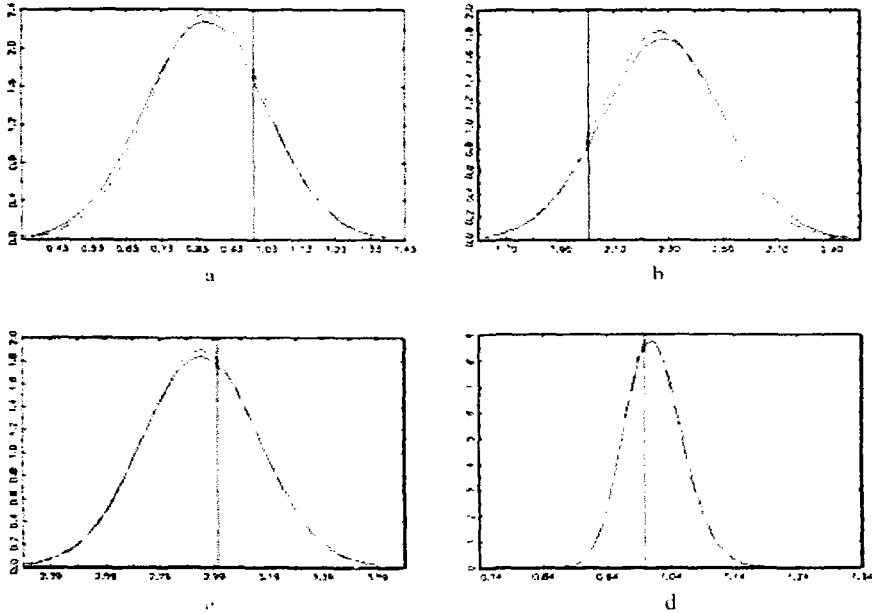
De hecho, esta familia de funciones de densidad *a priori* representan la familia conjugada de funciones de densidad *a priori* para la función de verosimilitud basadas en la distribución normal de la ecuación (32).

### 3.2. Simulación estocástica vía métodos MCMC

Las distribuciones *a priori* conjugadas están algunas veces relacionadas con una *a priori* cómoda en reconocimiento a su facilidad de integración, lo cual hace analíticamente conveniente al análisis posterior. Sin embargo existen ocasiones en donde no es posible encontrar las distribuciones posteriores de los parámetros en su forma analítica cerrada, entonces, es en estos casos cuando los esquemas de simulación estocástica como pueden ser los esquemas MCMC ayudan a realizar el análisis posterior correspondiente.

Los esquemas MCMC generan una gran cantidad de números aleatorios para representar la distribución posterior numéricamente. Sin embargo, cada reproducción MCMC es condicional al último efectuado, y cada réplica está correlacionada. Las propiedades estocásticas de las relaciones condicionales están descritas por una cadena de Markov, lo cual explica el origen del nombre de esta clase de algoritmo numérico.

De manera más formal, se necesita conocer la distribución  $p(\theta | X)$ , donde  $\theta \in \Theta$



**Figura 5:** Comparativo entre distribuciones posteriores. Las líneas continuas representan las distribuciones posteriores a partir de las *a priori* no informativas. Las líneas punteadas representan las distribuciones posteriores a partir de las *a priori* informativas. *a)*  $\beta_1$ , *b)*  $\beta_2$ , *c)*  $\beta_3$  y *d)*  $\sigma$ . La línea azul representa el valor real de los parámetros los cuales son 1, 2, 3 y 1, respectivamente. El tamaño de la muestra es igual a 2500

es el vector de parámetros y  $X$  son los datos. La idea de la simulación con cadenas de Markov es simular un proceso de Markov en  $\Theta$ , el cual converja a la distribución estacionaria  $p(\theta | X)$ . Entonces la clave de la simulación con cadenas de Markov es crear un proceso de Markov cuya distribución estacionaria sea  $p(\theta | X)$  y dejar que la simulación corra lo necesario para que los valores obtenidos sean lo suficientemente próximos a la distribución estacionaria. Para propósitos de este trabajo solo se mencionarán dos algoritmos, los cuales son justamente los esquemas de simulación estocástica que se utilizan con mayor frecuencia.

Así pues los métodos MCMC proporcionan una solución a los problemas difíciles de simulación a partir de distribuciones altamente dimensionales de las cantidades desconocidas que aparecen en modelos complejos.

En términos muy amplios, las cadenas de Markov son procesos estocásticos que describen trayectorias donde cantidades sucesivas son descritas probabilísticamente de acuerdo al valor de su antecesor inmediato. En algunos casos, esos procesos tienden a un equilibrio y las cantidades límites dan lugar a una distribución invariante. Las técnicas MCMC permiten simular a partir de una distribución para encajar éste como una distribución límite de una cadena de Markov y simular a partir de la cadena hasta que se aproxime a un equilibrio.

Antes de entender los algoritmos de simulación estocástica a través de cadenas de Markov es importante que el significado de un proceso estocástico y algunas propiedades de las cadenas de Markov queden bien entendidas.

### 3.2.1. Procesos Estocásticos

Un proceso estocástico se define como una colección de variables aleatorias  $\{x(t) \mid t \in T\}$  definidas sobre el mismo espacio, en donde para cada  $t \in T$  se tiene una variable aleatoria  $x(t) : \Omega \rightarrow \mathbb{R}$ . A  $t$  se le llama parámetro y casi siempre representa el tiempo. Entonces un proceso estocástico es un conjunto de variables aleatorias tal que para todo punto en el tiempo exista una variable aleatoria.

Existen procesos estocásticos para tiempo discreto y para tiempo continuo. Los más importantes en tiempo discreto son las cadenas de Markov, el proceso Poisson, proceso de renovación y modelos de colas. En tiempo continuo destacan las cadenas de Markov en tiempo continuo, el movimiento browniano o proceso de Wiener y las Martingalas.

Para conocer más acerca de estos tipos de procesos estocásticos, ver Ross (2000).

### 3.2.2. Cadenas de Markov

Una cadena de Markov es un proceso estocástico con parámetro discreto  $\{X_n : n \in \mathbb{N}\}$  donde dado el estado presente, los estados pasados y futuros son independientes. Esta propiedad puede ser más formalmente expresada como

$$\begin{aligned} P [\theta^{(n+1)} \in A \mid \theta^{(n)} = x, \theta^{(n-1)} \in A_{n-1}, \dots, \theta^{(0)} \in A_0] \\ = P [\theta^{(n+1)} \in A \mid \theta^{(n)} = x] \end{aligned} \quad (35)$$

para todos los conjuntos  $A_0, \dots, A_{n-1}, A \subset s$  y  $x \in s$ . La ecuación (35) se le conoce como propiedad Markoviana.

La importancia de la propiedad de Markov es que indica que la evolución del proceso

estocástico solo depende de la información más reciente, es decir, la memoria del proceso estocástico es de un estado.

En general, las probabilidades en (35) dependen de  $x$ ,  $A$  y  $n$ . Cuando la cadena de Markov no depende de  $n$ , se dice que la cadena es homogénea. En este caso, una función de transición ó kernel  $P(x, A)$  puede ser definido como:

- 1 Para todo  $x \in s$ ,  $P(x, \cdot)$  es una distribución de probabilidad sobre  $s$ ;
- 2 Para todo  $A \subset s$ , la función  $x \mapsto P(x, A)$  puede ser evaluada.

Es también útil cuando se trata de un espacio de estados discretos a identificar  $P(x, \{y\}) = P(x, y)$ . Esta función se le conoce con el nombre de probabilidad de transición y satisface:

- $P(x, y) \geq 0, \forall x, y \in s$ ;
- $\sum_{y \in s} P(x, y) = 1, \forall x \in s$  Como cualquier distribución de probabilidad

### 3.2.2.1. Distribución límite

Un problema fundamental en el contexto de simulación para las cadenas de Markov es el estudio del comportamiento asintótico de la cadena cuando  $n \rightarrow \infty$ , donde  $n$  es el número de pasos o iteraciones de la cadena. Una distribución  $\pi$  se dice que es una distribución estacionaria de una cadena con probabilidades de transición  $P(x, y)$  si

$$\sum_{x \in s} \pi(x) P(x, y) = \pi(y), \forall y \in s. \quad (36)$$

La ecuación (36) puede ser escrita en notación matricial como  $\pi = \pi P$ . Ahora bien, una vez que la cadena alcanza una iteración donde  $\pi$  es la distribución de la cadena, se retiene la misma en esta distribución para las iteraciones subsecuentes. A esta distribución se le conoce con el nombre de distribución de equilibrio o distribución invariante.

Se puede probar que si la distribución estacionaria  $\pi$  existe y  $\lim_{n \rightarrow \infty} P^n(x, y) = \pi(y)$  entonces, independientemente de la distribución inicial de la cadena,  $\pi^{(n)}$  se aproximará a  $\pi$ , cuando  $n \rightarrow \infty$ .

En este sentido, a la distribución se le conoce también con el nombre de distribución límite.



### 3.2.2.2. Simulación de una cadena de Markov

Ahora que ya se han presentado las propiedades fundamentales de una cadena de Markov es posible hablar más comodamente de la idea general de simular este tipo de proceso estocástico.

Se debe tomar una cadena de Markov  $(\theta^{(n)})_{n \geq 0}$  con espacio de estados  $s \subset \mathbb{R}^d$ , un kernel de transición  $P(x, y)$  y una distribución inicial  $\pi^{(0)}$  que como ya se ha mencionado, cumpla con la propiedad de que

$$\lim_{n \rightarrow \infty} \| P^n(x, \cdot) - \pi(\cdot) \| = 0, \forall x \in s$$

Una cadena de Markov que tiene esta característica recibe el nombre de ergódica.

La idea general es la generación de un valor de ésta cadena que comenzará con un valor para  $\theta^{(0)}$  muestreado a partir de  $\pi^{(0)}$ . El valor de  $\theta^{(1)}$  está entonces distribuido con densidad  $P(\theta^{(0)}, \cdot)$  y puede ser generado a partir de ésta. Para  $\theta^{(2)}$ , se repite el procedimiento para muestrear a partir de una distribución con densidad  $P(\theta^{(1)}, \cdot)$ . Al iterar este esquema a través de los pasos de la cadena se llega a muestrear  $\theta^{(n)}$  a partir de una distribución con una densidad  $P(\theta^{(n-1)}, \cdot)$ , para todo  $n$ .

Conforme el valor de  $n$  crece, se llega a estar cada vez más cerca de muestrear desde la distribución límite  $\pi$  y puede ser considerado como un muestreo de  $\pi$ . Es importante notar que todos los valores muestreados de la cadena son alcanzados después de converger y también son muestreados desde  $\pi$  debido a la estacionariedad de la cadena. Para conocer un mayor número de definiciones y propiedades de este tipo de proceso estocástico y explicaciones detalladas con respecto a las propiedades teóricas de una simulación de una cadena de Markov, ver Gamerman (1997).

### 3.2.3. Gibbs Sampler

El algoritmo Gibbs Sampler es uno de los más usados dentro de los esquemas MCMC y es una técnica para generar variables aleatorias indirectamente de una distribución (marginal) sin tener que calcular la densidad.

Para revisar los componentes del algoritmo, primero se debe de comprender que en la teoría de la probabilidad es posible que a partir de tener un conocimiento de las distribuciones condicionales se puede determinar la distribución conjunta, debido a esto considérese una distribución posterior  $p(\theta|y)$  para un vector de parámetros desconocidos. Primero represéntese un conjunto de distribuciones univariadas condicionales

para cada uno de los elementos de  $\theta$ , como

$$\begin{aligned}
 & p(\theta_1 | \theta_2, \theta_3, \dots, \theta_k, y) \\
 & p(\theta_2 | \theta_1, \theta_3, \dots, \theta_k, y) \\
 & \vdots \\
 & p(\theta_k | \theta_1, \theta_2, \dots, \theta_{k-1}, y)
 \end{aligned} \tag{37}$$

Bajo el algoritmo Gibbs Sampler, las distribuciones condicionales son usadas para generar secuencias de valores de los de parámetros aleatorios univariados, por cada uno de los elementos en  $\theta$ .

Ahora bien, dado un conjunto de distribuciones posteriores condicionales (37), el algoritmo Gibbs Sampler es iniciado mediante la especificación de un conjunto de valores para los parámetros  $\theta^0$ , que comienzan la cadena. Entonces, el siguiente conjunto de valores paramétricos son secuencialmente generados para muestrear los elementos de  $\theta$  a partir de las distribuciones condicionales

$$\begin{aligned}
 & p(\theta_1^1 | \theta_2^0, \theta_3^0, \dots, \theta_k^0, y) \\
 & p(\theta_2^1 | \theta_1^1, \theta_3^0, \dots, \theta_k^0, y) \\
 & \vdots \\
 & p(\theta_k^1 | \theta_1^1, \theta_2^1, \dots, \theta_{k-1}^1, y)
 \end{aligned} \tag{38}$$

Como ya se explicó en la sección anterior, la relación estocástica entre los muestreos subsiguientes está descrita por las probabilidades de transición markovianas de primer orden

$$\pi(\theta^{(m)}, \theta^{(m+1)}) = \prod_{j=1}^k p(\theta_j^{(m+1)} | \theta_m^{(h)} \text{ para } h > j, \theta_h^{(m+1)} \text{ para } h < j, y) \tag{39}$$

El cual es el producto de probabilidades condicionales en (37).

De igual forma como se mencionó en la sección pasada, se puede probar que cuando  $m \rightarrow \infty$ , el vector de resultados  $\theta^{(m)}$  converge en distribución a  $p(\theta | y)$ .

En todo momento se debe de tener en mente que la distribución de interés es  $\pi(\theta)$

donde  $\theta = (\theta_1, \theta_2, \dots, \theta_d)'$ , y también que las distribuciones condicionales totales  $\pi_i(\theta_i) = \pi(\theta_i | \theta_{-i})$ ,  $i = 1, \dots, d$  están disponibles. Esto significa que son completamente conocidas y se puede muestrear a partir de ellos.

El algoritmo Gibbs Sampler procede de la siguiente forma:

- 1 Inicializar el conteo de las iteraciones de la cadena con  $j = 1$  y el conjunto de valores iniciales  $\theta^{(0)} = (\theta_1^{(0)}, \dots, \theta_d^{(0)})'$ .
- 2 Obtener un nuevo valor  $\theta^{(j)} = (\theta_1^{(j)}, \dots, \theta_d^{(j)})'$  a partir de  $\theta^{(j-1)}$  a través de una generación sucesiva de valores

$$\begin{aligned}\theta_1^{(j)} &\sim \pi\left(\theta_1 \mid \theta_2^{(j-1)}, \dots, \theta_d^{(j-1)}\right) \\ \theta_2^{(j)} &\sim \pi\left(\theta_2 \mid \theta_1^{(j)}, \theta_3^{(j-1)}, \dots, \theta_d^{(j-1)}\right) \\ &\vdots \\ \theta_d^{(j)} &\sim \pi\left(\theta_d \mid \theta_1^{(j)}, \dots, \theta_{d-1}^{(j)}\right)\end{aligned}$$

- 3 Cambiar el contador de  $j$  a  $j + 1$  y regresar al paso 2 hasta que la cadena converja

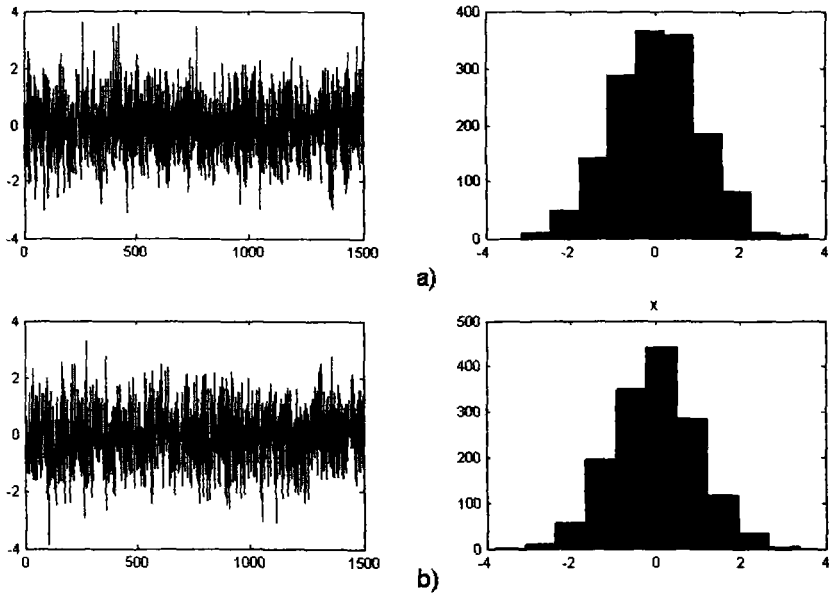
Cuando la cadena converja, el valor resultante  $\theta^{(j)}$  es un muestreo de  $\pi$ . Conforme el número de iteraciones se incrementa, la cadena se aproximará a su condición de equilibrio.

Para ver y estudiar con detalle la teoría existente del Gibbs Sampler, así como ejemplos de implementación, ver Casella (1992), Casella (1999) y Gamerman (1997).

Las gráficas (6) y (7) muestra un ejemplo de implementación Gibbs Sampler.

### 3.2.4. Metropolis-Hastings

Si el conjunto total de distribuciones condicionales (37) no está disponible, no es posible implementar el Gibbs Sampler. Por ejemplo, en modelos que no son lineales en los parámetros, entonces las distribuciones condicionales de los parámetros no se pueden conocer. En otros casos, la distribución podría ser conocida, pero no hay algoritmos eficientes para muestrear desde éste. Y es en tales casos donde podría ser empleado un esquema MCMC muy general que es conocido como Metropolis-Hastings.



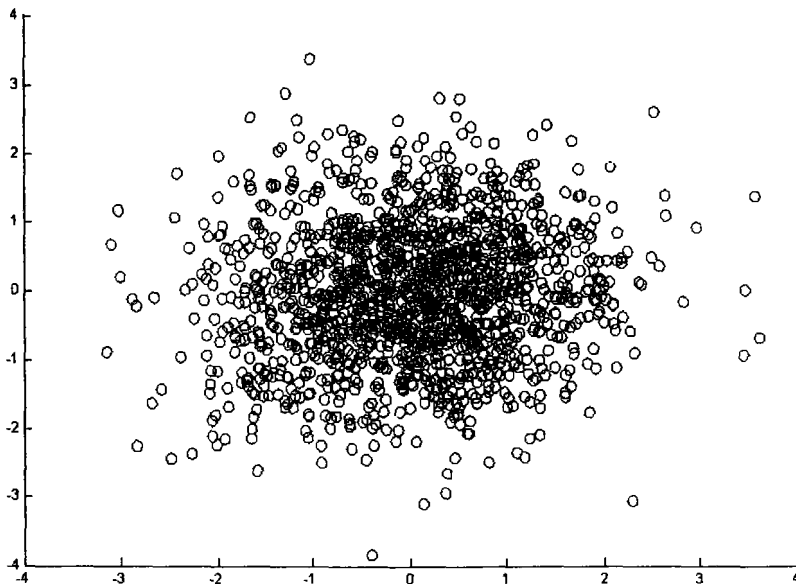
**Figura 6:** Implementación del Gibbs Sampler en una normal bivariada con medias  $(0,0.1)$  y varianzas  $(1,1)$  respectivamente y con una correlación de 0. a) Histograma posterior  $N(0.1,1)$  b) Histograma posterior  $N(0,1)$ .

Así como en el Gibbs Sampler, la idea básica debajo del Metropolis-Hastings es construir una cadena de Markov estacionaria que converja a  $p(\theta | y)$ . El componente principal del algoritmo es la *distribución instrumental, de salto o propuesta*  $q(\theta^{m+1} | \theta^m)$  a partir de la cual se genera un resultado de  $\theta^{m+1}$  condicional en  $\theta^m$ . Para propósitos prácticos, se debe seleccionar una distribución instrumental a partir de la cual se podrá generar convenientemente números aleatorios multivariados. Chib y Greenberg (1995) describen cinco alternativas.

Para asegurar que el algoritmo converja a  $p(\theta | y)$ , la cadena de Markov debe satisfacer la *condición de reversibilidad*

$$p(\theta^m | y) q(\theta^{m+1} | \theta^m) = p(\theta^{m+1} | y) q(\theta^m | \theta^{m+1}) \quad (40)$$

Donde el lado izquierdo es la probabilidad incondicional de generar  $\theta^{m+1}$  dado  $\theta^m$ , y el lado derecho es la probabilidad incondicional de generar  $\theta^m$  dado  $\theta^{m+1}$ . La condición



**Figura 7:** Scatterplot de las muestras de las normales de la figura (6).

de reversibilidad puede ser impuesta en (40) para inducir a un equilibrio en la ecuación

$$p(\theta^m | y) q(\theta^{m+1} | \theta^m) \alpha(\theta^m, \theta^{m+1}) = p(\theta^{m+1} | y) q(\theta^m | \theta^{m+1}) \quad (41)$$

Donde  $\alpha(\theta^m, \theta^{m+1})$  es conocida como la probabilidad de movimiento (*probability of move*) y está definida como

$$\alpha(\theta^m, \theta^{m+1}) = \min \left[ \frac{p(\theta^{m+1} | y) q(\theta^m | \theta^{m+1})}{p(\theta^m | y) q(\theta^{m+1} | \theta^m)}, 1 \right]$$

$$\text{si } \begin{cases} p(\theta^m | y) q(\theta^{m+1} | \theta^m) > 0 \\ \alpha(\theta^m, \theta^{m+1}) < 1 \end{cases} \quad (42)$$

Si el lado izquierdo en (41) es más grande que el derecho, entonces  $\alpha(\theta^m, \theta^{m+1}) < 1$  equilibra la probabilidad incondicional y la condición de reversibilidad es conocida.

De tal modo, el algoritmo Metropolis-Hastings queda determinado como sigue

- 1 Dado el valor de partida  $\theta^0$ , muestrear el instrumento aleatorio  $z$  a partir de  $q(\theta^1 | \theta^0)$  y  $u$  a partir de una distribución  $U(0, 1)$ .

- 2 Si  $u < \alpha(\theta^0, \theta^1)$ , fijar  $\theta^1 = z$ .
- 3 En otro caso, fijar  $\theta^1 = \theta^0$ .
- 4 Regresar al paso 1 y usar  $\theta^1$  para generar  $\theta^2$ .

Así como en el algoritmo Gibbs Sampler, el proceso se repite  $m$  veces, hasta que la cadena haya pasado del estado transitorio, a esta etapa de la implementación se le conoce con el nombre de etapa de calentamiento (*burn-in period*). Los resultados iniciales del proceso de calentamiento son descartados al momento de hacer inferencia estadística.

El algoritmo Metropolis-Hastings ofrece un número importante de aplicaciones que pueden ser analizadas puesto que no se necesita tener conocimiento de las densidades condicionales totales en su forma cerrada. Debido a esto, surgen tres puntos que es importante mencionarlos:

- 1 El Gibbs sampler es un caso especial del Metropolis-Hastings, donde la propuesta  $q(\theta^{m+1} | \theta^m) \propto \pi(\theta^{m+1})$  y a partir de (42) esto implica que la probabilidad de aceptación es siempre uno y el algoritmo siempre se mueve. Como el Gibbs sampler es un caso especial del Metropolis, uno puede diseñar algoritmos que consistan de pasos del Metropolis-Hastings o bien de Gibbs sampler, y estos comúnmente reciben el nombre de *algoritmos híbridos*.
- 2 El algoritmo Metropolis-Hastings permite que la forma funcional de la densidad no sea analítica, ejemplo de ello es cuando las funciones o modelos de activos financieros requieren de ecuaciones diferenciales parciales u ordinarias. Uno solamente tiene que evaluar la densidad real en dos puntos dados;
- 3 Hay una ventaja importante cuando existen problemas en el espacio paramétrico; Uno solo puede rechazar estos muestreos. Alternativamente, El muestreo puede ser realizado de manera condicional en regiones específicas, ver Gelfand (1992). Esto brinda un acercamiento conveniente al análisis de las restricciones de los parámetros impuestos por los modelos económicos.

Existen dos casos especiales más importantes de este algoritmo, el Metropolis-Hastings de cadenas independientes y Metropolis de caminata aleatoria. Sólo se explicará el primero debido a su utilización en el capítulo siguiente.

### 3.2.4.1. *Metropolis-Hastings independiente*

El algoritmo general Metropolis-Hastings muestrea  $\theta^{m+1}$  a partir de la propuesta,  $q(\theta^{(m+1)} | \theta^{(m)})$ , la cual depende del estado previo de la cadena de Markov  $\theta^{(m)}$ . Una alternativa es muestrear el candidato  $\theta^{(m+1)}$  desde una distribución independiente del estado previo de la cadena,  $q(\theta^{(m+1)} | \theta^{(m)}) = q(\theta^{(m+1)})$ . A esto se le conoce bajo el nombre del algoritmo de Metropolis-Hastings independiente, y queda determinado de la siguiente forma:

- 1 Muestrear  $\theta^{(m+1)}$  a partir de la propuesta independiente  $q(\theta^{(m+1)})$
- 2 Aceptar  $\theta^{(m+1)}$  con probabilidad  $\alpha(\theta^{(m)}, \theta^{(m+1)})$  donde

$$\alpha(\theta^{(m)}, \theta^{(m+1)}) = \min \left[ \frac{\pi(\theta^{(m+1)}) q(\theta^{(m)})}{\pi(\theta^{(m)}) q(\theta^{(m+1)})}, 1 \right]$$

Para revisar la información adicional acerca del algoritmo de Metropolis-Hastings, ver Casella (1999), Chib (1995), Geweke (1989) y Gamerman (1997).

Existen numerosas aplicaciones de simulación estocástica en diversas áreas. El área de interés particular en la tesis actual es la de la econometría financiera, ver Johannes (2003) para ejemplos de aplicación.

### 3.2.5. **Diagnóstico de convergencia**

Como ya se ha discutido con anterioridad, un valor a partir de la distribución de interés  $\pi$  es solamente obtenido cuando el número de iteraciones de la cadena tiende a infinito. En la práctica no es posible lograr esto, en vez de ello un valor obtenido a partir de una iteración suficientemente grande es tomado en vez de muestrearlo a partir de  $\pi$ . La dificultad es la determinación de cuan grande debería de ser esta iteración.

Hay dos formas de estudiar la convergencia. La primera es más teórica y trata de medir distancias y establecer límites en las funciones de distribución generadas a partir de la cadena, ver Tweedie (1994).

Sin embargo, el estudio de convergencia de la cadena también puede ser realizado a partir de una perspectiva estadística, es decir, analizando los resultados observados de la cadena, el problema de utilizar este tipo de métodos es que nunca se puede garantizar la convergencia debido a que solamente está basado en observaciones desde la misma cadena.

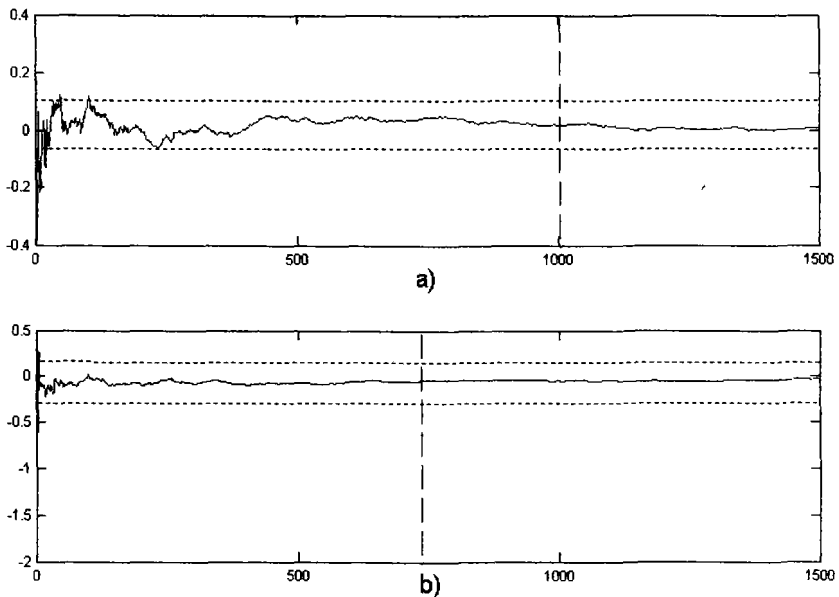
Aunque los dos métodos para estudiar la convergencia son validos y se complementan

el uno al otro, los resultados teóricos han probado ser más difíciles de obtener y para ser aplicados en problemas prácticos.

### 3.2.5.1. Monitores informales de convergencia

Gelfand y Smith (1990) sugirieron analizar la convergencia de manera un poco informal mediante técnicas gráficas. Después de  $m$  iteraciones en  $n$  cadenas paralelas, un histograma de los  $n$  valores de la  $m$ -ésima iteración de una función  $\theta$  dada puede ser graficada. Esta función puede ser uno de los componentes de  $\theta$  y el histograma podría ser suavizado si se desea. El procedimiento es repetido después de que  $k$  iteraciones son obtenidos en la cadena. El valor de  $k$  no necesita ser grande si uno sospecha acerca de la convergencia después de  $m$  iteraciones.

La figura (8) muestra un ejemplo de monitores informales de convergencia para el Gibbs Sampler de la figura (6).



**Figura 8:** Monitores informales de convergencia. a)Convergencia a la media de la distribución  $N(0.1,1)$ . b)Convergencia a la media de la distribución  $N(0,1)$ . El burn-in-period se compone con las muestras ubicadas a la izquierda de la línea vertical punteada.



### ***3.3. Series de Tiempo Financieras***

En las dos primeras secciones de este segundo capítulo se presentó a grandes rasgos los conceptos más importantes de la inferencia bayesiana, una vez hecho esto se prosiguió con un pequeño resumen de un par de esquemas de simulación MCMC que serán de gran ayuda en cuanto al análisis posterior que se realice cuando se quiera encontrar la distribución posterior de la volatilidad y su implementación en el modelo de Black & Scholes. El análisis posterior antes mencionado se hará en cuanto a un modelo de series de tiempo para volatilidad, es por esto que esta tercer sección está encaminada en la exposición de las series de tiempo financieras.

En esta sección se introducen algunos conceptos fundamentales que son necesarios para entender propiamente los modelos de series de tiempo. Se debe de comenzar con una simple introducción a los procesos estocásticos, la cual fue dada en la sección anterior y una definición clara de las series de tiempo conjuntamente con una introducción de las series de tiempo de interés particular en las finanzas.

#### **3.3.1. Definición de series de tiempo**

Las series de tiempo son un conjunto de observaciones  $x_t$  que son registrados en un tiempo específico  $t$ . Los modelos de series de tiempo proporcionan un método sofisticado de proyectar series históricas ya que se basan en la noción de que las series que se van a pronosticar se han generado por un proceso estocástico, con una estructura que puede caracterizarse y describirse. La descripción no se da en función de una relación causa y efecto (como sería en el caso de un modelo de regresión) sino en función de cómo está incorporada la aleatoriedad en el proceso.

Los modelos de series de tiempo que se explicarán a lo largo de ésta sección se basan en la suposición de que las series que se van a pronosticar se han generado por un proceso estocástico; es decir, que cada valor  $x_1, x_2, \dots, x_t$  en la serie de tiempo es extraído al azar de una distribución de probabilidad. De hecho, si se quisiera generalizar este detalle, se debería decir que la serie observada  $x_1, x_2, \dots, x_t$  es extraída de un conjunto de variables aleatorias distribuidas en forma conjunta. Una de las características de mayor relevancia es que en las series de tiempo no se cumple el supuesto de independencia.

Las series de tiempo se catalogan en discretos y continuos. Esta designación se debe a como se encuentra el conjunto de datos; es decir, si el conjunto de datos es muestreado

en intervalos fijos de tiempo se le denomina *serie de tiempo discreta*, o si es muestreado en intervalos de tiempo continuas se le denomina *serie de tiempo continua*. Esta diferencia marca una enorme pauta en el trabajo actual y se comentará más adelante cuando se relacione la modelación por medio de series de tiempo con el modelo de Black & Scholes.

El análisis de series de tiempo ha sido usado en diversos quehaceres de la vida humana, ejemplos que van desde aspectos puramente teóricos hasta totalmente prácticos. Dentro de la práctica se ha utilizado en diversas áreas del conocimiento como son la física, la meteorología, la industria en general y no puede quedar atrás el área financiera, en donde se ha vuelto herramienta indispensable cuando se busca evaluar de manera más precisa cualquier tipo de activo financiero, hasta en los más modernos productos financieros derivados, como es el caso actual.

El análisis de series de tiempo financieras está enfocado a lo que en el argot financiero se le conoce como datos de altas frecuencias, ver Dacorogna (2001). Algunos de las series de tiempo financieras que son analizados en las finanzas de altas frecuencias son el precio, los retornos, la volatilidad realizada (histórica), la propagación (*spread*) y la volatilidad (estocástica). Sin embargo durante el presente trabajo solo se discutirán las series de tiempo financieras correspondientes al precio, los retornos y la volatilidad estocástica.

### 3.3.2. Precio

Los precios de los activos son las variables más importantes exploradas en las finanzas. Dependiendo de la estructura del mercado y el distribuidor de datos, los precios están disponibles en diferentes formas:

- Precio par *bid-ask*:  $P_{bid}$  y  $P_{ask}$ . Un dato individual de este tipo en un momento particular del tiempo se le conoce como *tick*.
- Precios de transacción.
- Bid, ask y precios de transacción en secuencias irregulares. (No en pares)
- Precios medios.

### 3.3.3. Retornos

El retorno debe ser definido como la utilidad generada sobre una inversión de capital o sobre una inversión en valores. Es por esto que en cuanto al inversionista

corresponde es más útil analizar el retorno de una inversión que su precio nominal. A parte de que estadísticamente la serie de retornos tiene propiedades más interesantes para el modelaje econométrico, como por ejemplo, de manera general las series de retornos son estacionarias, en comparación con la serie de precios que no lo son. Además, la distribución de los retornos es más simétrica y estable a lo largo del tiempo que la distribución de los precios.

Financieramente hablando existen dos tipos más generales de retornos: el retorno simple y el compuesto. Se definen a continuación:

Sea  $P_t$  el precio de un activo financiero al tiempo  $t$ .

El retorno simple se define de la siguiente manera:

$$r_t = \frac{P_t - P_{t-1}}{P_{t-1}}$$

y el retorno compuesto o Log-retorno:

$$r_t = \ln \left( \frac{P_t - P_{t-1}}{P_{t-1}} \right) = \ln \left( \frac{P_t}{P_{t-1}} \right)$$

El Log-retorno es comúnmente más utilizado que el retorno simple debido a que el origen de esta fórmula tiene que ver con el interés continuo, el cual es utilizado en muchos ámbitos financieros como el tipo de interés que opera diversos instrumentos, en este caso el rendimiento de las acciones.

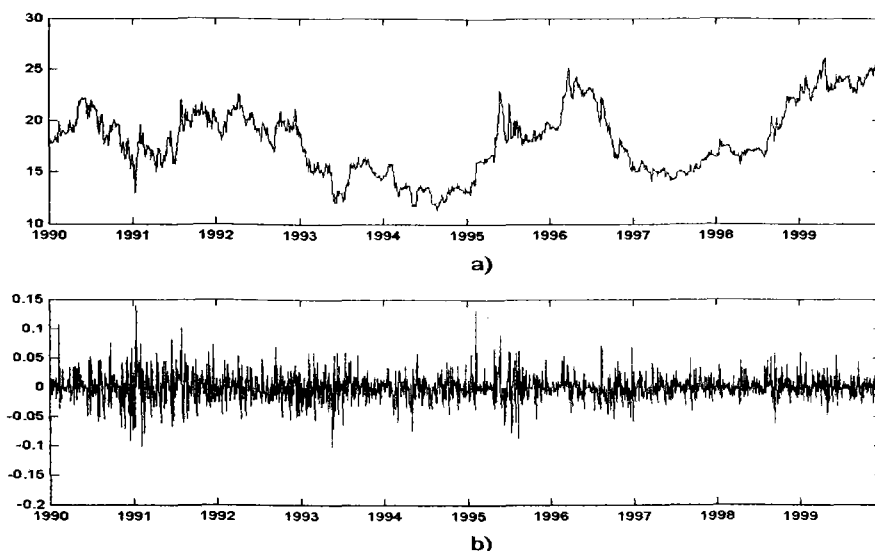
Para ver las distribuciones usadas comúnmente para los retornos, ver Tsay (2002).

Estadísticamente la serie de retornos cumple con las siguientes propiedades, un ejemplo de una serie de retornos se observa en la figura (9).

- **Curtosis.** Hay un exceso de curtosis en comparación con la distribución normal, por lo tanto, la serie de retornos es de colas pesadas.
- **Autocorrelación.** No hay una autocorrelación significativa, sin embargo, el cuadrado de dichos retornos si la presentan.
- **Heterocedasticidad.** La varianza evoluciona a través del tiempo.

### 3.3.4. Volatilidad

Este concepto ha sido discutido brevemente en el primer capítulo. En él se mencionó el error que existe al tratar de definir la volatilidad. Sin más preámbulos solo



**Figura 9:** a) Serie histórica 1990-1999 Bimbo A. b) Serie de retornos

debe uno fijarse en el modelo de Black & Scholes para percibir como está tradicionalmente definida la volatilidad; es decir con la desviación estándar, la cual sería correcta siempre y cuando la volatilidad no evolucionara a través del tiempo.

Es por ello que la volatilidad juega un papel muy importante en el mercado de derivados, en especial en el de opciones. Sin embargo el estudio de la volatilidad es necesario en otras áreas como la administración de riesgos, en donde se pueden plantear modelos de volatilidad para aproximar el valor en riesgo de una posición financiera, o en macroeconomía en donde se podrían plantear políticas económicas más adecuadas por medio de la volatilidad en los agregados macroeconómicos permitiendo así una mayor estabilidad económica y podría darse el caso de anunciar un posible colapso económico antes de que éste ocurra.

En conclusión, la modelación de la volatilidad por medio de series de tiempo puede mejorar la eficiencia en la estimación de los parámetros y la exactitud en los intervalos de pronóstico que sirvan como punto de partida para una mejora en la toma de decisiones ya sean tanto financieras como económicas.

Siendo la volatilidad indirectamente observable no existe una forma directa de medirla. Sin embargo una definición más acorde con el planteamiento del problema

que da origen al presente trabajo es considerar a la volatilidad como la *varianza condicional de los retornos de un activo*.

Más adelante se estudiarán los distintos modelos de series de tiempo que se utilizan en el estudio de la volatilidad, sin embargo la mayoría de ellos se basan en lo siguiente:

Sea  $y_t$  la serie de Log-retornos, entonces la media condicional está dada por  $\mu_t = \mathbb{E}(y_t | F_{t-1})$  y la varianza condicional por  $h_t = \sigma^2(y_t | F_{t-1}) = \mathbb{E}[(y_t - \mu)^2 | F_{t-1}]$ , donde  $F_{t-1}$  denota toda la información disponible hasta el tiempo  $t - 1$ .

Algunas de las características de la volatilidad que comúnmente se encuentran en los Log-retornos son:

- La volatilidad evoluciona a través del tiempo y de manera continua.
- Se encuentran períodos de alta volatilidad y períodos de baja volatilidad, a estos períodos se les denomina *clusters*.
- La volatilidad no diverge a infinito, varía dentro de un rango fijo; es decir que la volatilidad es estacionaria.
- La volatilidad parece reaccionar de manera diferente a una alza en los precios que a una caída de estos.

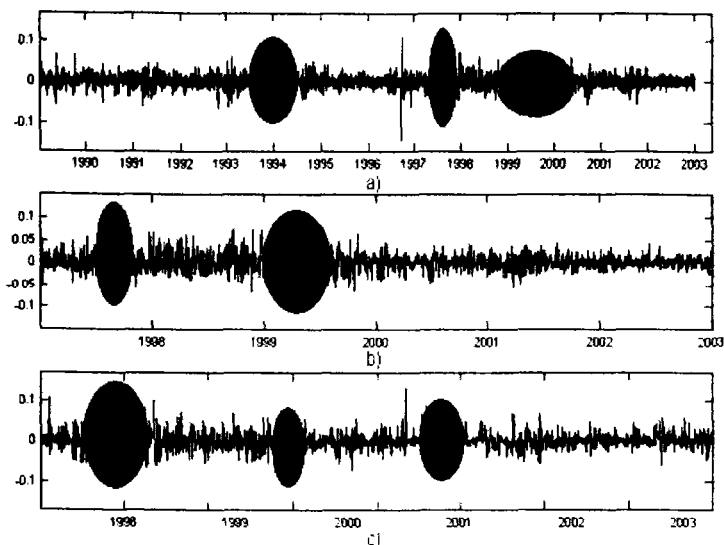
Un ejemplo de la existencia de clusters de volatilidad se observa en la figura (10). En ella se muestran los clusters de mayor volatilidad (área dibujada con rojo), en estos espacios la transición entre los períodos de volatilidad se realiza de manera gradual y continua.

### 3.3.5. Modelos de series de tiempo

Existen dos tipos de modelos de series de tiempo, los modelos determinísticos y estocásticos.

Dentro de los modelos determinísticos encontramos dos ramas, una que se basa en modelar con respecto a la media del proceso y la otra con respecto a la varianza del mismo.

Los modelos que se explicarán a continuación son modelos de series de tiempo financieras y su utilización es indispensable cuando se habla de datos financieros de altas frecuencias, es decir cuando la obtención de la información histórica se obtiene de manera diaria, en horas, de hecho podrían ser en segundos, como se espera que



**Figura 10:** Clusters de mayor volatilidad para las series de retornos a)IPC, b)TelmexL y c) BimboA

ocurra en un mercado bursátil de alto funcionamiento.

Por la naturaleza por como están contruidos estos modelos no pueden ser utilizados cuando la obtención de la información solo está disponible mensualmente, trimestralmente o hasta anualmente, como ocurre con la mayoría de los agregados macroeconómicos. Cuando esto ocurre es recomendable utilizar los llamados modelos de series de tiempo temporales, una bibliografía que sirve como punto de partida para el estudio de estos modelos de series de tiempo es Ghysels(2001).

A continuación se dará una breve descripción con respecto a los modelos determinísticos de series de tiempo financieras.

Se hará mucho más énfasis en los modelos con respecto a la volatilidad ya que son justamente este tipo de modelos sobre los que se practicarán los esquemas MCMC.

### 3.3.6. Modelos de series de tiempo para la media

#### 3.3.6.1. *AR (Autoregressive)*

En el proceso autorregresivo de orden  $p$ ,  $AR(p)$ , la observación actual  $y_t$  es generada por un promedio ponderado de observaciones pasadas que se remontan  $p$  períodos,

junto con una perturbación aleatoria en el período actual. El proceso  $AR(p)$  se define como

$$y_t = \phi_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \varepsilon_t \quad (43)$$

En el modelo  $AR(p)$  (y también en los demás modelos de series de tiempo que se verán a continuación) se supone que las perturbaciones aleatorias ( $\varepsilon_t$ ) están distribuidas en forma independiente a lo largo del tiempo; es decir, son generados por un proceso de ruido blanco. En particular, se asume que cada término de perturbación es una variable aleatoria normal con media 0, varianza  $\sigma_\varepsilon^2$  y covarianza  $\gamma_k = 0$  para  $k \neq 0$ .

### 3.3.6.2. *MA (Moving Average)*

En el proceso de media móvil de orden  $q$ ,  $MA(q)$ , cada observación  $y_t$  es generada por un proceso ponderado de perturbaciones aleatorias que se remontan  $q$  períodos. Queda definido como

$$y_t = c_0 + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_q \varepsilon_{t-q} \quad (44)$$

Los procesos AR y MA deben satisfacer ciertas restricciones en los parámetros para garantizar estacionariedad, ver Chatfield (1989).

### 3.3.6.3. *ARMA (Autoregressive Moving Average)*

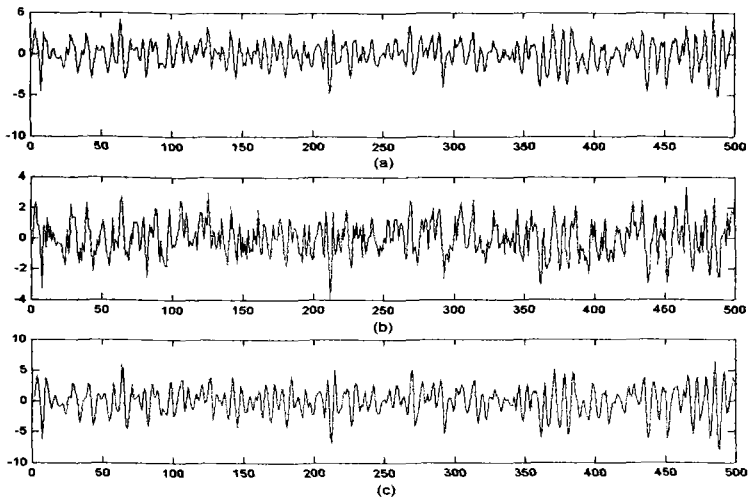
Muchos procesos estocásticos estacionarios no pueden modelarse como promedios móviles o autorregresivos puros, ya que tienen las cualidades de ambos procesos. La extensión lógica de los dos modelos presentados anteriormente es el proceso mixto autorregresivo-promedio móvil  $ARMA(p, q)$  y se representa como

$$y_t = \phi_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_q \varepsilon_{t-q} \quad (45)$$

La figura (11) muestra una simulación de un proceso autorregresivo, un proceso de media móvil y el proceso mixto ARMA.

### 3.3.6.4. *ARIMA (Autoregressive Integrated Moving Average)*

En la práctica muchas de las series de tiempo que se tienen no son estacionarias, de modo que las características del proceso estocástico subyacente cambian con el tiempo.



**Figura 11:** Simulación de procesos a) AR(2) con parámetros  $\phi_1 = 0.9$  y  $\phi_2 = -0.7$ , b) MA(2) con parámetros  $\varepsilon_1 = 0.6$  y  $\varepsilon_2 = -0.3$  y c) ARMA(2,2) con los parámetros anteriores

Se dice que  $y_t$  es *no estacionaria homogénea de orden  $d$*  si  $w_t = \Delta^d y_t$ , es una serie estacionaria. Aquí  $\Delta$  denota diferenciación; es decir,  $\Delta y_t = y_t - y_{t-1}$ ,  $\Delta^2 y_t = \Delta y_t - \Delta y_{t-1}$  y así sucesivamente.

Entonces, después de haber diferenciado la serie  $y_t$  para producir la serie estacionaria  $w_t$ , se puede modelar  $w_t$  como un proceso ARMA. Si  $w_t = \Delta^d y_t$  y  $w_t$  es un proceso  $ARMA(p, q)$ , entonces se dice que  $y_t$  es un *proceso autorregresivo integrado de promedio móvil de orden  $(p, d, q)$* , o tan solo  $ARIMA(p, d, q)$ .

Cuando  $p = 0$ ,  $d = 1$  y  $q = 1$ , el modelo se llama  $IMA(1, 1)$ .

De igual manera cuando  $p = 1$ ,  $d = 1$  y  $q = 0$  se le denomina  $ARI(1, 1)$ .

Para revisar más detalles acerca de la teoría y propiedades de toda esta gama de modelos con respecto a la media del proceso, a parte de las bibliografías ya mencionadas, ver Diggle (1990) y Wei (1990).



### 3.3.7. Modelos de series de tiempo para la volatilidad

#### 3.3.7.1. ARCH (Autoregressive Conditional Heteroscedastic)

Lo primero que se debe mencionar es que este modelo es con el cual se trabajarán los esquemas MCMC a lo largo del tercer capítulo.

Los modelos de series de tiempo han sido inicialmente introducidos para propósitos descriptivos como predicción y corrección estacional ó para el control dinámico, ver Gouriéroux (1997). A principio de la década de los setenta, el rol de investigación en cuanto a los modelos de series de tiempo se refiere se enfocó en una clase específica de modelos, los modelos ARMA, los cuales eran muy fácil de implementarlos. Sin embargo, esta propuesta tiene dos desventajas:

- 1 Es esencialmente una situación lineal, la cual automáticamente restringe el tipo de dinámica que es aproximada;
- 2 Es generalmente aplicado sin la imposición de una restricción a priori en los parámetros autorregresivos o de promedio móvil, lo cual acarrea problemas estructurales.

Las series de tiempo financieras presentan diversas formas de dinámica no-lineal. Sin embargo, las teorías financieras basadas en el comportamiento racional del inversionista naturalmente sugerirían el incluir y el probar algunas restricciones en los parámetros. En la teoría económica los agentes económicos no responden solamente a la media y en la práctica financiera frecuentemente se sugiere que tanto la varianza como la media de la tasa de retornos son determinantes para las decisiones de inversión. Existen diversas investigaciones al respecto, por ejemplo el Premio Nóbel de economía, Milton Friedman, argumentó que como una alta inflación estará generalmente asociada con una variabilidad alta de la misma variable, las relaciones estadísticas entre inflación y desempleo deberían tener una pendiente positiva, no una negativa como en la tradicional curva de Phillips. Para más información al respecto, ver Friedman (1977).

En este contexto los modelos ARCH, introducidos por Engle (1982), surgen como una estructura apropiada para estudiar este tipo de problemas.

A la par con el desarrollo de los modelos ARCH, dentro de la teoría financiera se ha ido dando una evolución o más bien una reexaminación o refinamiento de la mayoría de los modelos estructurales. De hecho el enorme éxito de los modelos ARCH también se debe a los errores de algunas prácticas basadas en análisis estático de fenómenos financieros.

No obstante el gran reconocimiento, existe una lista de limitaciones tanto teóricas como prácticas en cuanto a la modelación ARCH, para más detalles ver Gouriéroux (1997).

El modelo ARCH, como ya se había mencionado, fue propuesto por Engle en 1982 y lo usó para estimar las medias y las varianzas de la inflación en el Reino Unido durante la década de los setentas.

La idea básica de los procesos ARCH es que los retornos de una serie de tiempo  $y_t$  tienen media cero, carecen de una dependencia serial pero son dependientes y pueden representarse mediante una función cuadrática. Es así por lo que un modelo  $ARCH(r)$  se especifica de la siguiente manera

$$y_t = \varepsilon_t \sqrt{h_t}$$

$$h_t = \alpha_0 + \alpha_1 y_{t-1}^2 + \dots + \alpha_r y_{t-r}^2 \quad (46)$$

donde  $\varepsilon_t$  es una sucesión de variables aleatorias independientes e idénticamente distribuidas (valid) con media cero y varianza uno. Para  $\alpha_0 > 0$  y  $\alpha_i > 0 \forall i > 0$ . De acuerdo a la definición de  $h_t$  es como se logra que la varianza no sea constante a través del tiempo.

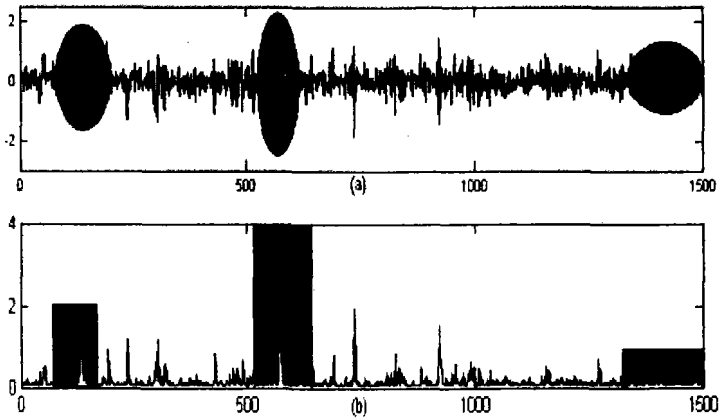
Para observar a detalle la especificación del modelo y sus parámetros, ver Engle (1982).

En la figura (12) se presenta la simulación de un proceso ARCH y la varianza condicional  $h_t$ . Los óvalos dibujados con rojo representan dos de las etapas donde los retornos alcanzaron mayor desempeño, estas etapas se ven acompañadas de épocas de mayor volatilidad (rectángulos en rojo). Las áreas dibujadas en verde representan a su vez donde la volatilidad fue pequeña.

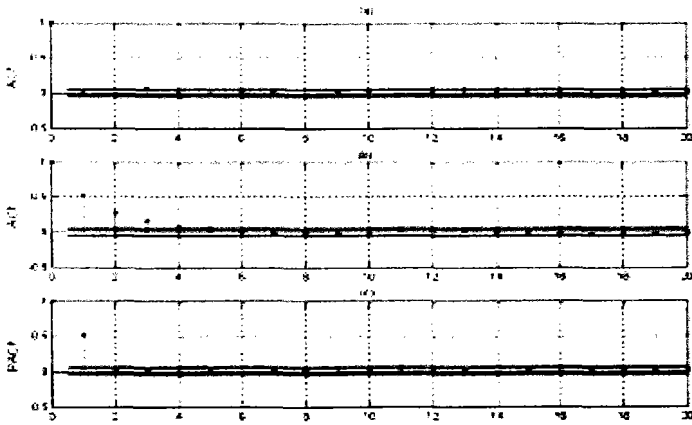
Se debe de observar también que los retornos grandes se ven seguidos de retornos grandes, y viceversa. Esto ejemplifica la existencia de los clusters de volatilidad.

En la figura (13) se muestran las funciones de autocorrelación de los retornos del proceso simulado ARCH(2) y el cuadrado de los mismos, aunado con la función de correlación parcial de éste último. De hecho en la misma gráfica se aprecia otra propiedad de los retornos, ya que al parecer no hay correlación significativa en la serie, sin embargo, los cuadrados de los retornos si presentan correlación.

Propiedades de los modelos ARCH, ver Tsay(2002).



**Figura 12:** a) Simulación de un proceso  $ARCH(2)$  con  $\alpha_0 = 1$ ,  $\alpha_1=0.03$  y  $\alpha_2=0.05$ .  
 b) Varianza condicional  $h_t$  de la serie simulada



**Figura 13:** a) Función de autocorrelación de la serie de retornos simulada en la figura (12). b) Función de autocorrelación de los cuadrados de los retornos. c) Función de autocorrelación parcial de los cuadrados de los retornos

- El modelo trata los retornos positivos y negativos de igual manera. En la práctica, se sabe que para las series de tiempo financieras los precios reaccionan de manera diferente a retornos positivos que a negativos.
- El modelo es restrictivo en los parámetros.

- El modelo no provee un modo de entender el comportamiento de la serie de tiempo financiera. Sólo provee una manera mecánica de describir la varianza condicional.
- El modelo puede requerir muchos parámetros para describir de manera adecuada la volatilidad de la serie.

Así como se han presentado las riquezas de los modelos ARCH, existen una serie de debilidades, para un análisis detallado ver Tsay (2002).

### 3.3.7.2. *GARCH (Generalized Autoregressive Conditional Heteroscedastic)*

El proceso GARCH fue propuesto por Bollerslev (1986) para permitir una mayor flexibilidad en los procesos ARCH. Como ya fue descrito, los modelos ARCH están basados en una representación autorregresiva de la varianza condicional. Sin embargo se podría anexar una forma de promedio móvil, esto se hace para disminuir la parsimonia del modelo ARCH. Esto es lo que hace el proceso GARCH. Entonces el modelo  $GARCH(\tau, s)$  está definido por

$$\begin{aligned} \varepsilon_t &= \sigma_t \epsilon_t \\ \sigma_t^2 &= \alpha_0 + \sum_{i=1}^m \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^s \beta_j \sigma_{t-j}^2 \end{aligned} \quad (47)$$

donde  $\epsilon_t$  es una sucesión de vaaid con media cero y varianza uno, sus parámetros deben de ser  $\alpha_0 > 0$ ,  $\alpha_i > 0$ ,  $\beta_j > 0$  y  $\sigma_{i=1}^{max(m,s)} (\alpha_i + \beta_i) < 1$ , esta es la condición que garantiza que la varianza incondicional sea finita.

Para revisar propiedades y la teoría de los modelos GARCH, ver Bollerslev (1986) y Tsay (2002).

### 3.3.7.3. *ARCH-M (ARCH in MEAN)*

Fue propuesto por Engle, Lilien y Robins (1987). El modelo ARCH es extendido para permitir a la varianza condicional afectar a la media de la  $Y_T$ . De esta forma los cambios en la varianza condicional afectarán directamente el valor esperado de un retorno.

### 3.3.7.4. *IGARCH (Integrated GARCH)*

Al igual que el modelo ARIMA propone una forma de modelar la media de una serie no estacionaria, este modelo propuesto por Bollerslev (1992) es realizado bajo

el hecho de que no todos los retornos se modelan de la mejor manera con un modelo estacionario. En finanzas es complicado pensar en utilizar un modelo IGARCH ya que su consecuencia es el hecho de la varianza incondicional no es finita.

#### 3.3.7.5. *HARCH (Heterogeneous ARCH)*

Este modelo fue propuesto por Müller (1997) para modelar los cuadrados de los retornos, con un buen manejo analítico. Los procesos HARCH corresponden a la familia de los ARCH pero difieren de todos los otros tipos de procesos ya que poseen la única propiedad de considerar la medición de la volatilidad sobre diferentes tamaños de intervalos.

Otros tipos de modelos determinísticos para volatilidad son EGARCH, FIGARCH, GARCH-M, FACTOR-ARCH, FACTOR-GARCH, AGARCH, VGARCH, GJR, NGARCH, QARCH, EMA-HARCH, TARCH, ARMA-GARCH etc.

Para más información acerca de estas generalizaciones del ARCH, ver Engle (1995), Gouriéroux (1997) y Mills (1999).

#### 3.3.8. Proceso ARCH y el modelo Black & Scholes

A lo largo del primer capítulo se presentó el modelo de Black & Scholes, así como en este segundo capítulo se ha ostentado la inferencia bayesiana seguido de los esquemas de simulación estocástica y consumado con una breve exposición de las series de tiempo financieras, haciendo mayor énfasis en el modelo ARCH ya que es éste con el cual se trabajará en el tercer capítulo de la tesis.

En el primer capítulo se distinguió que la volatilidad representaba un factor de suma importancia en lo que a la valuación de opciones se refiere, en el modelo de Black & Scholes este parámetro es constante, se ha discutido ya en demasía acerca de la veracidad de este argumento, y lo que se propone en esta tesis es ajustar un modelo de series de tiempo para volatilidad que sea capaz de solventar esta disyuntiva.

Sin embargo no es tan sencillo implementar un modelo discreto en el modelo de Black & Scholes por dificultades teóricas. Estas dificultades empiezan porque el modelo de Black & Scholes es un claro ejemplo de un modelo continuo, ver Hull (1997). Y con respecto a esto se debería de usar un modelo de este tipo para la volatilidad. Sin embargo, la profundización teórica en este tipo de modelos hace a su estimación muy complicada, ya que los datos reales ocurren en intervalos de tiempo discretos. Se ha intentado hacer inferencia indirecta ajustando un modelo ARCH para deducir un modelo de volatilidad continuo. El esfuerzo computacional es demasiado grande y

complica su implementación. La inferencia indirecta trata de encontrar un modelo de volatilidad continuo que haya sido observado de manera discreta en el tiempo y que produzca las mismas estimaciones que un modelo discreto. Ver *Gourieroux, Monfort & Renault (1993)*

Entonces utilizando inferencia bayesiana sobre el proceso ARCH e implementando los métodos MCMC se pueden obtener muestras de la distribución posterior de la volatilidad y a través de estas obtener muestras de la distribución posterior del precio de la opción por medio de la fórmula de Black & Scholes.

Este punto es justamente la base sobre la cual se trabaja todo el tercer capítulo, de hecho la finalidad de la tesis.

## Capítulo IV

# INFERENCIA BAYESIANA PARA LA VOLATILIDAD EN EL MODELO BLACK & SCHOLES

El presente capítulo está encaminado a mostrar la posibilidad de utilizar métodos MCMC para inferencia bayesiana en modelos de precios de activos en tiempo continuo.

La solución bayesiana al problema de inferencia, en este tipo de modelos, es la distribución de los parámetros,  $\Theta$ , y las variables condicionales latentes,  $X$ , en datos observados  $Y$ . Es así como la distribución posterior,  $p(\Theta, X | Y)$ , combina la información en el modelo y los precios observados, y es el punto de partida para la inferencia estadística. Los métodos MCMC proveen de una herramienta para explorar éstas distribuciones posteriores y a partir de éstas generar una cadena de Markov  $(\Theta, X)$ ,  $\{\Theta^{(g)}, X^{(g)}\}_{g=1}^G$  cuya distribución límite es  $p(\Theta, X | Y)$ . El método Monte Carlo usa esas muestras para integración numérica, para la estimación de los parámetros, para la estimación de los estados de la cadena y también para comparar entre modelos.

Caracterizar  $p(\Theta, X | Y)$  en modelos de activos en tiempo continuo es difícil por varias razones. Primero porque los precios son observados discretamente mientras que los modelos teóricos se especifican en tiempo continuo. Segundo porque en algunos casos las variables estado son latentes a partir de la perspectiva del investigador. Tercero porque la distribución posterior  $p(\Theta, X | Y)$  es típicamente de muy alta dimensión y por tanto los muestreos realizados por métodos estándar comúnmente fallan. Cuarto, en algunos modelos de tiempo continuo se generan distribuciones de transición para precios y variables estado que no son normales ni estándares, es por ello que la estimación a partir de métodos tales como el de máxima verosimilitud resulta muy complicado. Finalmente, en modelos de precios para las opciones, como lo es el de Black & Scholes, y en modelos estructurales, los parámetros resultan ser no lineales y a veces podrían no tener una forma analítica que pueda representar la solución de una ecuación diferencial ordinaria o parcial. Los métodos MCMC como pueden ser el Gibbs Sampling o el Metropolis-Hastings enfrentan todas estas dificultades.

Ahora bien, matemáticamente los métodos MCMC son particularmente bien adap-

tables para aplicaciones financieras en tiempo continuo por varias razones, la más importante es que estos tipos de modelos especifican que los precios y las variables estados resuelven ecuaciones diferenciales estocásticas los cuales son construidas a partir de movimientos brownianos, procesos Poisson, etc. Cuyas distribuciones resultan fáciles de caracterizar. Cuando se trabajan de manera discreta y en cualquier intervalo de tiempo finito, los modelos toman la forma de modelos familiares de series de tiempo con distribuciones de errores de manera normal o como una mezcla de normales discretas. Esto implica que las herramientas de la inferencia bayesiana pueden ser directamente aplicables a esos modelos. Para ver un mayor número de razones propuestos en la teoría de la econometría financiera, ver Johannes (2003).

Se han cimentado las razones por las cuales la utilización de los esquemas MCMC facilitan la inferencia bayesiana en los modelos financieros en tiempo continuo. Ahora es tiempo de abordar la finalidad de la tesis, es tiempo de comenzar la inferencia bayesiana en el modelo de Black & Scholes.

## ***4.1. Alternativas realizadas y propuesta actual al modelo de Black & Scholes***

### **4.1.1. Alternativas realizadas**

En diversas ocasiones se ha trabajado ya realizando inferencia bayesiana en el modelo de Black & Scholes.

Johannes (2003) usó inferencia bayesiana para el precio de las opciones, se tomó a Black & Scholes como el modelo que especifica ese precio y se le anexa un término de error con distribución normal, de tal modo que ésta será la serie de tiempo.

Es decir, como ya se especificó en el primer capítulo el precio de las opciones a partir de Black & Scholes queda determinado por la siguiente ecuación

$$C_t = BS(\sigma, S_t) = S_t N(d_1) - e^{r(T-t)} K N(d_1 - \sigma\sqrt{T-t})$$

donde

$$d_1 = \frac{\log(S_t/K) + (r + \sigma^2/2)(T-t)}{\sigma\sqrt{T-t}}$$

y entonces se especifica que

$$\log(S_t/S_{t-1}) = \mu + \sigma\epsilon_t$$



$$C_t = BS(\sigma, S_t) + \epsilon_t^c$$

donde  $\epsilon_t \sim N(0, 1)$  y  $\epsilon_t^c \sim N(0, \sigma_c^2)$ .

A partir de estas ecuaciones se realiza la inferencia bayesiana y se desarrolla un esquema MCMC para inferir la volatilidad, es decir  $\sigma^2$ . De tal forma se obtienen distribuciones posteriores para cada uno de los parámetros, y después se puede obtener la distribución posterior para  $C$ , es decir  $p(C_t | S_t, \sigma^2, \sigma_c^2)$ .

No ha sido la única vez que se ha hecho inferencia bayesiana en el modelo Black & Scholes, Bauwens y Lubrano (1998) infirieron sobre la media predictiva del precio de las opciones vía este mismo modelo. En este trabajo se calcula la media predictiva como

$$E(BS_{T+s} | y) = E_{\theta|y} [E(BS_{T+s} | \theta, y)]$$

para su cálculo se definió primero a la volatilidad mediante un modelo GARCH con errores student, y después se utilizó un método diferente MCMC llamado Gridy Gibbs Sampler para el cálculo de dicha media predictiva. Para un estudio más detallado, ver Bauwens (1998).

#### 4.1.2. Propuesta actual

En el presente trabajo se muestra una propuesta diferente de inferencia bayesiana en el modelo Black & Scholes. Se trata de reunir aspectos importantes de investigaciones como las mencionadas anteriormente y complementarlas, es decir Johannes (2003) encontró una distribución posterior para el precio de las opciones vía Black & Scholes pero no legitimó la veracidad de que la volatilidad fuera constante, en contraparte Bauwens (1998) utilizó un modelo GARCH para la volatilidad del modelo y encontró una media predictiva a lo largo del tiempo, sin embargo en ningún momento se desarrolla alguna distribución posterior para el precio de las opciones.

Se propone que el hecho de que la volatilidad sea considerada como constante a lo largo del tiempo es erróneo y es por esto que es indispensable definirla mediante un modelo de series de tiempo para volatilidad. Para lograr esto se utilizará inferencia bayesiana para encontrar una distribución posterior de un modelo ARCH mediante mecanismos MCMC, en particular el algoritmo Metropolis-Hastings de propuesta independiente, y una vez obtenida ésta distribución posterior se implementará sobre el modelo Black & Scholes para encontrar la distribución del precio de las opciones.

El hecho de que se utilicen herramientas bayesianas sobre el modelo ARCH es porque ésta medida de la incertidumbre es más poderosa que solo una estimación puntual del modelo.

También, en el más puro sentido financiero, una medida del precio de las opciones mediante el modelo Black & Scholes o algún modelo binomial que se pueda utilizar carece de toda información de la incertidumbre, lo que se pretende con encontrar una distribución posterior del precio de las opciones es obtener una mejora sobre dicha información.

#### ***4.2. Elementos para el desarrollo de la inferencia bayesiana sobre el modelo ARCH***

Como ya se ha mencionado con anterioridad los modelos discretos de ecuación determinista de series de tiempo para estudiar la volatilidad de un activo parecen ser los menos complicados para hacer inferencia en los parámetros. Sin embargo parece que brindar una información más completa que una estimación puntual, aún con éstos modelos, de la estructura de la volatilidad resultase ser complicado. Con la versatilidad de los mecanismos MCMC y de la inferencia bayesiana estos contratiempos pueden superarse.

Para ello es necesario definir primero un modelo sobre el cual se practicará lo anterior, en el presente trabajo se propone utilizar un modelo ARCH(2) para describir la volatilidad de la serie financiera, la cuál será el Índice de Precios y Cotizaciones de la Bolsa Mexicana de Valores (IPC) en el período 1990-2004. Se muestra en la figura (14).

Ahora bien, para lograr realizar inferencia bayesiana se necesitan de los siguientes elementos:

- 1 Parametrización exacta del modelo.
- 2 Función de verosimilitud del modelo.
- 3 Distribución inicial del modelo.
- 4 Distribución posterior.

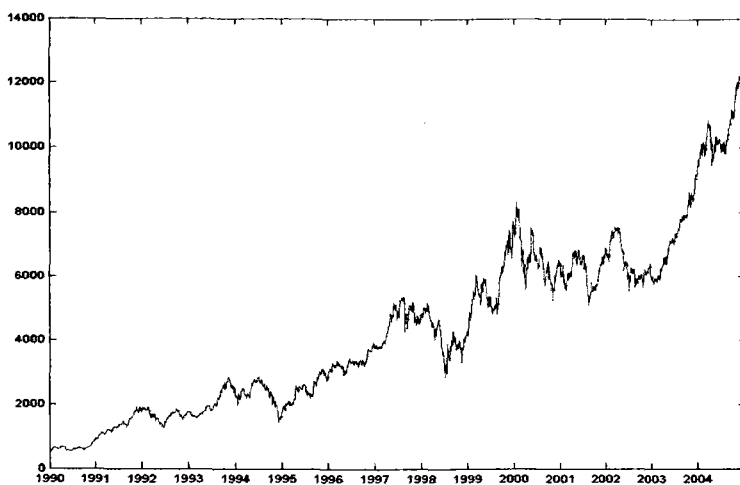


Figura 14: a) Serie histórica del IPC 1990-2004

## 4.2.1. Elementos para inferencia bayesiana

### 4.2.1.1. Parametrización del modelo

Como ya fue definido en la sección (2.3.7) el modelo ARCH( $r$ ) está expresado de la siguiente forma

$$y_t = \varepsilon_t \sqrt{h_t} \quad \varepsilon_t \sim N(0, 1)$$

$$h_t = \alpha_0 + \alpha_1 y_{t-1}^2 + \dots + \alpha_r y_{t-r}^2$$

Por tanto el modelo ARCH(2) queda definido como sigue

$$y_t = \varepsilon_t \sqrt{h_t} \quad \varepsilon_t \sim N(0, 1)$$

$$h_t = \alpha_0 + \alpha_1 y_{t-1}^2 + \alpha_2 y_{t-2}^2$$

$$\theta = (\alpha_0, \alpha_1, \alpha_2)'$$
(48)

### 4.2.1.2. Función de verosimilitud

La función de verosimilitud para un ARCH( $r$ ) bajo el supuesto de normalidad es

$$f(y_1, \dots, y_n | \Theta) = \prod_{t=1}^n \frac{1}{\sqrt{2\pi h_t}} e^{-\frac{y_t^2}{2h_t}}$$
(49)

donde  $\Theta = (\alpha_0, \alpha_1, \dots, \alpha_r)$ .

En términos de densidades, la ecuación (49) puede ser reescrita como

$$f(y_1, \dots, y_n | \Theta) = f(y_n | F_{n-1}) f(y_{n-1} | F_{n-2}), \dots, f(y_{r+1} | F_r) f(y_1, y_2, \dots, y_r | \Theta)$$

La forma exacta de  $f(y_1, y_2, \dots, y_r | \Theta)$  es complicada así que comunmente se condiciona con respecto a las primeras  $r$  observaciones para trabajar solamente con la verosimilitud condicional, ver Tsay(2002).

$$f(y_{r+1}, \dots, y_n | \Theta) = \prod_{t=r+1}^n \frac{1}{\sqrt{2\pi h_t}} e^{-\frac{y_t^2}{2h_t}} \quad (50)$$

Por tanto a partir de (50), la verosimilitud que se trabajará con el modelo ARCH(2) es

$$L(\theta | \underline{y}) = \left(\frac{1}{2\pi}\right)^{n/2} \left[ \prod_{t=1}^n \left(\frac{1}{h_t}\right)^{\frac{1}{2}} \right] \exp\left\{-\frac{1}{2} \sum_{t=1}^n \frac{y_t^2}{h_t}\right\} \quad (51)$$

El hecho de que la función de verosimilitud condicional para los modelos ARCH sea de este modo calculada es una mayor razón para la rápida adopción de ésta clase de modelos en econometría, en la economía financiera y en la macroeconomía. Aún ésto, una forma moderna de llevar a cabo inferencia estadística es vía MCMC. Para revisar extensivamente el uso de los métodos MCMC en el contexto de la econometría, ver Chib (2001) y Florentini (2002).

#### 4.2.1.3. Distribuciones iniciales

La definición de las distribuciones iniciales para los parámetros del modelo ARCH(2) están basadas en las restricciones de los mismos para que cumplan diversas condiciones de regularidad para asegurar que la varianza incondicional de  $y_t$  sea finita.

Dado un modelo ARCH( $r$ )

$$y_t = h_t \varepsilon_t, \quad h_t^2 = \alpha_0 + \alpha_1 y_{t-1}^2 + \dots + \alpha_r y_{t-r}^2$$

se debe de cumplir que  $\alpha_0 > 0$ ,  $\alpha_i \geq 0 \forall i > 0$  y  $\sum_{i>0} \alpha_i < 1$ .

Por tanto para un modelo ARCH(2) se debe de cumplir que  $\alpha_0 > 0$ ,  $\alpha_1 > 0$ ,  $\alpha_2 > 0$  y  $\alpha_1 + \alpha_2 < 1$ . Esto conlleva lógicamente a que  $0 < \alpha_1 < 1$  y  $0 < \alpha_2 < 1$ .

De esta forma las distribuciones a priori para los parámetros quedan definidas como

$$p(\theta) = p(\alpha_0)p(\alpha_1)p(\alpha_2) = I_{(0,\infty)}^{(\alpha_0)} I_{(0,1)}^{(\alpha_1)} I_{(0,1)}^{(\alpha_2)} \quad (52)$$

#### 4.2.1.4. Distribuciones posteriores

Como se explicó en la sección (2.1.3) la distribución posterior se calcula de la siguiente forma

$$p(\theta | \underline{y}) \propto L(\theta | \underline{y}) p(\theta)$$

Entonces, a partir de (51) y (52) la distribución posterior para nuestro modelo de inferencia queda establecido como

$$p(\theta | \underline{y}) \propto \left[ \prod_{t=1}^n \left( \frac{1}{h_t} \right)^{\frac{1}{2}} \right] \exp \left\{ -(1/2) \sum_{t=1}^n \frac{y_t^2}{h_t} \right\} I_{(0,\infty)}^{(\alpha_0)}, I_{(0,1)}^{(\alpha_1)}, I_{(0,1)}^{(\alpha_2)} \quad (53)$$

Sin embargo la distribución posterior  $\pi(\theta, \underline{y})$  de los parámetros del modelo ARCH(2) no tiene una forma analítica cerrada, entonces se propone plantear un esquema MCMC para obtener muestras de la distribución posterior de los parámetros. Ya ha sido explicado la metodología de la simulación estocástica en la sección (2.2). Ahora lo que se necesita calcular son los elementos que se necesitan para desarrollarla.

El método MCMC que se utilizará es el algoritmo Metropolis-Hastings de propuesta independiente. Se necesitan los siguientes elementos:

- 1 Distribución de salto (Propuesta independiente).
- 2 Cálculo de la probabilidad de salto.  $\alpha(\theta^{(j-1)}, \theta^{(j)})$
- 3 Cálculo del logrativo.  $Ln(\alpha(\theta^{(j-1)}, \theta^{(j)}))$

### 4.2.2. Elementos para MCMC

#### 4.2.2.1. Propuesta independiente

Como se explicó en la sección (2.2.4) el algoritmo Metropolis-Hastings puede tener casos especiales en su distribución instrumental.

- 1  $q(z | y) = q(y | z)$ : Algoritmo Metropolis.
- 2  $q(z | y) = \prod f(z_i | z_{<i}, y_{>i})$ : Gibbs sampler.
- 3  $q(z | y) = q(z)$ : Propuesta independiente.

Esta última es la que se utilizará para el cálculo de la probabilidad de salto  $\alpha(\theta^{(j-1)}, \theta^{(j)})$ .

La propuesta independiente es una normal trivariada

$$q(\theta^{(j)} | \theta^{(j-1)}) = q(\theta^{(j)}) \sim N_3(\theta, c\Sigma) \quad (54)$$

Aquí  $\theta = (\alpha_0, \alpha_1, \alpha_2)$  y  $\theta^{(j)}$  es el vector de parámetros en la iteración  $j$ ,  $c$  es una constante para calibrar la tasa de rechazo y  $\Sigma$  es una matriz de covarianzas muestral que se estima por medio de una corrida exploratoria de la cadena de Markov. Se explicarán a detalle más adelante.

#### 4.2.2.2. Cálculo de la probabilidad de salto

La probabilidad de salto, utilizando una propuesta independiente se define como

$$\begin{aligned} \alpha(\theta^{(j-1)}, \theta^{(j)}) &= \min \left[ \frac{\pi(\theta^{(j)} | \cdot) q(\theta^{(j-1)} | \theta^{(j)})}{\pi(\theta^{(j-1)} | \cdot) q(\theta^{(j)} | \theta^{(j-1)})}, 1 \right] \\ &= \min \left[ \frac{\pi(\theta^{(j)}) q(\theta^{(j-1)})}{\pi(\theta^{(j-1)}) q(\theta^{(j)})}, 1 \right] \end{aligned} \quad (55)$$

Ahora utilizando (54) y (53) en (55), tenemos que

$$\begin{aligned} \alpha(\theta^{(j-1)}, \theta^{(j)}) &= \\ \min \left[ \frac{\left[ \prod_{t=1}^n \left( \frac{1}{h_t} \right)^{1/2} \right] \exp \left\{ -\frac{1}{2} \sum_{t=1}^n \frac{y_t^2}{h_t} \right\} c \Sigma^{-1} (2\pi)^{-3/2} \exp \left\{ -\frac{1}{2} (\theta^{(j-1)} - \theta_0)' c \Sigma^{-1} (\theta^{(j-1)} - \theta_0) \right\}}{\left[ \prod_{t=1}^n \left( \frac{1}{h_t} \right)^{1/2} \right] \exp \left\{ -\frac{1}{2} \sum_{t=1}^n \frac{y_t^2}{h_t} \right\} c \Sigma^{-1} (2\pi)^{-3/2} \exp \left\{ -\frac{1}{2} (\theta^{(j)} - \theta_0)' c \Sigma^{-1} (\theta^{(j)} - \theta_0) \right\}}}, 1 \right] \end{aligned} \quad (56)$$

En la ecuación (56) se deben tomar en cuenta las siguientes consideraciones.

En el numerador:

- $\prod_{t=1}^n \left( \frac{1}{h_t} \right)^{1/2}$  y  $\exp \left\{ -\frac{1}{2} \sum_{t=1}^n \frac{y_t^2}{h_t} \right\}$  pertenecen al salto (j)
- La componente de la normal trivariada pertenece al salto (j-1)

En el denominador

- $\prod_{t=1}^n \left( \frac{1}{h_t} \right)^{1/2}$  y  $\exp \left\{ -\frac{1}{2} \sum_{t=1}^n \frac{y_t^2}{h_t} \right\}$  pertenecen al salto (j-1)
- La componente de la normal trivariada pertenece al salto (j)

La importancia de ésto se comentará más a detalle cuando se tenga calculado el logratio.

#### 4.2.2.3. Cálculo del logratio

El logratio se utiliza como una forma más sencilla de calcular las probabilidades de transición al momento de simular computacionalmente. Se define como

$$Ln(\alpha(\theta^{(j-1)}, \theta^{(j)})) \quad (57)$$

Entonces, a partir de (56) se define

$$\Upsilon_1 = \left[ \prod_{t=1}^n \left( \frac{1}{h_t} \right)^{1/2} \right] \exp \left\{ -\frac{1}{2} \sum_{t=1}^n \frac{y_t^2}{h_t} \right\} c \sum^{-1} (2\pi)^{-3/2} \exp \left\{ -\frac{1}{2} (\theta^{(j-1)} - \theta_0)' c \sum^{-1} (\theta^{(j-1)} - \theta_0) \right\} \quad (58)$$

$$\Upsilon_2 = \left[ \prod_{t=1}^n \left( \frac{1}{h_t} \right)^{1/2} \right] \exp \left\{ -\frac{1}{2} \sum_{t=1}^n \frac{y_t^2}{h_t} \right\} c \sum^{-1} (2\pi)^{-3/2} \exp \left\{ -\frac{1}{2} (\theta^{(j)} - \theta_0)' c \sum^{-1} (\theta^{(j)} - \theta_0) \right\} \quad (59)$$

La forma de calcular el logratio se realiza teniendo presente que a partir de propiedades algebraicas se puede deducir la fórmula de la siguiente forma

$$\begin{aligned} Ln[\alpha(\theta^{(j-1)}, \theta^{(j)})] &= Ln \left[ \min \left( \frac{\Upsilon_1}{\Upsilon_2}, 1 \right) \right] = \\ &= \min \left[ Ln \left( \frac{\Upsilon_1}{\Upsilon_2} \right), 0 \right] = \min [Ln(\Upsilon_1) - Ln(\Upsilon_2), 0] \end{aligned}$$

Trabajemos primero con el término  $\Upsilon_1$ .

$$\begin{aligned} \text{Sea } \xi_1 &= \left[ \prod_{t=1}^n \left( \frac{1}{h_t} \right)^{1/2} \right], \quad \xi_2 = \exp \left\{ -\frac{1}{2} \sum_{t=1}^n \frac{y_t^2}{h_t} \right\} \text{ y} \\ \xi_3 &= c \sum^{-1} (2\pi)^{-3/2} \exp \left\{ -\frac{1}{2} (\theta^{(j-1)} - \theta_0)' c \sum^{-1} (\theta^{(j-1)} - \theta_0) \right\} \end{aligned}$$

Teniendo en mente que  $Ln(\Upsilon_1) = Ln(\xi_1 \xi_2 \xi_3) = Ln(\xi_1) + Ln(\xi_2) + Ln(\xi_3)$

$$\begin{aligned} Ln(\xi_1) &= Ln \left[ \prod_{t=1}^n \left( \frac{1}{h_t} \right)^{1/2} \right] = Ln \left[ \left( \frac{1}{h_1} \right)^{1/2} \left( \frac{1}{h_2} \right)^{1/2} \dots \left( \frac{1}{h_n} \right)^{1/2} \right] \\ &= Ln \left[ \left( \frac{1}{h_1} \right)^{1/2} \right] + Ln \left[ \left( \frac{1}{h_2} \right)^{1/2} \right] + \dots + Ln \left[ \left( \frac{1}{h_n} \right)^{1/2} \right] = \sum_{t=1}^n Ln \left[ \left( \frac{1}{h_t} \right)^{1/2} \right] \end{aligned}$$

$$Ln(\xi_2) = Ln \left[ \exp \left\{ -\frac{1}{2} \sum_{t=1}^n \frac{y_t^2}{h_t} \right\} \right] = -\frac{1}{2} \sum_{t=1}^n \frac{y_t^2}{h_t}$$

$$\begin{aligned} Ln(\xi_3) &= Ln \left[ c \sum^{-1} (2\pi)^{-3/2} \exp \left\{ -\frac{1}{2} (\theta^{(j-1)} - \theta_0)' c \sum^{-1} (\theta^{(j-1)} - \theta_0) \right\} \right] \\ &= Ln \left[ c \sum^{-1} \right] + Ln \left[ (2\pi)^{-3/2} \right] + Ln \left[ \exp \left\{ -\frac{1}{2} (\theta^{(j-1)} - \theta_0)' c \sum^{-1} (\theta^{(j-1)} - \theta_0) \right\} \right] \end{aligned}$$

$$Ln(\xi_3) = Ln \left[ c \sum^{-1} \right] + Ln \left[ (2\pi)^{-3/2} \right] - \frac{1}{2} (\theta^{(j-1)} - \theta_0)' c \sum^{-1} (\theta^{(j-1)} - \theta_0)$$

Por tanto  $Ln(\Upsilon_1)$  queda expresado finalmente como

$$Ln(\Upsilon_1) = \sum_{t=1}^n Ln \left[ \left( \frac{1}{h_t} \right)^{1/2} \right] - \frac{1}{2} \sum_{t=1}^n \frac{y_t^2}{h_t} + Ln \left[ c \sum^{-1} \right] + Ln \left[ (2\pi)^{-3/2} \right] - \frac{1}{2} (\theta^{(j-1)} - \theta_0)' c \sum^{-1} (\theta^{(j-1)} - \theta_0)$$

El cálculo de  $Ln(\Upsilon_2)$  se calcula de la misma forma y queda expresado como

$$Ln(\Upsilon_2) = \sum_{t=1}^n Ln \left[ \left( \frac{1}{h_t} \right)^{1/2} \right] - \frac{1}{2} \sum_{t=1}^n \frac{y_t^2}{h_t} + Ln \left[ c \sum^{-1} \right] + Ln \left[ (2\pi)^{-3/2} \right] - \frac{1}{2} (\theta^{(j)} - \theta_0)' c \sum^{-1} (\theta^{(j)} - \theta_0)$$

Al calcular  $Ln(\Upsilon_{21}) - Ln(\Upsilon_2)$  se obtiene que

$$Ln(\Upsilon_1) - Ln(\Upsilon_2) = \sum_{t=1}^n Ln \left[ \left( \frac{1}{h_t} \right)^{1/2} \right] - \frac{1}{2} \sum_{t=1}^n \frac{y_t^2}{h_t} - \frac{1}{2} (\theta^{(j-1)} - \theta_0)' c \sum^{-1} (\theta^{(j-1)} - \theta_0) - \sum_{t=1}^n Ln \left[ \left( \frac{1}{h_t} \right)^{1/2} \right] + \frac{1}{2} \sum_{t=1}^n \frac{y_t^2}{h_t} + \frac{1}{2} (\theta^{(j)} - \theta_0)' c \sum^{-1} (\theta^{(j)} - \theta_0) \quad (60)$$

Por tanto a partir de (60) se obtiene el logratio de la distribución de salto como

$$Ln[\alpha(\theta^{(j-1)}, \theta^{(j)})] = \min \left[ \sum_{t=1}^n Ln \left[ \left( \frac{1}{h_t} \right)^{1/2} \right] - \frac{1}{2} \sum_{t=1}^n \frac{y_t^2}{h_t} - \frac{1}{2} (\theta^{(j-1)} - \theta_0)' c \sum^{-1} (\theta^{(j-1)} - \theta_0) - \sum_{t=1}^n Ln \left[ \left( \frac{1}{h_t} \right)^{1/2} \right] + \frac{1}{2} \sum_{t=1}^n \frac{y_t^2}{h_t} + \frac{1}{2} (\theta^{(j)} - \theta_0)' c \sum^{-1} (\theta^{(j)} - \theta_0), 0 \right] \quad (61)$$

En comparación la forma que presenta la ecuación (61) resulta computacionalmente más sencilla calcularla que la que se presenta (56). Esto debe de tomarse en cuenta al momento de implementarlo en un sistema o programa de simulación, ya que un error en el cálculo desemboca directamente en un error en las muestras de la distribución posterior vía Metropolis-Hastings.

### 4.3. Desarrollo MCMC

Antes de implementar el esquema MCMC con datos reales se estudia la eficiencia del algoritmo utilizando una serie de retornos simulada directamente de un proceso ARCH(2) de tamaño  $n$  con parámetros preestablecidos  $\alpha_0$ ,  $\alpha_1$  y  $\alpha_2$ . A partir de aquí el



mecanismo es el mismo que cuando se trabaja con los retornos de los datos reales.

Antes de la implementación del Metropolis-Hastings de propuesta independiente se debe de estimar puntualmente a los parámetros y a los errores estándar del modelo ARCH. Existen paquetes estadísticos y econométricos para este propósito como Matlab, Eviews, Gauss, SPSS, etc.

Con la estimación máxima verosímil de los parámetros se calcula la volatilidad  $h_t$  del primer estado de la cadena y el vector de medias para la propuesta independiente, en este caso una normal trivariada. También los valores de tales estimaciones serán la primera entrada de la columna de la matriz de las iteraciones de los parámetros posteriores  $M_{(T)}$ . Una vez realizado esto, se procede con la iteración del algoritmo Metropolis-Hastings muestreando el vector  $\theta$  en un sólo paso a través de la propuesta independiente  $q(\theta^{(j)} | \theta^{(j-1)})$  que se obtiene a partir de la normal trivariada  $N_3(\theta^{(j)}, c\Sigma)$ .

Se debe de tener en cuenta dos asuntos importantes acerca del cálculo de ésta normal trivariada. La primera de ellas es que la matriz de covarianzas muestral  $\Sigma$  se debe de estimar a través de una corrida exploratoria de la cadena de Markov, es decir que debe de hacerse una corrida de la cadena de tamaño  $n$  de tal forma que la matriz  $M_{(T)}$  sea de dimensión  $3 \times n$  y después calcularse la matriz de covarianzas muestrales de  $M_{(T)}$ . La dimensión  $3 \times n$  representan los  $n$  valores obtenidos vía Metropolis-Hastings de los tres parámetros del modelo ARCH(2). El asunto de la calibración de la constante  $c$  se hace de manera posterior, es decir una vez que ya se tiene calculada  $\Sigma$ . Se busca de manera iterativa el valor de  $c$  que satisfaga una tasa de aceptación alrededor del 45%. Se trata de mantener dicha tasa alrededor de este valor porque así se le permite a la propuesta independiente tener mayor libertad para muestrear distintas secciones de lo que será la distribución posterior y evitar que solo muestree en modas locales. La especificación de la tasa de aceptación depende del tipo de modelos que se esté utilizando, hay veces que se trabajan con tasas entre el 30 y 45%, existen ocasiones que no es posible mantener las tasas por abajo del 50%. Para el presente trabajo tanto para los retornos simulados como para los reales se trabajaron tasas por debajo del 45% y por encima del 40%. Resumiendo lo anterior, una buena tasa de aceptación dependerá de una buena elección de la matriz de covarianzas y una buena calibración de  $c$ .

Una vez que se ha fijado una tasa de aceptación a trabajar, ahora se necesita fijar el número de iteraciones del Metropolis-Hastings. Este es un nuevo problema y no hay

un número de iteraciones con la cual se trabaje de forma general ya que dependerá del tipo de modelo que se esté utilizando. Pedersen(2004) sugiere que para modelos estadísticos de series de tiempo se utilicen un número de iteraciones no menores a 100,000, sin embargo este número de iteraciones puede reducirse utilizando métodos de aceleración de convergencia para MCMC, ver Gamerman (1997). Sin embargo Bauwens(1998) sugiere iterar 10,000 veces el Metropolis-Hastings para encontrar convergencia en la cadena, por si fuera poco Fiorentini(2002) y Zhang(2003) concuerdan que una iteración de tamaño 500,000 y un burn-in-period de 50,000 asegurará dicha convergencia. El número de iteraciones y el burn-in-period depende del modelo que se utilice y de la potencia del equipo computacional con el que se cuente. Para este trabajo y se comentará a detalle más adelante se hicieron pruebas desde 10,000 hasta 300,000 iteraciones encontrando los mejores resultados en los que se presentan.

#### 4.3.1. Implementación del Metropolis-Hastings de propuesta independiente con retornos simulados

Para estudiar el funcionamiento del algoritmo en retornos simulados se simula un proceso ARCH(2) con parámetros  $\alpha_0 = 0.5$ ,  $\alpha_1 = 0.3$  y  $\alpha_2 = 0.1$ , siguiendo el mecanismo que se acaba de explicar e iterando el algoritmo Metropolis-Hastings de propuesta independiente con un tamaño de  $N = 100,000$  se obtiene la matriz

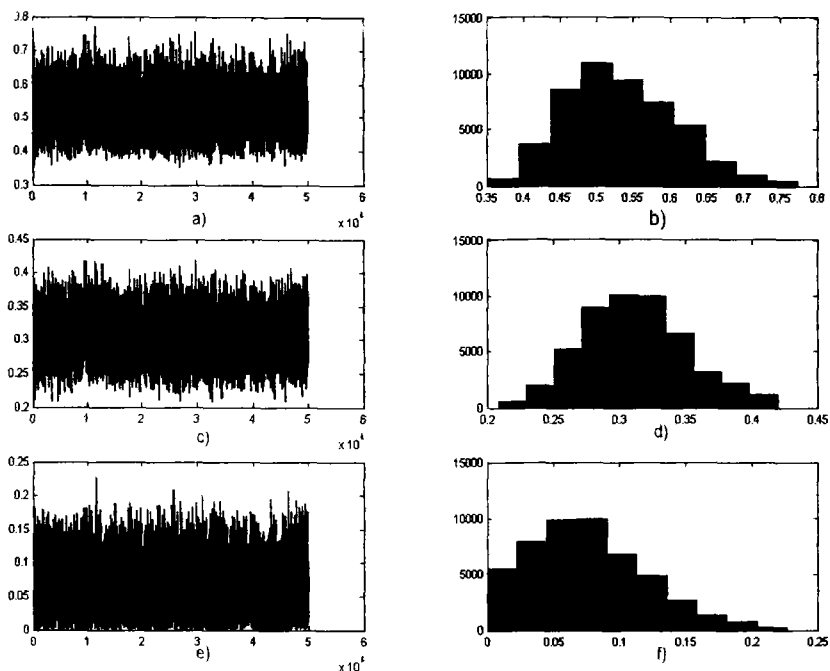
$$M_{(T)} = \begin{pmatrix} \alpha_0^{(0)} & \alpha_0^{(1)} & \dots & \alpha_0^{(50000)} & \alpha_0^{(50001)} & \dots & \alpha_0^{(100000)} \\ \alpha_1^{(0)} & \alpha_1^{(1)} & \dots & \alpha_1^{(50000)} & \alpha_1^{(50001)} & \dots & \alpha_1^{(100000)} \\ \alpha_2^{(0)} & \alpha_2^{(1)} & \dots & \alpha_2^{(50000)} & \alpha_2^{(50001)} & \dots & \alpha_2^{(100000)} \end{pmatrix} \quad (62)$$

En (62) se desechan las iteraciones marcadas con rojo que componen el burn-in-period para proseguir la inferencia estadística. Se calcula la media y la desviación estándar para cada una de los parámetros a partir de la iteración 50001 en adelante. Esto compone la media y la desviación estándar posterior de los parámetros.

La figura (15) representa el histograma posterior de los parámetros del modelo ARCH(2). Observese que la cantidad de las iteraciones de las trazas de la distribución posterior de los parámetros del modelo es de 50000, lo que equivale a los valores de los parámetros en color negro de la matriz  $M_{(T)}$ .

Hasta aquí se ha obtenido la información que se presenta en el cuadro(4).

Se debe de observar en el cuadro(4) que el algoritmo Metropolis-Hastings estina de manera adecuada a los parámetros del modelo ARCH(2) utilizado.



**Figura 15:** a) Iteraciones  $\alpha_0$  b) Histograma posterior  $\alpha_0$  c) Iteraciones  $\alpha_1$  d) Histograma posterior  $\alpha_1$  e) Iteraciones  $\alpha_2$  f) Histograma posterior  $\alpha_2$

Ahora utilizando las medias posteriores ( $\alpha_0^*$ ,  $\alpha_1^*$  y  $\alpha_2^*$ ) se encuentra el siguiente modelo

$$h_t^* = \alpha_0^* + \alpha_1^* y_{t-1}^2 + \alpha_2^* y_{t-2}^2 \quad (63)$$

La ecuación (63) es la ecuación de la media posterior para la volatilidad. La figura (16) representa los retornos del proceso ARCH(2) y su media posterior de la volatilidad.

Entonces a partir de este diseño de MCMC se toma una sección de los retornos simulados entre el retorno 90 y 170 de 1500 y se estima la volatilidad que en forma figura queda representada por la media posterior de la volatilidad  $h_t^*$  y los intervalos de credibilidad al 95 %. La figura (17) muestra un comparativo entre el diseño MCMC de la volatilidad incluyendo los intervalos de credibilidad y la volatilidad estimada a partir de métodos de máxima verosimilitud mostrados en el cuadro (4).

Ahora como a partir del esquema MCMC empleado se obtienen muestras de la distribución posterior para los parámetros, entonces es posible obtener muestras de

Parámetro	Valor real	EMV	Media posterior	DE	DE posterior
			M-H	EMV	M-H
$\alpha_0$	0.5	0.51274	0.5332	0.032838	.0748
$\alpha_1$	0.3	0.30361	0.3122	0.045017	.0393
$\alpha_2$	0.1	0.06389	0.0756	0.030595	.0436

**Cuadro 4:** Estimación Maxima Verosimilitud (EMV) para la media y desviación estándar (DE) vs estimación por medio del Metropolis-Hastings (MH) de los valores reales de la simulación de un proceso ARCH(2).

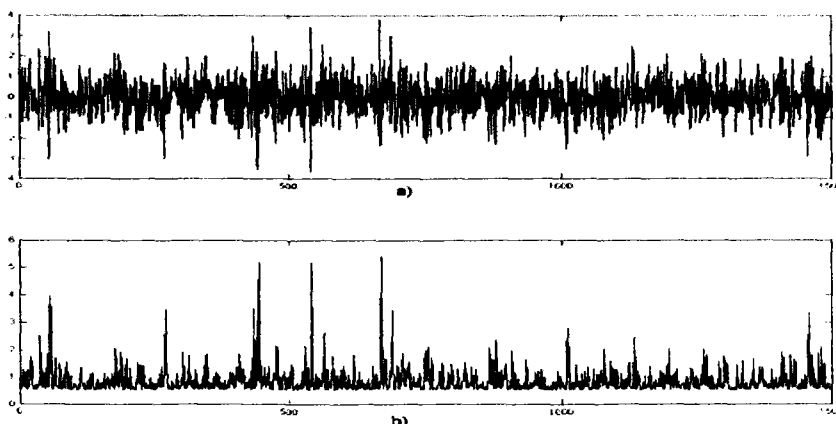
la distribución posterior para el modelo ARCH(2). Es decir, de la matriz  $M_{(T)}$  se obtiene de manera individual

$$\alpha_0^D = \begin{pmatrix} \alpha_0^{(50001)} \\ \vdots \\ \alpha_0^{(100000)} \end{pmatrix}^t \quad \alpha_1^D = \begin{pmatrix} \alpha_1^{(50001)} \\ \vdots \\ \alpha_1^{(100000)} \end{pmatrix}^t \quad \alpha_2^D = \begin{pmatrix} \alpha_2^{(50001)} \\ \vdots \\ \alpha_2^{(100000)} \end{pmatrix}^t$$

y ahora se calcula dichas muestras de la siguiente forma

$$h_t^D = \alpha_0^D + \alpha_1^D y_{n-1}^2 + \alpha_2^D y_{n-2}^2 \tag{64}$$

donde  $y_n$  representa el retorno del día  $n$ . Para los efectos que se buscan en el presente trabajo  $n$  representa la última observación de la serie de retornos. Es decir que a partir de la ecuación (64) se obtienen las muestras de la distribución posterior de la



**Figura 16:** a) Retornos proceso ARCH(2) con  $\alpha_0=0.5$ ,  $\alpha_1=0.3$  y  $\alpha_2=0.1$ . b) Media posterior de la volatilidad

volatilidad al retorno  $n$ . El kernel suavizado de dicha distribución posterior para los retornos simulados se muestra en la figura (18).

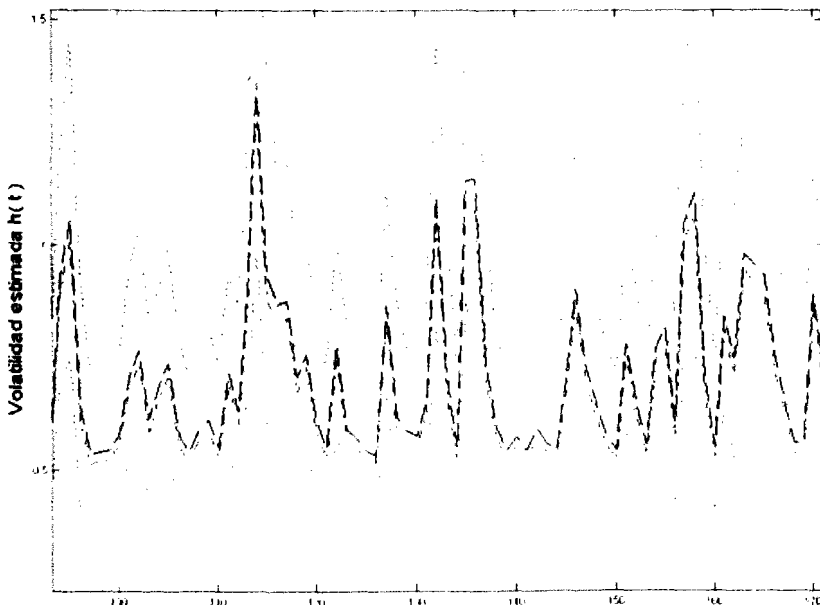
El significado más importante de la figura (18) radica en que es a partir de esta distribución con la que se buscará encontrar las muestras de la distribución posterior para el precio de la opción por medio del modelo Black & Scholes. Ahora bien, aunque el sentido de obtener dichas muestras utilizando datos simulados es limitado sirve muy bien este ejemplo para verificar su funcionamiento antes de encontrarlo con datos reales.

Las muestras de la distribución posterior para el precio de una opción de compra mediante el modelo Black & Scholes son calculados mediante la fórmula (65)

$$C = S_0 N(d_1) - X e^{-rT} N(d_2) \quad (65)$$

Donde

$$d_1 = \frac{\ln(S_0/X) + (r + (h_t^D)^2/2) T}{(h_t^D)\sqrt{T}} \quad d_2 = d_1 - (h_t^D)\sqrt{T}$$



**Figura 17:** Estimador Máxima Verosimilitud (línea punteada negra) Media posterior (línea punteada azul) Intervalos de credibilidad al 95% (líneas punteadas rojas)

$h_t^D$  representa el vector de las muestras de la distribución posterior de la volatilidad al retorno  $n$ . Los demás parámetros ya fueron definidos en el primer capítulo.

La figura (19) representa a la distribución posterior para el precio de la opción de compra con parámetros hipotéticos.

Una vez que ya ha sido explicado el desarrollo del mecanismo MCMC con retornos simulados se prosigue en la implementación del algoritmo Metropolis-Hastings de propuesta independiente con retornos reales.

#### 4.3.2. Implementación del Metropolis-Hastings de propuesta independiente con retornos reales

El subyacente financiero elegido corresponde a la serie histórica del IPC 1990-2004. El IPC es el principal indicador del comportamiento del mercado accionario de la Bolsa Mexicana de Valores (BMV), el cual expresa el rendimiento de este mercado tomando como referencia las variaciones de precios de una canasta balanceada, ponderada y representativa del total de los títulos accionarios cotizados en la BMV. El IPC expresa en forma fidedigna la situación del mercado bursátil y constituye un indicador altamente confiable. En cuanto a negociaciones bursátiles, al invertir en el IPC se está diversificando el riesgo de una cartera de inversión con los títulos con

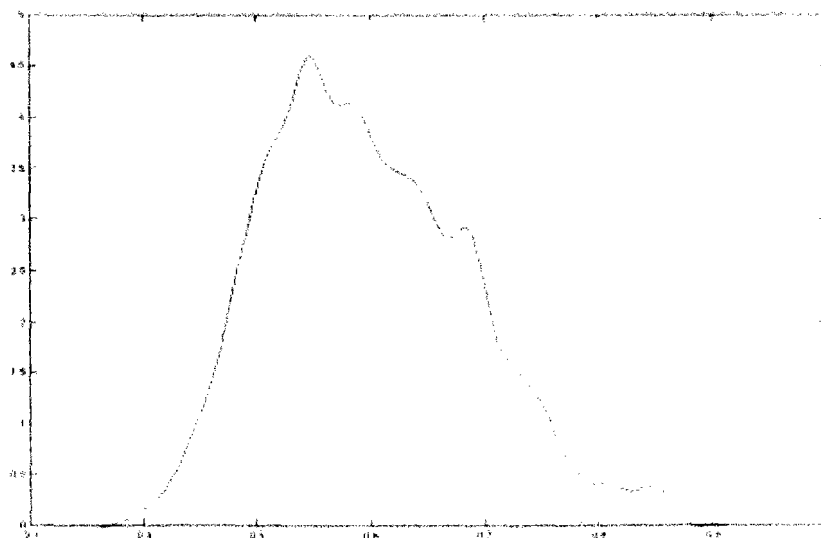
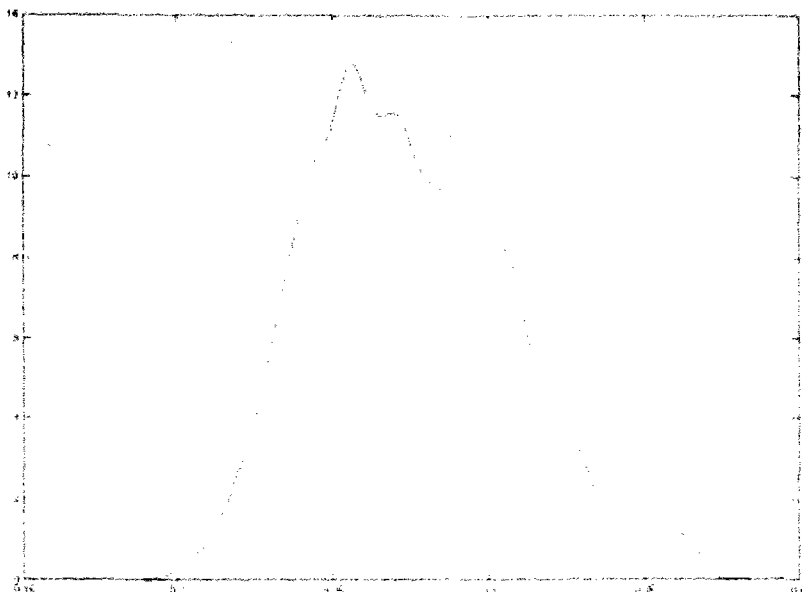


Figura 18: Distribución posterior de la volatilidad al retorno 1500



**Figura 19:** Distribución posterior del precio de una opción de compra vía Black & Scholes

mayor índice de bursatilidad en el país, es por ello que la forma de inversión en el IPC es exactamente la misma como si se estuviera invirtiendo en una cartera de inversión. El mecanismo de utilización de las primas de opciones en el IPC tiene un significado completamente igual que cuando se utilizan sobre acciones individuales, tasas de interés, bonos financieros, tipos de cambio, etc.

De la misma forma en que se desarrolló el mecanismo MCMC para los retornos simulados, se implementa con retornos reales.

Se comienza por estimar los parámetros  $\alpha_0$ ,  $\alpha_1$  y  $\alpha_2$  del modelo ARCH(2) por métodos máxima verosimilitud.

Aquí se debe de hacer hincapié en que la parametrización del modelo a utilizar para el algoritmo MCMC es

$$y_t = \epsilon_t \sqrt{h_t} \quad \epsilon_t \sim N(0, 1)$$

$$h_t = \hat{\alpha}_0 + \hat{\alpha}_1 y_{t-1}^2 + \hat{\alpha}_2 y_{t-2}^2$$

El hecho en que se tome que la varianza del error sea uno es correcto en este caso en particular ya que al ser el IPC una canasta bursátil controlada los retornos no mantienen una alta variación en comparación con las que se podría encontrar con acciones en momentos de alta incertidumbre. La figura (20) muestra el histograma de los retornos del IPC 1990-2004 en la cual se muestra que la variación de los retornos está acotada muy por debajo de uno. Sin embargo en un subyacente cuya variación en los retornos sea muy alta, la hipótesis de que la varianza sea uno sería errónea y es en estos casos donde optar por distribuciones más completas como una Gamma sobre los retornos sería altamente recomendable.

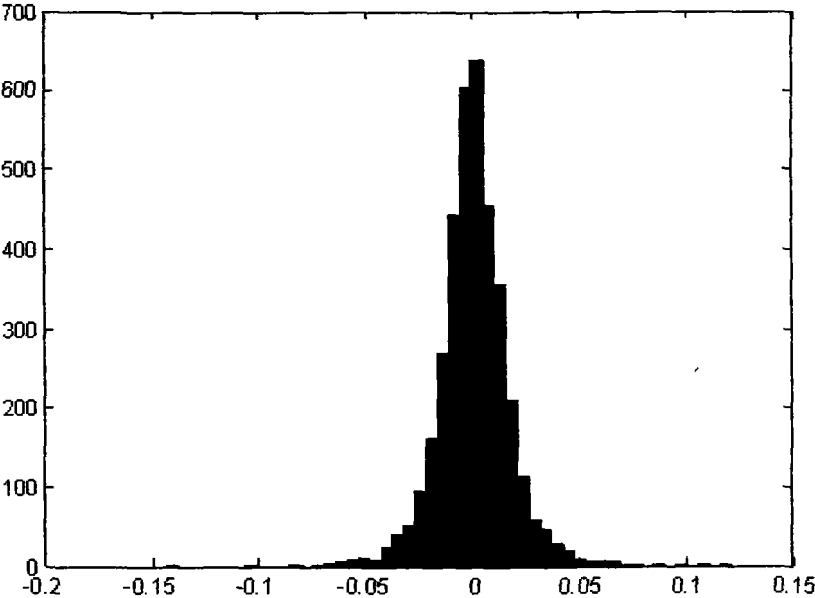


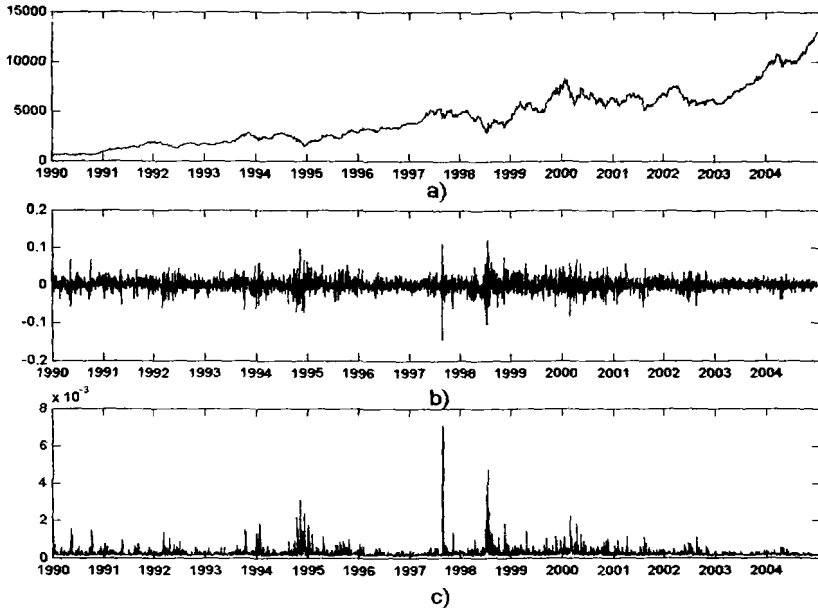
Figura 20: Histograma de los retornos del IPC 1990-2004

Una vez estimados los parámetros del modelo se continua con el mismo desarrollo que con los retornos simulados.

Se debe calibrar  $c\Sigma$  para proponer una tasa que esté alrededor del 45 % de aceptación.

La estimación máxima verosimil de los parámetros del modelo ARCH(2) se encuentran en el cuadro(5). La figura (21) ilustra al IPC con su serie de retornos y el modelo ARCH(2) estimado por máxima verosimilitud para la misma serie.





**Figura 21:** a) Serie IPC 1990-2004 b) Serie de retornos c) Volatilidad estimada por máxima verosimilitud

En la figura (22) se muestran los kernels suavizados de estimación para los parámetros del modelo para volatilidad utilizado.

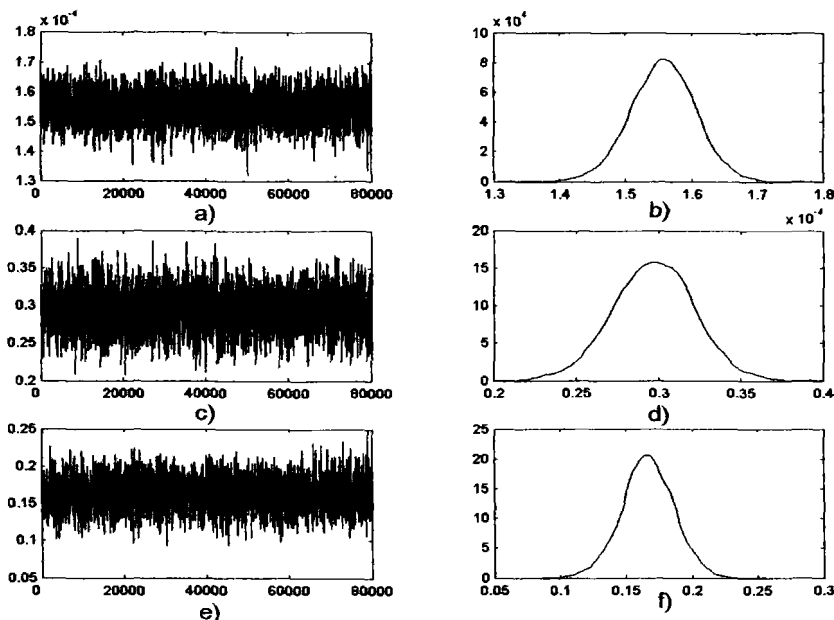
Se calcula a partir de la matriz  $M(T)$  en datos reales la media posterior de la volatilidad. El cuadro (5) muestra el comparativo entre los valores de los parámetros estimados por los métodos de máxima verosimilitud y el método MCMC empleado.

Parámetro	EMV	Media posterior M-H	DE EMV	DE posterior M-H
$\alpha_0$	0.0001551	0.000156	$0.39324 \times 10^{-6}$	$.5147 \times 10^{-6}$
$\alpha_1$	0.29368	0.2970	0.019534	.0248
$\alpha_2$	0.16431	0.1662	0.015759	.0198

**Cuadro 5:** Estimación Maxima Verosimilitud (EMV) para la media y desviación estándar (DE) vs estimación por medio del Metropolis-Hastings (MH)

Con las medias posteriores de los parámetros ( $\alpha_0^*$ ,  $\alpha_1^*$  y  $\alpha_2^*$ ) se calcula

$$h_t^* = \alpha_0^* + \alpha_1^* y_{t-1}^2 + \alpha_2^* y_{t-2}^2 \quad (66)$$



**Figura 22:** a) Iteraciones  $\alpha_0$  b) Distribución posterior  $\alpha_0$  c) Iteraciones  $\alpha_1$  d) Distribución posterior  $\alpha_1$  e) Iteraciones  $\alpha_2$  f) Distribución posterior  $\alpha_2$

La figura (23) muestra un comparativo entre los métodos utilizados para estimar la volatilidad. La importancia de los intervalos de credibilidad radica en que brindan una medida más informativa de la incertidumbre en comparación de cuando se usan estimadores puntuales como los máximo verosímiles.

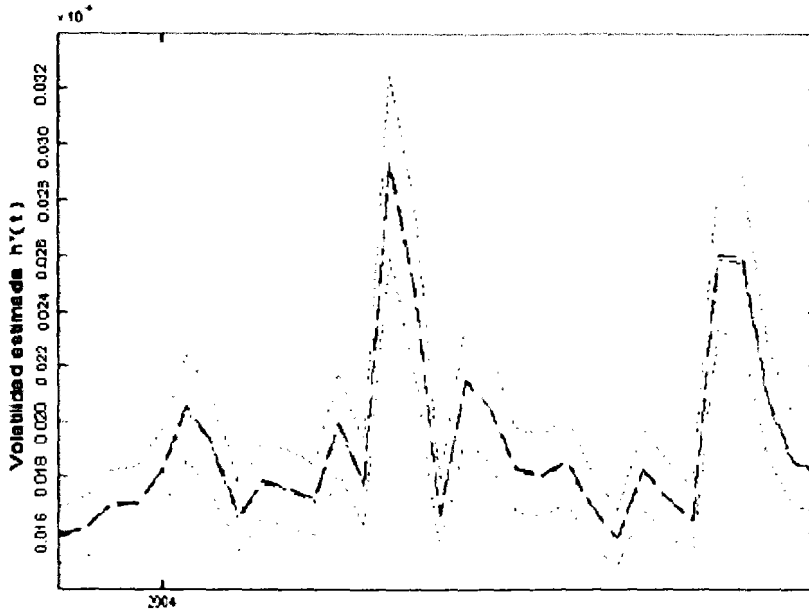
Después se sigue con la búsqueda de la muestras que compondrán la distribución posterior de la volatilidad. Como se hizo anteriormente

$$\alpha_0^D = \begin{pmatrix} \alpha_0^{(75001)} \\ \vdots \\ \alpha_0^{(100000)} \end{pmatrix}^t \quad \alpha_1^D = \begin{pmatrix} \alpha_1^{(75001)} \\ \vdots \\ \alpha_1^{(100000)} \end{pmatrix}^t \quad \alpha_2^D = \begin{pmatrix} \alpha_2^{(75001)} \\ \vdots \\ \alpha_2^{(100000)} \end{pmatrix}^t$$

y después

$$h_t^D = \alpha_0^D + \alpha_1^D y_{3686}^2 + \alpha_2^D y_{3685}^2 \quad (67)$$

La ecuación (67) se calcula de igual forma que (64) solo que el tamaño de observaciones de la serie histórica del IPC es de 3688, esto hace que el tamaño de la serie



**Figura 23:** Estimador Máxima Verosimilitud (línea punteada negra) Media posterior (línea punteada azul) Intervalos de credibilidad al 95 % (líneas punteadas rojas)

de retornos sea de 3687.

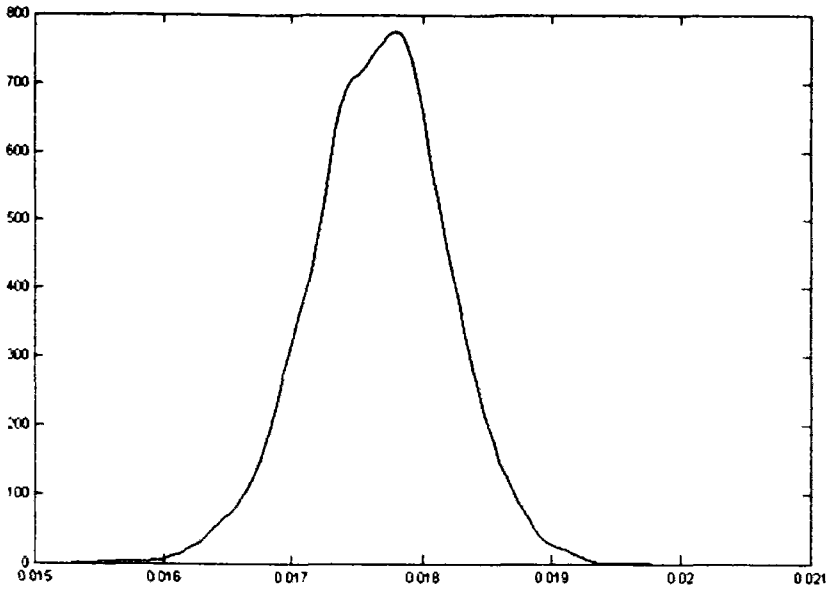
La figura (24) muestra el kernel suavizado de la distribución posterior de la volatilidad de la ecuación (67).

A partir de las muestras que generan (24) se obtienen las muestras de la distribución posterior del precio de una opción call y put mediante el modelo Black & Scholes. Las figuras (25) (26) muestran dichas distribuciones.

Los datos utilizados para el cálculo del modelo Black & Scholes se encuentran el cuadro (6).

Los datos del cuadro (6) y la forma de negociación de las opciones sobre el IPC tienen fundamento en los Términos y Condiciones Generales de Contratación de los Contratos de Opción sobre el Índice de Precios y Cotizaciones de la Bolsa Mexicana de Valores, S.A. de C.V., ver Mexder (2004). Y se especifican de forma muy general de la siguiente forma:

- El tamaño del contrato es de \$ 10 multiplicado por el precio o prima del contrato



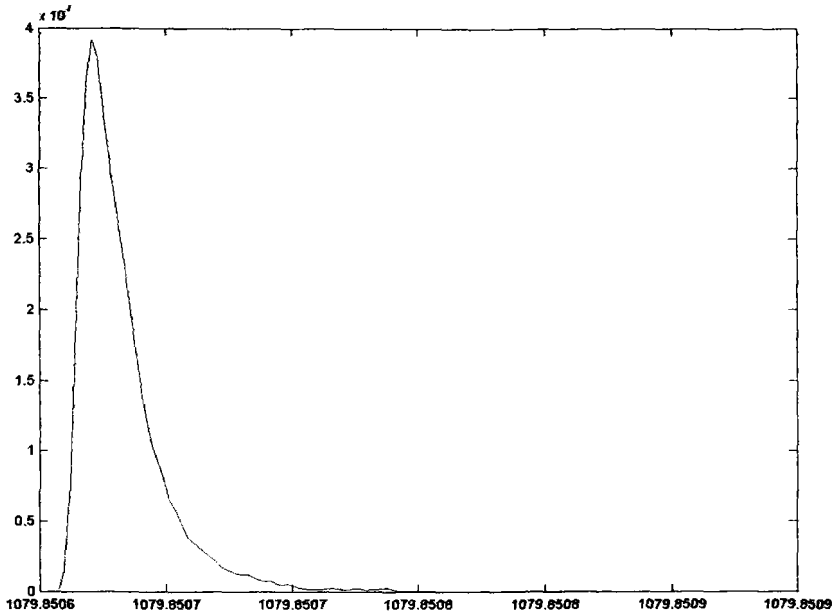
**Figura 24:** Distribución posterior de la volatilidad al 30 de diciembre de 2004

Parámetro	Valor
Precio actual ( $S_0$ )	12918
Precio de ejercicio (strike)	12900
Tasa de interés libre de riesgo	8.59 %
Período de ejercicio	1 año

**Cuadro 6:** Datos reales para el cálculo del precio de compra y precio de venta de una opción para el IPC

de opción.

- El estilo del contrato de opción es europeo. Esto le da mayor fuerza al modelo Black & Scholes propuesto en el presente trabajo.
- Los precios de ejercicio se expresan en puntos enteros del IPC y serán múltiplos de 50 puntos más próximo al último valor del índice del día hábil inmediato anterior.
- La unidad de cotización de la prima del contrato de opción estará expresada en puntos índice.

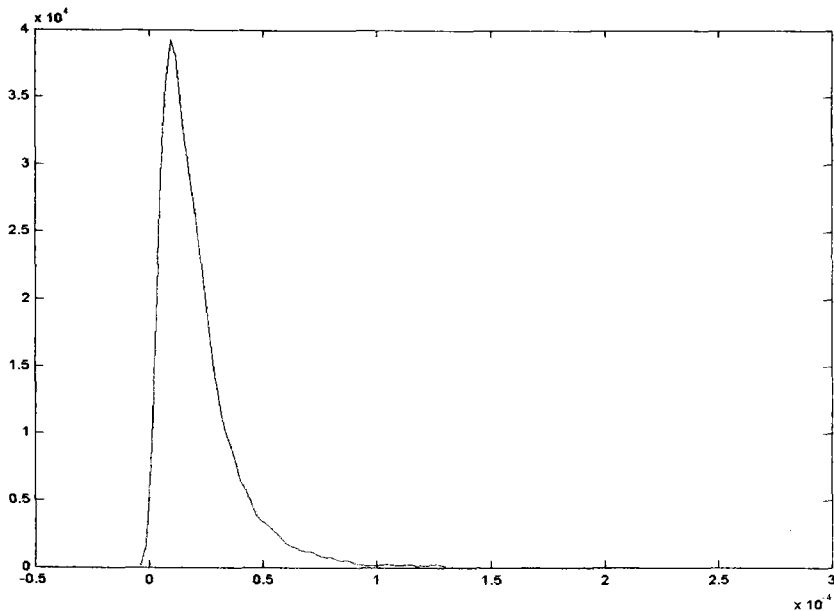


**Figura 25:** Distribución posterior de una opción de compra para el IPC contratada el 31 de diciembre de 2004 vía Black & Scholes

- Un contrato de opción de compra tendrá valor intrínseco cuando el precio de ejercicio sea inferior al precio de cierre del IPC que dé a conocer la BMV, en la fecha de vencimiento. Un contrato de opción de venta tendrá valor intrínseco cuando el precio de ejercicio sea inferior al precio de cierre del IPC que dé a conocer la BMV, en la fecha de vencimiento. En los casos contrarios el valor intrínseco al vencimiento será de cero.
- La fecha de vencimiento se realiza en períodos mensuales.
- La tasa de interés libre de riesgo corresponderá a los Cetes a 28 días del día en que se efectúa la negociación, será determinada mediante la información del Banco de México.

#### 4.3.3. Interpretación de resultados

Como ya se comentó, las figuras (25) y (26) representan las distribuciones posteriores para el precio de una opción de compra y venta vía Black & Scholes, ahora es necesario saber cómo interpretar los resultados.



**Figura 26:** Distribución posterior de una opción de venta para el IPC contratada el 31 de diciembre de 2004 vía Black & Scholes

Para esto se debe de entender el mecanismo de utilización de las opciones financieras desde un punto de vista estratégico. Volviendo a comentar la definición de una opción call de compra que especifica que es el derecho de poseer un determinado subyacente a una fecha de vencimiento a un precio strike por lo que se pagará una prima. Dicho de otra forma a través de un contrato de opción y a diferencia de una acción normal, no se es poseedor del subyacente sino sólo del derecho sobre el bien, el cual andando el tiempo se verá si se interesa o no ejercerlo.

El punto de interés es cuando se debe de ejercer el derecho. Se especifica el *Break Even* como el punto de inicio de ganancias y para una opción de compra estará definido como el precio strike más la prima de la opción. Si el precio del subyacente a la fecha de vencimiento está por arriba del *Break Even* entonces existirá ganancia si está por abajo la pérdida será la prima. Esto se explicó en el primer capítulo.

Con los datos mostrados en el cuadro (6) se calcula el valor de la opción de compra y venta vía Black & Scholes utilizando el concepto de volatilidad constante que se

obtiene a partir del cálculo de la varianza de toda la serie de retornos del IPC y el cual es de 1.65%. Los valores son:

call	1079.9 pts
put	.0000022073 pts

Esto significa que el *Break Even* para el inversionista que asume una posición call largo estará en 13979 pts (strike + call). Ahora, el problema a discutir es la medida de probabilidad que tiene el *Break Even* a través del valor de la prima, es decir conocer la incertidumbre de que el IPC esté por arriba o por debajo del *Break Even*. La desventaja de utilizar el modelo original de Black & Scholes es que no brinda ninguna información acerca de dicha incertidumbre, de hecho la probabilidad de que el IPC esté en los 13979 pts resulta ser cero ya que Black & Scholes sugiere que el precio de las acciones se mueve de manera continua y una probabilidad puntual será siempre cero.

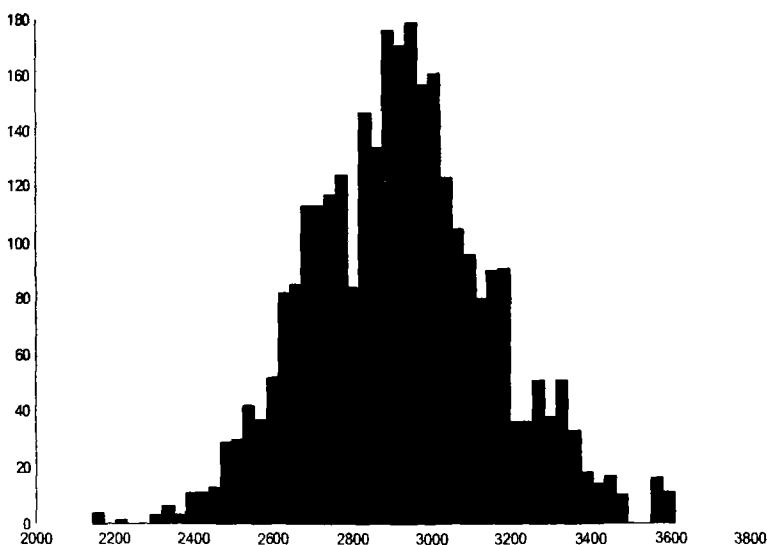
Con la implementación de métodos MCMC en la volatilidad del modelo de Black & Scholes se supera ésta falta de información de la incertidumbre. Con las distribuciones posteriores mostradas en las figuras (25) y (26) ya es posible vencer esta adversidad y entonces se pueden especificar intervalos de credibilidad para el valor de la prima.

Un ejemplo de ello puede ser que a partir de la distribución posterior del call (25) se especifican intervalos de credibilidad al 95%. Esto es que la prima del IPC tiene una probabilidad del 95% de encontrarse entre 1079.85062 pts y 1079.85075 pts. Esto se refleja mejor en un resultado pensado de la siguiente manera. El *Break Even* para un inversionista que el 31 de diciembre de 2004 adquiriera un contrato de opción de compra para el IPC y con fecha de vencimiento de un año a partir de ese mismo día estará entre 13979.85062 y 13979.85075 con una probabilidad del 95%.

Además el método MCMC propuesto en este trabajo permite actualizar la volatilidad en tiempo real, entonces se pueden encontrar distribuciones posteriores para el call y el put en cada cálculo de prima. En el ejemplo anterior la diferencia entre el cuantil .975 y .025 es pequeña ya que la volatilidad de la serie de retornos del IPC 1990-2004 es también pequeña, sin embargo para subyacentes con alta volatilidad o en momentos de alta volatilidad este rango crecerá y tomará una mayor fuerza esta información de la incertidumbre.

**ESTA TESIS NO SALE  
DE LA BIBLIOTECA**

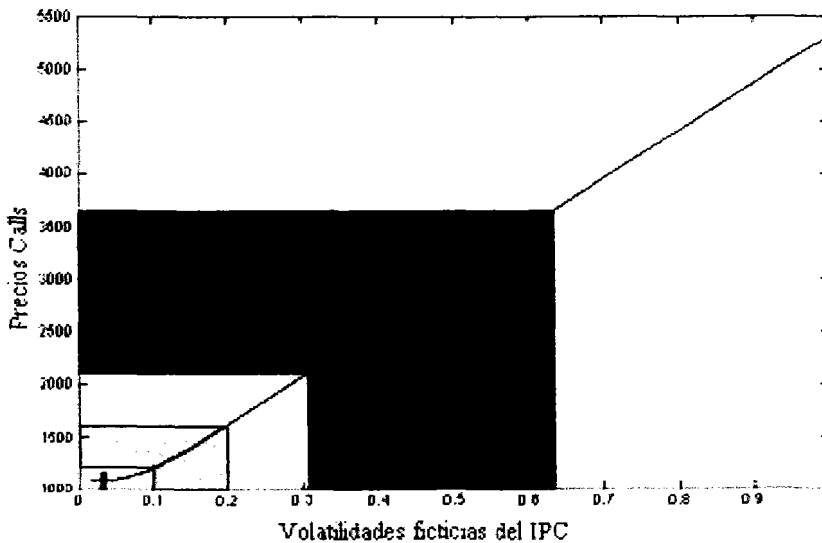
Para entender este hecho, antes que nada, es importante mencionar el modelo Black & Scholes por como está formulado no responde a variabilidades pequeñas de la volatilidad, es decir que el precio call o put no tiende a variar con movimientos pequeños en la volatilidad del subyacente, esto limita mucho al modelo porque en el mundo real es difícil encontrar series financieras que se sujeten a cambios tan bruscos en la volatilidad, para ejemplificar este detalle obsérvese la escala de la volatilidad en la figura (17) y compárese con la misma escala de la figura (23). A partir de los datos simulados se obtiene como ya se ha explicado la distribución posterior para el call. En la figura (27) se puede observar que la escala para el precio del call ha crecido con respecto a la mostrada por la figura (25), y aunque el precio de la opción de compra ha sido calculada con los mismos parámetros mencionados en el cuadro (6) a excepción, claro, de la volatilidad que se implementó a partir de datos simulados como se explicó en la sección (4.3.1), el tamaño intercuantil ha cambiado notablemente. Sólo de esta forma se puede distinguir de buena forma la importancia del uso de una distribución posterior para el precio de una opción call vía Black & Scholes, mejorando la información de la incertidumbre en el *Break Even* del IPC.



**Figura 27:** Distribución posterior de una opción de compra utilizando los datos simulados de la sección (4.3.1) vía Black & Scholes



Para ejemplificar más este importante detalle obsérvese la figura (28). Lo que se hizo fue calcular el valor del call vía el modelo original Black & Scholes mediante los mismos parámetros que los mostrados en el cuadro (6) solamente cambiando la varianza en el modelo. Se utilizaron varianzas ficticias desde el 0% hasta el 100% con el propósito de estrezar al modelo para averiguar la elasticidad entre el call y la varianza del IPC.



**Figura 28:** Inspección del efecto de los cambios en la volatilidad del IPC sobre su precio Call utilizando los parámetros del cuadro (6)

El rectángulo en color rojo representa el espacio ocupado por las volatilidades tal como se obtuvieron en el presente trabajo, es decir en la figura (24) se puede ver que el rango de la muestra posterior de la volatilidad al 30 de diciembre de 2004 es de 0.005 -(0.015,0.02)-. Dicho rango y su efecto sobre los precios calls está expresado en dicho rectángulo en color rojo. Aquí se obtiene el mismo resultado de una variación casi nula, representada por un comportamiento horizontal en los precios de compra de la opción. Lo anterior es lo que le da firmeza a los resultados obtenidos en la figura (25). El rectángulo en color lila representa las volatilidades que se simularon en la sección (4.3.1) cuya muestra posterior de probabilidad de sus calls se encuentra en la figura (27). Sin embargo decir que un subyacente tiene variaciones entre el 31% y el 65% es prácticamente imposible. De hecho, y ya son casos extraños, se han llegado

a encontrar activos financieros que suelen tener variaciones entre 10% y 20% de su valor nominal. Dicho caso se encuentra expresado por el rectángulo en gris.

#### 4.3.4. Out-of-Sample

El estudio *Out-of-Sample* se refiere a la realización de pruebas hacia atrás en el tiempo para verificar si un modelo funciona correctamente. En este caso en particular se presentan diversos resultados tomados en momentos específicos del tiempo para corroborar y analizar el funcionamiento de las muestras posteriores de probabilidad de los precios de compra de las opciones.

Como ya se mencionó, las muestras posteriores de la volatilidad y del precio de compra de una opción al 30 de diciembre de 2004 se calcularon vía MCMC a través de datos obtenidos durante 15 años (1990-2004). Nuestro primer estudio *Out-of-Sample* comprende un periodo menor para que se le pueda dar una mayor movilidad a la volatilidad del IPC, dicho periodo comprende solamente un año (02-01-03 al 31-12-03). Este intervalo se compone de 250 datos (año financiero).

El objeto de este primer estudio es el determinar si el avance sobre la volatilidad en el modelo Black & Scholes sigue cubriendo al resultado que se obtendría con el modelo original y también inspeccionar, cuan largo o cuan corto debe ser el plazo de tiempo para que las muestras de la distribución posterior del call funcionen en plenitud.

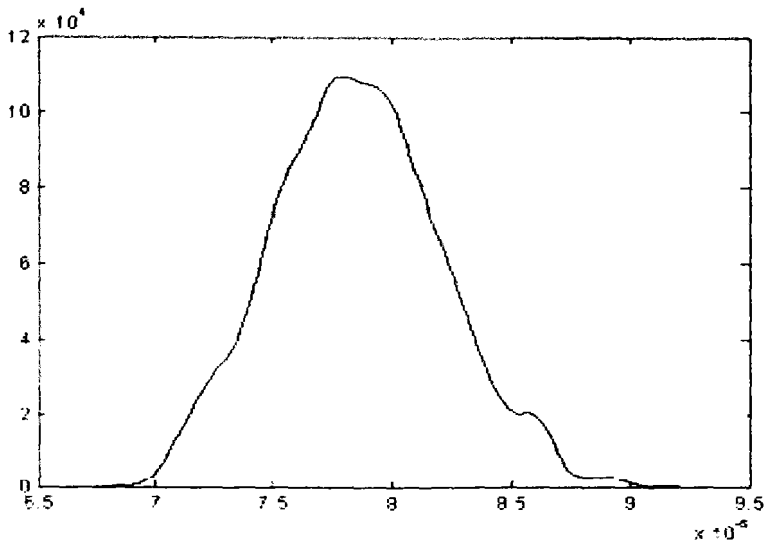
En la figura (29) se observa el kernel posterior del modelo ARCH(2). A continuación se presenta en la figura (30) el kernel posterior de una opción de compra utilizando los siguientes datos:  $S_0=8818.19$ ,  $X=8800$ ,  $r=0.0859$  y  $t=1$ .

Utilizando el modelo original Black & Scholes con los datos anteriores se obtiene que el valor del call es de 742.36335 puntos índice del IPC. En la figura (30), la línea vertical de color rojo representa este valor, por tanto las muestras de la distribución posterior del call cubren al valor real.

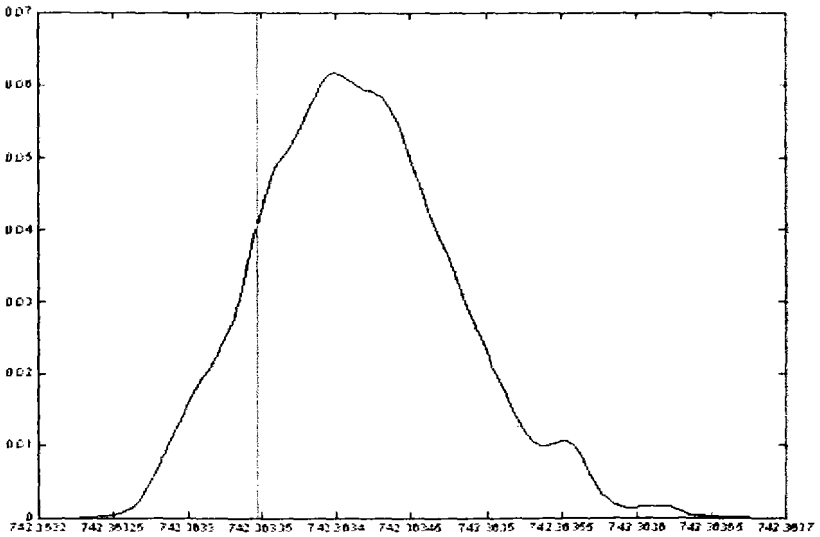
Ahora bien, el intervalo de credibilidad al 95% del *Break Even*, de acuerdo con las muestras de distribución posterior obtenidas, se compone entre los cuantiles 9542.3633 y 9542.36355.

En este ejemplo el contrato de compra de la opción se pactó a un año. El IPC al día 31 de diciembre de 2004 fue de 12918 puntos índice. Como se puede observar el pronóstico Black & Scholes estuvo muy por debajo y eso acarreó que nuestro intervalo de credibilidad al 95% también lo estuviera.

Sin embargo un plazo de un año podría ser demasiado largo para poder haber

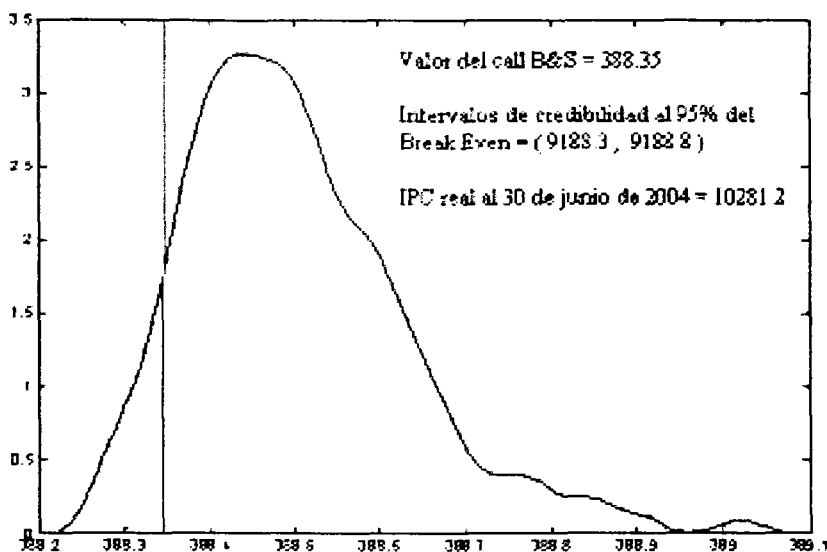


**Figura 29:** Distribución posterior de la volatilidad al 31 de diciembre del 2003



**Figura 30:** Distribución posterior de una opción de compra para el IPC contratada el 31 de diciembre de 2003 vía Black & Scholes

obtenido un buen pronóstico del *Break Even* debido a que mientras más largo sea el horizonte de pronóstico más incertidumbre habrá. Lo que a continuación se presenta son pruebas de un horizonte de tiempo más corto. Las figuras (31) y (32) expresan las distribuciones posteriores para el valor del call utilizando los mismo datos anteriores pero a 6 meses y un mes respectivamente.

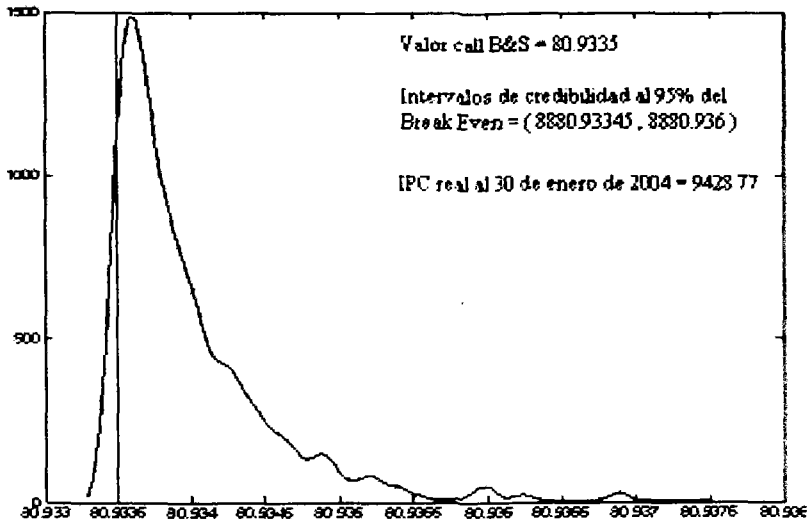


**Figura 31:** Distribución posterior de una opción de compra para el IPC contratada el 31 de diciembre de 2003 vía Black & Scholes. Fecha de expedición de 6 meses

Como se puede ver en ambas figuras, la diferencia entre el *Break Even* y el valor real del IPC disminuye mientras más corto es el horizonte de pronóstico.

El segundo estudio *Out-of-Sample* se refiere al hecho de la elección de un cluster de alta volatilidad en la serie histórica del IPC con el propósito de realizar un estudio acerca del aumento de la volatilidad y su efecto en la distribución posterior del precio del call negociado a corto plazo para poder determinar si la propuesta actual al modelo Black & Scholes funcionaría de manera exitosa. La muestra tomada es de todo el año 2000.

La figura (33) representa el comportamiento del IPC a lo largo del 2000, su serie de retornos y el modelo ARCH(2) estimado por máxima verosimilitud.

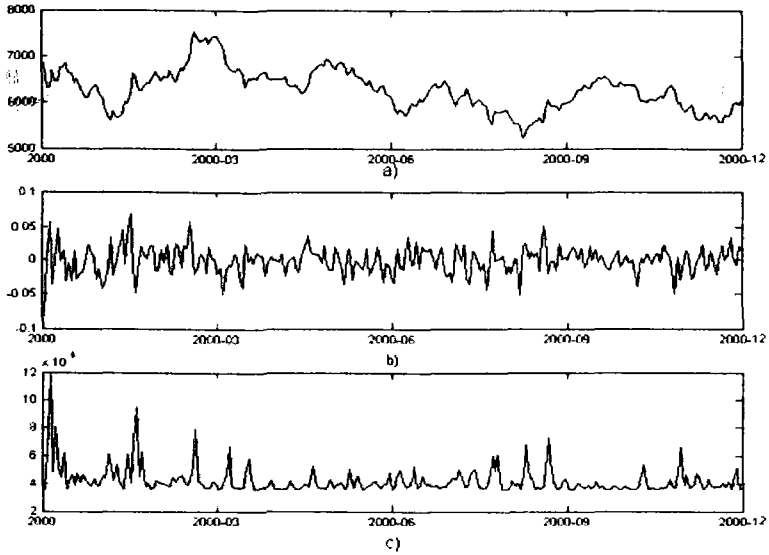


**Figura 32:** Distribución posterior de una opción de compra para el IPC contratada el 31 de diciembre de 2003 vía Black & Scholes. Fecha de expedición de un mes

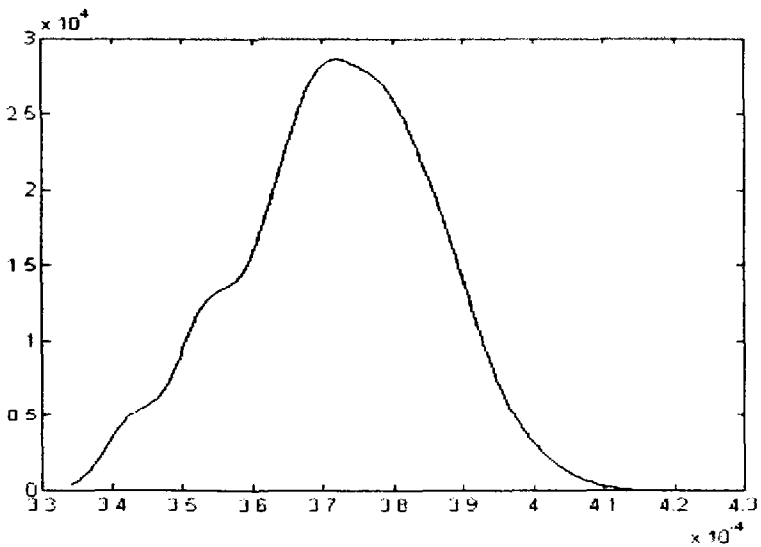
En la figura (33) se debe de observar que la escala de la volatilidad ha crecido con respecto a la mostrada por la figura (21), esto es debido al cluster de volatilidad tomado. Las figuras (34) y (35) muestran las distribuciones posteriores de la volatilidad y del precio de compra de la opción sobre el IPC respectivamente.

De la figura (35) se puede observar que existe una mayor distancia intercuantil entre los precios de compra, esto es repercusión del aumento de la volatilidad del IPC durante el 2000.

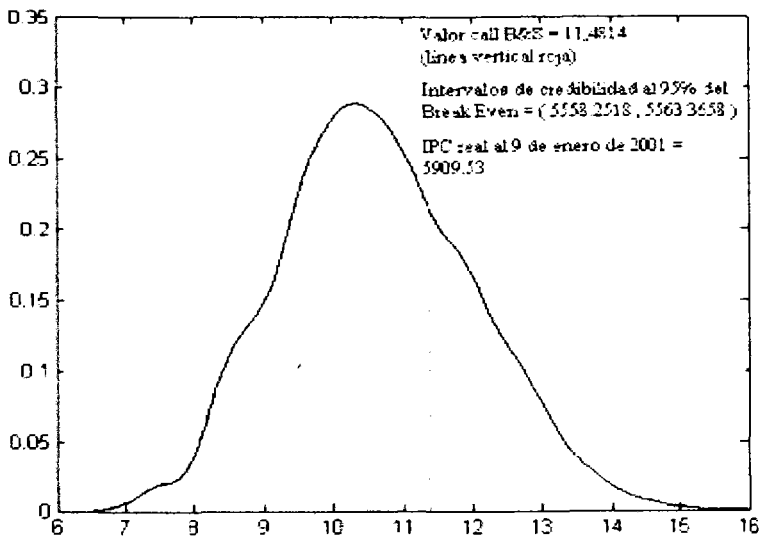
A través del estudio *Out-of-Sample* mostrado se determinan un par de cosas importantes acerca del modelo, la primera es que las muestras de la distribución posterior del precio de compra de una opción pactada siempre cubrirá al mismo valor obtenido directamente del modelo Black & Scholes, y segundo, cuando existan clusters de alta volatilidad el tamaño intercuantil de la distribución posterior del precio de compra de la opción crecerá.



**Figura 33:** a)IPC 2000 b) Serie de retornos c) Volatilidad estimada por máxima verosimilitud



**Figura 34:** Distribución posterior de la volatilidad al 14 de diciembre de 2000



**Figura 35:** Distribución posterior de una opción call vía Black & Scholes negociada el 15 de diciembre de 2000. Fecha de expedición 9 de enero de 2001 (15 días)

## Capítulo V

### CONCLUSIONES

En este trabajo se exhibe al modelo Black & Scholes como el responsable de la evolución de la ingeniería financiera, no obstante se legitimó la veracidad de su supuesto acerca de la volatilidad. Por otra parte se discute el cimiento y el funcionamiento de la inferencia bayesiana y los métodos MCMC así como el aprovechamiento de estas herramientas estadísticas en las series de tiempo financieras. A partir de ello se muestra como superar la volatilidad constante en el modelo original debatido, hallando muestras de la distribución posterior del modelo ejecutado de serie de tiempo para volatilidad. Aún así la tesis no culmina con esta novedad, fue de interés particular el poder encontrar muestras de la distribución posterior del precio de compra o venta de una opción mediante el modelo Black & Scholes, lo que nunca había sido especificado de esta manera y que propone ubicar en la vanguardia financiera a un modelo tan prominente como resulta serlo.

Uno de los aspectos más notables que el presente trabajo deja, es la posibilidad de generar conocimiento en la aplicación de técnicas bayesianas tanto en modelos financieros como en modelos econométricos o macroeconómicos, implementando métodos MCMC que si bien no han sido tan explotados en la investigación tienen un sin fin de aplicaciones. En su momento se expusieron de manera muy breve un par de ejemplos de ello.

Es indispensable mencionar que la investigación actual abre posibilidades de perfeccionamiento, ya que en ningún momento se ha indicado que la actual propuesta establecida sea claramente la mejor, solo se ha realizado un ligero progreso sobre dos aspectos cruciales en el modelo financiero de fijación de primas para opciones. Si bien se propone el aprovechamiento de un modelo ARCH para explicar la volatilidad de los subyacentes financieros, podría mejorarse con la implementación de modelos más completos como los GARCH y sus generalizaciones. Por otra parte se utiliza el modelo ARCH con una suposición de normalidad en  $\epsilon_t$  lo cual hace al modelo relativamente sencillo, sin embargo esta suposición acarrea distintas dificultades estadísticas y es por ello que en la actualidad se comienza por utilizar cada vez más modelos ARCH



con distribuciones t-student. Otra de las mejoras que podrían realizarse son formas más completas de realizar inferencia bayesiana sobre el movimiento de los valores de las opciones, es decir suponer que el precio de las opciones está dada por el modelo Black & Scholes y a partir de éste, pronosticar los valores futuros anexando un término de error a dicho modelo.

Por último, se sugiere que la curva de aprendizaje de los métodos MCMC deba ser llevada a cabo de manera gradual de tal manera que su entendimiento madure con el tiempo. Su implementación en un lenguaje de computación o software especializado debería ser la culminación de la investigación. Dicha implementación fue desarrollada bajo la plataforma de computación técnica Matlab.

## REFERENCIAS

- [1] BAUWENS, L., (1998). *Bayesian inference on GARCH models using the Gibbs sampler*. *Econometrics Journal*, Vol.1. Pág C23-C46.
- [2] BERGER, J., (1988). *Statistical Decision Theory and Bayesian Analysis*. Springer Series in Statistic, Nueva York.
- [3] BOLLERSLEV, T., (1986). *Generalized Autoregressive Conditional Heteroskedasticity*. *Journal of Econometrics*, Vol.31. Pág 307-327.
- [4] CASELLA, G., (1999). *Monte Carlo Statistical Methods*. Springer Series in Statistic, Nueva York.
- [5] CASELLA, G., (1992). *Explaining the Gibbs Sampler*. *The American Statistician*, Vol.46. Pág.167.
- [6] CHATFIELD, C., (1989). *The Analysis of Time Series. An Introduction*. Chapman & Hall, Gran Bretaña.
- [7] CHEN, M.-H., (2000). *Monte Carlo Methods in Bayesian Computation*. Springer Series in Statistics, Nueva York.
- [8] CHIB, S., (1995). *Understanding the Metropolis-Hastings Algorithm*. *The American Statistician*, Vol.49. Pág 327-335.
- [9] CHIB, S., (2001). *Markov chain Monte Carlo methods: Computation and inference*. *Handbook of Econometrics*, Volume 5, pp. 3569-3649. Amsterdam: North-Holland.
- [10] DACOROGNA, M., (2001). *An introduction to high-frequency finance*. Academic Press, San Diego, Cal.
- [11] DIGGLE, J., (1990). *Time Series. A Biostatistical Introduction*. Oxford Press, E.U.A.
- [12] ENGLE, R., (1995). *ARCH selected readings. Advanced text in econometrics*. Oxford University Press, E.U.A.
- [13] ENGLE, R., (1982). *Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of U.K. Inflation*. *Econometrica*, Vol.50. Pág 987-1008.
- [14] FLORENTINI, G., (2002). *Likelihood-based estimation of latent generalised ARCH structures*. Unpublished paper. Universidad de Alicante, Italia.

- [15] FRIEDMAN, M., (1977). *Nobel lecture. Inflation and unemployment*. Journal of political economy, Vol 85. 451-472.
- [16] GAMERMAN, D., (1997). *Markov Chain Monte Carlo. Stochastic simulation for Bayesian inference*. Chapman & Hall, Londrés.
- [17] GELFAND, A., (1990). *Sampling-based approaches to calculating marginal densities*. Journal of the American Statistical Association, Vol.85. Pág 398-409.
- [18] GELFAND, A., (1992). *Bayesian Analysis of constrained parameters and truncated data problems using Gibbs Sampling*. Journal of the American Statistical Association, Vol.87. Pág 523-532.
- [19] GEWEKE, J., (1989). *Bayesian Inference in Econometric Models Using Monte Carlo Integration*. Econometrica, Vol.57. Pág 1317-1339.
- [20] GHYSELS, E., (2001). *The Econometric Analysis of Seasonal Time Series*. Cambridge University Press, Cambridge.
- [21] GOURIÉROUX, C., (1997). *ARCH Models and Financial Applications*. Springer Series in Statistic, Nueva York.
- [22] HULL, J., (2000). *Option, Futures and other Derivatives*. Prentice Hall, Canadá.
- [23] JARRW, (1983). *Option pricing*. Irwing, E.U.A.
- [24] JOHANNES, M., (2003). *MCMC Methods for Continous-Time Financiál Econometrics*. To appear in Handbook of Financial Econometrics., 2003.
- [25] LEE, B., (1989). *Bayesian Statistic: An Introduction*. Oxford University Press, Reino Unido, 1989.
- [26] LUENBERGER, D., (1999). *Investment Science*. Oxford University Press, Nueva York.
- [27] MEXDER, (2004). *Términos y Condiciones Generales de Contratación de los Contratos de Opción sobre el Índice de Precios y Cotizaciones de la Bolsa Mexicana de Valores, S.A. de C.V. Mercado Mexicano de Derivados, México D.F.*
- [28] MEYN, S., (1994). *Computable bounds for convergence rates of Markov Chain*. Annals of Applied Probability, Vol.4. Pág 124-148.
- [29] MITTELHAMMER, R., (2000). *Econometrics Foundations*. Cambridge University Press, Nueva York.
- [30] PEDERSEN, J., (2004). *Markovkæde Monte Carlo(MCMC) med Metropolis-Hastings algoritmen*. Stockholm School of Economics, Technical Report.
- [31] PINDYCK, R., (1998). *Econometric Models and Econometric Forecasts*. McGraw-Hill, E.U.A.

- [32] ROSS, S., (2000). *Probability Models*. Academic Press, San Diego, Cal.
- [33] ROSS, S., (1999). *Finanzas corporativas*. McGraw-Hill, México.
- [34] SABAU, H., (1997). *Derivados financieros*. Operadora de bolsa Serfín, México.
- [35] STAMPFLI, J., (2002). *Matemáticas para las finanzas*. Thomson, México.
- [36] TINOCO, D., (2002). *Futuro y opciones financieras*. Editorial Limusa, México.
- [37] TSAY, R., (2002). *Analysis of Financial Time Series. Financial Econometrics*. John Wiley & Sons, E.U.A.
- [38] VILLAGRÁN, A., (2003). *Modelos Mezcla para Volatilidad. Tesis de maestría en estadística*. Centro de Investigaciones en Matemáticas. Universidad de Guanajuato., México.
- [39] VILLAGRÁN, A., (2005). *Bayesian Inference on Mixture-of-Expert for Estimation of Stochastic Volatility*. To appear in *Advanced in Econometrics*, Volumen in honor to Engle and Granger.
- [40] WEI, W., (1989). *Time Series Analysis. Univariate and Multivariate Methods*. Addison-Wesley, E.U.A.
- [41] Y TIAO, B., (1970). *Bayesian inference in statistical analysis*. Wiley Classics Library, E.U.A.
- [42] ZHANG, X., (2003). *Estimation of Asymmetric Box-Cox Stochastic Volatility Models Using MCMC Simulation*. Monash University. Department of econometrics and business statistic, Working paper, Australia.