

LA INFERENCIA BAYESIANA EN LA ADMINISTRACIÓN DE RIESGOS¹

Carlos Vladimir Rodríguez-Caballero²

Resumen

Cuando se realiza inferencia a través de enfoques bayesianos la mayoría de las veces se recurre a formar distribuciones posteriores conjugadas por su facilidad de manejo e integración, sin embargo estas formas analíticas cerradas no son posibles alcanzarlas en la mayoría de los modelos econométricos o financieros usados comúnmente en la práctica de la administración de riesgos, por lo que los algoritmos de simulación estocástica, como son los algoritmos Markov Chain Monte Carlo (MCMC) deben ser implementados. En este trabajo se utiliza el algoritmo Metropolis-Hastings y un algoritmo diseñado para simular una cadena de Markov con saltos denominado Reversible Jump MCMC (RJMCMC).

En el presente artículo, se encuentran muestras de la distribución posterior de un modelo de volatilidad propuesto y se ejemplifican posibles implementaciones en dos modelos usados frecuentemente en el ejercicio del riesgo de crédito y riesgo de mercado como son el modelo Black & Scholes y el valor en riesgo (VaR).

Finalmente se muestra una propuesta bayesiana para el desarrollo de un scoring de crédito, otro punto central en la administración de riesgo actual.

Palabras clave: MCMC, RJMCMC, ARCH, scorecard, mezclas finitas

¹ Este artículo es una versión preliminar del publicado.

² Estudiante de primer semestre de la maestría en economía (econometría) en el departamento de Economía y Finanzas. Universidad de Guanajuato. E-mail: vladimir.rodriguez@ugto.org.

El autor agradece a la Universidad de Guanajuato por el apoyo recibido al momento de presentar este trabajo en el II Foro de Finanzas y Administración de Riesgos realizado en septiembre de 2009 en la Universidad Autónoma de México.

I. Introducción

Las técnicas estadísticas o econométricas han sido usadas dentro de la industria financiera para dar soporte en distintas facetas de la administración de riesgos como pueden ser la originación de crédito vía modelos de scorecard, controlando la calidad crediticia conforme a las metodologías de probabilidad de incumplimiento y pérdida esperada, fijando límites de riesgo de crédito vía los modelos VaR o haciendo distintos modelajes en el manejo del portafolio de series accionarias o índices bursátiles. El modelo Black & Scholes ha sido usado enormemente para el cálculo de primas de opciones y planeación de inversiones. Por otro lado, desde hace algunos años, la econometría financiera se revolucionó en gran medida gracias a los avances sobre modelación no lineal, tal como la familia de modelos de heteroscedasticidad condicional y los modelos de volatilidad estocástica.

En la última década se han desarrollado bastantes propuestas tanto teóricas como aplicadas en este sector, algunas de ellas en los acuerdos de Basilea I y II en donde se cimentaron las bases para la administración integral de riesgos. Siendo a partir de dichos acuerdos que las instituciones bancarias a lo largo del mundo empezaron a desarrollar modelos estadísticos o matemáticos para el control del riesgo de crédito.

Sin embargo, dichos estudios en su forma original, encuentran solamente estimaciones puntuales en los parámetros del modelo en cuestión y por consiguiente solo brindan información puntual acerca del problema sin hacer mención alguna acerca de la probabilidad asociada.

Uno de los principales intereses en la modelación actual de la administración de riesgos es el buscar robustecer los modelos ya existentes. Aprovechando el paradigma bayesiano, donde se valora más la información a través de la probabilidad de un parámetro, pronóstico

o estadístico que simplemente a través de la estimación puntual del mismo, es posible alcanzar una mejora sustancial en la modelación de riesgos.

Al realizar inferencia a través de enfoques bayesianos la mayoría de las veces se recurre a formar distribuciones posteriores conjugadas por su facilidad de manejo, sin embargo estas formas analíticas cerradas no son posibles alcanzarlas en la mayoría de los modelos en la administración de riesgos, entonces los algoritmos de simulación estocástica, entre los que destacan los mecanismos MCMC tienen que ser implementados.

En el presente trabajo se muestran dichos mecanismos al modelar fenómenos diversos concernientes a la administración de riesgos y se desarrolla de la siguiente forma: en la segunda sección se muestra un esbozo del paradigma bayesiano y se plantean los algoritmos esenciales para el desarrollo de las simulaciones que se usarán más adelante, en la tercer parte del trabajo se obtienen muestras de la distribución posterior de la volatilidad de un modelo sencillo de heteroscedasticidad condicional (ARCH) y sus posibles implementaciones en el modelo Black & Scholes y VaR, en la cuarta sección se ejemplifica una metodología bayesiana para la modelación de un scoring de crédito usando mezclas finitas de poissones y simulación RJMCMC para lograr el objetivo, finalmente se dan las conclusiones y extensiones del trabajo y la bibliografía correspondiente.

II. Simulación estocástica para inferencia bayesiana

Esta sección se enfoca en mostrar los resultados teóricos más importantes tanto en la inferencia bayesiana como en la simulación estocástica necesaria para muestrear las distribuciones posteriores de probabilidad.

No se busca presentar toda la teoría existente ya en libros de texto porque no es la finalidad de la sección, sin embargo se marcarán bibliografías para que aquel lector que quisiese ahondar en la teoría pueda hacerlo.

A) Elementos de inferencia bayesiana. El teorema de Bayes

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.

Para una discusión profunda acerca del papel que juega la inferencia bayesiana en el análisis estadístico, ver Box & Jenkins (1970).

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 los k parámetros $\theta = \{\theta_1, \dots, \theta_k\}$. Supóngase también que θ tiene por sí mismo una función de distribución $p(\theta)$. Entonces,

$$p(y|\theta)p(\theta) = p(y, \theta) = p(\theta|y)p(y) \quad (1)$$

Dado los datos observados y , la distribución condicional de θ es $p(\theta|y) = \frac{p(y|\theta)p(\theta)}{p(y)}$,
alternativamente

$$p(\theta|y) \approx L(\theta|y)p(\theta) \quad (2)$$

A $p(\theta)$ se le conoce como distribución a priori y a $p(\theta|y)$ distribución posterior de probabilidad. La función de verosimilitud $L(\theta|y)$ 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 θ .

En la inferencia bayesiana cuando se carece de todo conocimiento acerca de la distribución a priori y es supuesto que esta 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.

B) 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.

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

$$1 - \alpha \leq p(C|x) = \int_C dF^{\pi(\theta|x)} d\theta \quad (3)$$

A partir de que la distribución posterior π es una probabilidad actual en Θ , uno puede hablar significativamente de la probabilidad de que θ 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 θ .

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).

C) 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 convenientes 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. 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 $\theta \in \Theta$ es el vector de parámetros y X son los datos. La idea de la simulación es crear un proceso de Markov en Θ , el cual converja a la distribución estacionaria $p(\theta|X)$.

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.

Se aconseja a aquel lector que no esté familiarizado con teoría de procesos estocásticos, en específico al relacionado con las cadenas de Markov consultar Ross (2000).

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 π se dice que es una distribución estacionaria de una cadena con probabilidades de transición $P(x, y)$ si

$$\sum_{x \in S} \pi(x)P(x, y) = \pi(y), \forall y \in S \quad (4)$$

La ecuación (4) puede ser escrita en notación matricial como $\pi = \pi P$.

Ahora bien, una vez que la cadena alcanza una iteración donde π 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 π 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 π , cuando $n \rightarrow \infty$. En este sentido, a la distribución se le conoce también con el nombre de distribución límite.

Este importante resultado es un resultado brillante que permite usar este proceso estocástico para resolver los problemas encontrados en inferencia bayesiana, la prueba de dicho teorema necesita de la definición de algunas propiedades de una cadena de Markov y conocimientos de teoría ergódica, el teorema y su demostración puede ser encontrado en Ross (2000) y Gamerman (2006), entre otros.

2) Cadenas reversibles

Sea $(\theta^{(n)})_{n \geq 0}$ una cadena homogénea con probabilidades de transición $P(x, y)$ y distribución estacionaria π . Asúmase que se desea investigar la sucesión $\theta^{(n)}, \theta^{(n-1)}, \dots$, en sentido contrario. Se prueba directamente que esta sucesión satisface

$$P(\theta^{(n)} = y | \theta^{(n-1)} = x_1, \theta^{(n-2)} = x_2, \dots) = P(\theta^{(n)} = y | \theta^{(n-1)} = x_1) \quad (5)$$

Y entonces define una cadena de Markov. Las probabilidades de transición $P_n^*(x, y)$ están definidas por

$$P_n^*(x, y) = P(\theta^{(n)} = y | \theta^{(n-1)} = x) = \frac{P(\theta^{(n+1)} = x | \theta^{(n)} = y) P(\theta^{(n)} = y)}{P(\theta^{(n+1)} = x)}$$

$$\frac{\pi^{(n)}(y) P(x, y)}{\pi^{(n+1)}(x)}$$

En general esta cadena no es homogénea. Si $n \rightarrow \infty$, o alternativamente $\pi^{(0)} = \pi$, entonces $P_n^*(x, y) = P^*(x, y)$ y entonces la cadena es homogénea. Si $P_n^*(x, y) = P(x, y) \forall x, y \in S$, la cadena de Markov en sentido inverso tiene las mismas probabilidades de transición que la cadena original. Las cadenas de Markov con esta propiedad se dicen que son reversibles y la condición de reversibilidad es usualmente escrita como

$$\pi(x)P(x, y) = \pi(y)P(y, x) \quad \forall x, y \in S \quad (6)$$

La condición (6) es de suma importancia para los algoritmos MCMC como el Metropolis-Hastings y los algoritmos RJMCMC.

3) Simulación de una cadena de Markov

Una cadena de Markov se dice ergódica si tiene la siguiente propiedad

$$\lim_{n \rightarrow \infty} \|P^n(x, \cdot) - \pi(\cdot)\| = 0, \quad \forall x \in S \quad (7)$$

Para simular una cadena de Markov $(\theta^{(n)})_{n \geq 0}$ se debe de considerar que sea ergódica con espacio de estados $s \subset \mathbb{R}^d$, kernel de transición $P(x, y)$ y distribución inicial $\pi^{(0)}$.

El valor de $\theta^{(1)}$ está distribuido con densidad $P(\theta^{(0)}, \cdot)$ y puede ser generado desde un número aleatorio. Para $\theta^{(2)}$, este procedimiento se repite pero se muestrea desde una densidad $P(\theta^{(1)}, \cdot)$. Iterando este esquema a través de los pasos de la cadena de Markov se llega a muestrear $\theta^{(n)}$ desde una distribución con densidad $P(\theta^{(n-1)}, \cdot) \forall n$.

Conforme el valor de n crece, se llega a estar cada vez más cerca de muestrear desde la distribución límite π y puede ser considerado como un muestreo de π . Es importante notar que todos los valores muestreados de la cadena son alcanzados después de converger y también son muestreados desde π 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 (2006).

D) Algoritmo Metropolis-Hastings

Antes de presentar el algoritmo Metropolis-Hastings es importante comentar que existe un algoritmo particular con el que se recomienda el estudio de simulación MCMC, el algoritmo Gibbs Sampler quien 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. El hecho primordial por el que en este trabajo no es utilizado este algoritmo es que se desconocen las distribuciones condicionales totales de los parámetros del modelo, diferencia circunstancial entre ambos algoritmos mencionados.

Para ver y estudiar con detalle la teoría existente del Gibbs Sampler, así como ejemplos de implementación, ver Casella (1992), Casella (1999) y Gamerman (2006).

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 θ^{m+1} condicional en θ^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 expuesta en (6) quedando de la forma siguiente

$$p(\theta^m|y)q(\theta^{m+1}|\theta^m) = p(\theta^{m+1}|y)q(\theta^m|\theta^{m+1}) \quad (8)$$

Donde el lado izquierdo es la probabilidad incondicional de generar θ^{m+1} dado θ^m , y el lado derecho es la probabilidad incondicional de generar θ^m dado θ^{m+1} . La condición de reversibilidad puede ser impuesta en (8) para inducir a un equilibrio en la ecuación

$$p(\theta^m|y)q(\theta^{m+1}|\theta^m)\alpha(\theta^m|\theta^{m+1}) = p(\theta^{m+1}|y)q(\theta^m|\theta^{m+1}) \quad (9)$$

Donde $\alpha(\theta^m|\theta^{m+1})$ es conocida como la probabilidad de movimiento y está definida como

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

$$si \begin{cases} p(\theta^m|y) q(\theta^{m+1}|\theta^m) > 0 \\ \alpha(\theta^m|\theta^{m+1}) < 1 \end{cases} \quad (10)$$

Si el lado izquierdo en (9) es más grande que el derecho, entonces $\alpha(\theta^m|\theta^{m+1})$ 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 θ^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 θ^1 para generar θ^2 .

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.

Existen dos casos especiales más importantes de este algoritmo, el Metropolis-Hastings de cadenas independientes y Metropolis de caminata aleatoria. En este trabajo se utiliza el de cadenas independientes.

Para revisar la información adicional acerca del algoritmo de Metropolis-Hastings, ver Casella (1999), Chib (1995), Geweke (1989) y Gamerman (2006).

Existen numerosas aplicaciones de simulación estocástica en diversas áreas. Para el área de interés particular, ver Johannes (2003) para ejemplos de aplicación.

E) Simulación RJMCMC

Uno de los desarrollos más importantes en la investigación MCMC fue el propuesto en el artículo seminal Green (1995) en el cual se demostró como los métodos MCMC podrían ser aplicados a una clase de problemas más amplios, incluyendo aquellos donde se desconoce profundamente todo el espacio general del problema. Más formalmente, la generalización encontrada permitió la consideración de problemas estadísticos donde el espacio parametral pudiera ser escrito como la unión de subespacios, cada uno con una posible dimensión. A ésta clase de métodos MCMC fue denominado como simulaciones de saltos reversibles Reversible Jump MCMC (RJMCMC) y en esencia provee la posibilidad de hacer inferencia acerca de modelos estocásticos mucho más generales que aquellos logrados solo con los métodos MCMC tradicionales.

Desde su nacimiento ha sido usado con gran éxito en diversas áreas de conocimiento como son en la biotecnología, la genética, prevención de desastres, etcétera. En temas relacionados a inferencia estadística y mucho más cercano al interés propio del trabajo, los métodos RJMCMC han sido usado en la modelación bayesiana de series de tiempo cuando el orden del proceso estocástico no es conocido y modelos mezcla cuando el número de

componentes de la mezcla no es conocido, caso específico por el cual se debe implementar este método en la propuesta bayesiana del scoring de crédito.

A pesar la enorme funcionalidad de estos métodos, aún no han sido ampliamente divulgados y/o implementados debido a la complejidad computacional que se dice que tienen. Sin embargo no se ha apreciado que los métodos RJMCMC únicamente vuelven más estricto el algoritmo Metropolis-Hastings al permitirle considerar problemas que involucren espacios de estados generales, incluso aún cuando ellos contengan subespacios de dimensiones diferentes.

El enfoque utilizado para analizar estos detalles requiere un tratado especial sobre teoría de la medida, solo así fue posible demostrar el hecho de que los algoritmos RJMCMC no son más que generalizaciones sobre el Metropolis-Hastings en el artículo seminal de Green. Hastie (2005) especifica el hecho de que no es necesario ahondar al respecto para utilizar el método en ejercicios reales y que existen explicaciones más sencillas que las originales, sin embargo el enfoque basado en teoría de la medida juega un rol supremo al entender a detalle el algoritmo. El lector puede consultar Green (1995), Tierney (1998), Green (2003) para más detalles.

El algoritmo se mostrará en específico en la última parte del trabajo cuando se proponga su implementación en un scoring bayesiano.

III. Inferencia bayesiana en la volatilidad

A) Elementos para la simulación MCMC

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 estos 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 (1).

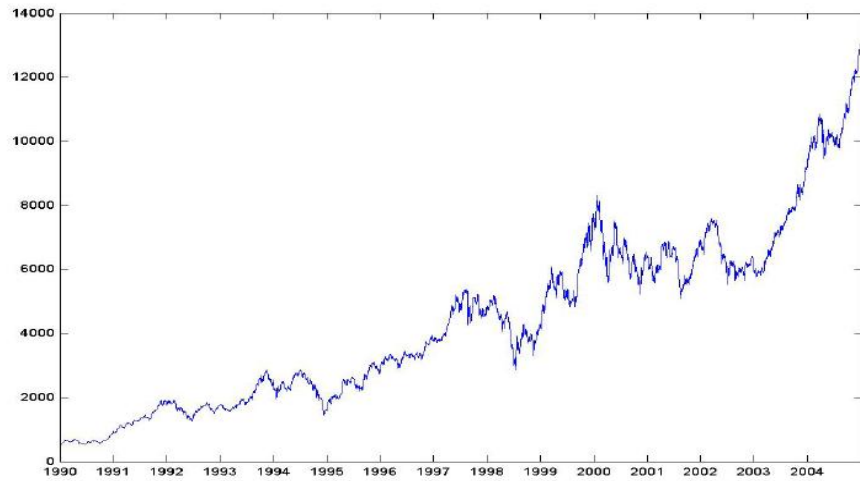


Figura 1. Serie histórica del IPC 1990-2004

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

- Parametrización exacta del modelo.

$$\begin{aligned}
 y_t &= \varepsilon_t \sqrt{h_t} & \varepsilon_t &\sim N(0,1) \\
 h_t &= \alpha_0 + \alpha_1 y_{t-1}^2 + \alpha_2 y_{t-2}^2 \\
 \theta &= (\alpha_0, \alpha_1, \alpha_2)'
 \end{aligned}
 \tag{11}$$

- Función de verosimilitud del modelo.

$$L(\theta|y) = \left(\frac{1}{2\pi}\right)^{\frac{n}{2}} \left[\prod_{i=1}^n \left(\frac{1}{h_t}\right)^{\frac{1}{2}} \right] e^{\left\{-\frac{1}{2} \sum_{i=1}^n \frac{y_t^2}{h_t}\right\}} \quad (12)$$

Para detalles de cálculo, ver Tsay (2002)

Para revisar extensivamente el uso de los métodos MCMC en el contexto de la econometría, ver Chib (2001) y Florentini (2002).

- Distribución inicial del modelo.

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.

$$p(\theta) = p(\alpha_0)p(\alpha_1)p(\alpha_2) = I_{(0,\infty)}^{\alpha_0} I_{(0,1)}^{\alpha_1} I_{(0,1)}^{\alpha_2} \quad (13)$$

- Distribución posterior.

De acuerdo a (2), la componemos mediante la multiplicación de (12) y (13)

$$p(\theta|y) \approx \left(\frac{1}{2\pi}\right)^{\frac{n}{2}} \left[\prod_{i=1}^n \left(\frac{1}{h_t}\right)^{\frac{1}{2}} \right] e^{\left\{-\frac{1}{2} \sum_{i=1}^n \frac{y_t^2}{h_t}\right\}} I_{(0,\infty)}^{\alpha_0} I_{(0,1)}^{\alpha_1} I_{(0,1)}^{\alpha_2} \quad (14)$$

Como se observa en (14), la distribución posterior no tiene una forma cerrada, por lo que se necesitan de los siguientes elementos para poder construir la simulación MCMC. Se implementa el algoritmo Metropolis-Hastings de cadena independiente.

- Propuesta independiente

La propuesta independiente es una normal trivariada

$$q(\theta^j|\theta^{j-1}) = q(\theta^j) \sim N_3(\theta, c\Sigma) \quad (15)$$

Aquí $\theta = (\alpha_0, \alpha_1, \alpha_2)$ y θ^j es el vector de parámetros en la iteración j , c es una constante para calibrar la tasa de rechazo y Σ es una matriz de covarianzas muestral que se estima por medio de una corrida exploratoria de la cadena de Markov. Rodríguez-Caballero

(2005) mantuvo tasas de rechazo alrededor del 45% para mantener la transición de estados, esta se logra calibrando conjuntamente c y Σ .

- Probabilidad de salto

La probabilidad de salto, mostrada en (10), utilizando una propuesta independiente se define como

$$\alpha(\theta^{j-1}|\theta^j) = \min \left[\frac{\pi(\theta^{(j)}) q(\theta^{(j-1)})}{\pi(\theta^{(j-1)}) q(\theta^{(j)})}, 1 \right] \quad (16)$$

Usando (14) y (15) en (16), se obtiene

$$\alpha(\theta^{(j-1)}, \theta^{(j)}) = \min \left[\frac{\left[\prod_{t=1}^n \left(\frac{1}{h_t} \right)^{1/2} \right] \exp \left\{ -\frac{1}{2} \sum_{t=1}^n \frac{y_t^2}{h_t} \right\} c \sum^{-1} (2\pi)^{-3/2} \exp \left\{ -\frac{1}{2} (\theta^{(j-1)} - \theta_0)' c \sum^{-1} (\theta^{(j-1)} - \theta_0) \right\}}{\left[\prod_{t=1}^n \left(\frac{1}{h_t} \right)^{1/2} \right] \exp \left\{ -\frac{1}{2} \sum_{t=1}^n \frac{y_t^2}{h_t} \right\} c \sum^{-1} (2\pi)^{-3/2} \exp \left\{ -\frac{1}{2} (\theta^{(j)} - \theta_0)' c \sum^{-1} (\theta^{(j)} - \theta_0) \right\}}, 1 \right] \quad (17)$$

Para facilidad computacional se simula bajo la transformación $\ln(\alpha(\theta^{j-1}|\theta^j))$ desde (17) y se le conoce como logratio.

Hasta aquí se ha concluido con la definición de elementos necesarios para simular la distribución posterior del modelo ARCH (2).

Para ver las especificaciones de cálculo del logratio y los detalles de la simulación computacional, ver Rodríguez-Caballero (2005)

B) Inferencia bayesiana en la volatilidad del IPC

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.

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 θ 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 definida en (15).

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 mil sin embargo este número de iteraciones puede reducirse utilizando métodos de aceleración de convergencia para MCMC, ver Gamerman (2006). Sin embargo Bauwens (1998) sugiere iterar 10 mil veces el Metropolis-Hastings para encontrar convergencia en la cadena, por si fuera poco Florentini (2002) y Zhang(2003) concuerdan que una iteración de tamaño 500 mil y un burn-in-period de 50 mil 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 en Rodríguez-Caballero (2005) se analiza a detalle se hicieron pruebas desde 10 mil hasta 300 mil iteraciones encontrando los mejores resultados en un burn-in-period de 75 mil e iteraciones hasta 100 mil.

Como ya se había comentado con anterioridad se utiliza el subyacente IPC 1990-2004 para ejemplificar el uso. Se obtienen los siguientes resultados.

La estimación máxima verosímil de los parámetros del modelo ARCH(2) se encuentran en la figura(3). La figura (2) ilustra al IPC con su serie de retornos y el modelo ARCH(2) estimado por máxima verosimilitud para la misma serie.

Se calcula a partir de la matriz M_T con las 100 mil iteraciones de la cadena de Markov y quitando las primeras 75 mil correspondientes al burn-in-period, se calcula la media posterior de la volatilidad. La figura (3) muestra el comparativo.

En la figura (4) se muestran los kernels suavizados de estimación para los parámetros del modelo para volatilidad utilizado.

La figura (5) muestra un comparativo entre los métodos utilizados para estimar la volatilidad. La importancia de los intervalos de credibilidad radica en que brindan una

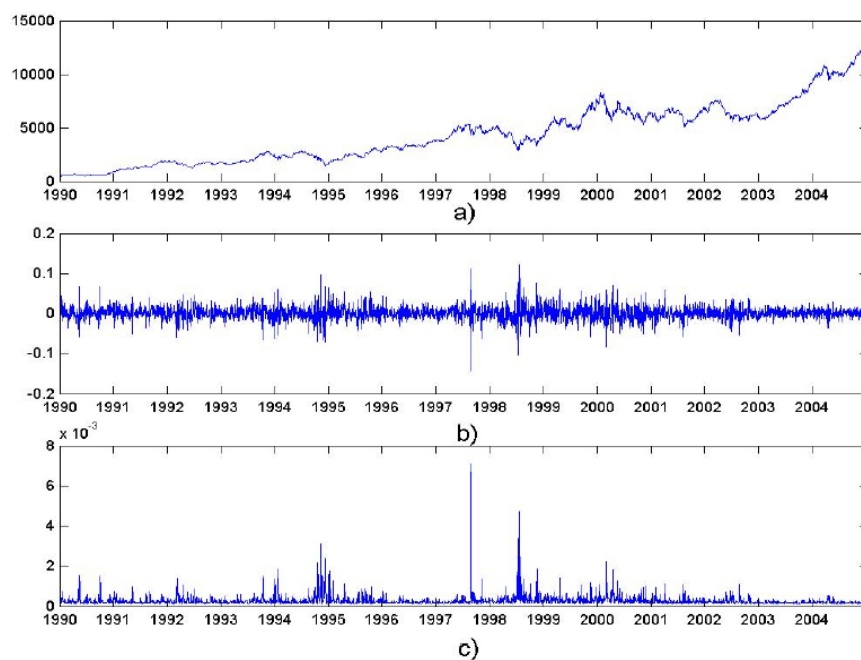


Figura 2. a) Serie IPC 1990-2004 b) Serie de retornos
c) Volatilidad estimada por máxima verosimilitud

Parámetro	EMV	Media posterior M-H	DE EMV	DE posterior M-H
α_0	0.0001551	0.000156	0.39324×10^{-6}	$.5147 \times 10^{-6}$
α_1	0.29368	0.2970	0.019534	.0248
α_2	0.16431	0.1662	0.015759	.0198

Figura 3. Estimación Máxima Verosimilitud (EMV) para la media y desviación estándar (DE) vs estimación por medio del Metropolis-Hastings (MH)

medida más informativa de la incertidumbre en comparación de cuando se usan estimadores puntuales como los máximo verosímiles.

La figura (6) muestra el kernel suavizado de la distribución posterior de la volatilidad conjuntamente con el monitor informal de convergencia en el cual se verifica la estacionariedad de la cadena de Markov.

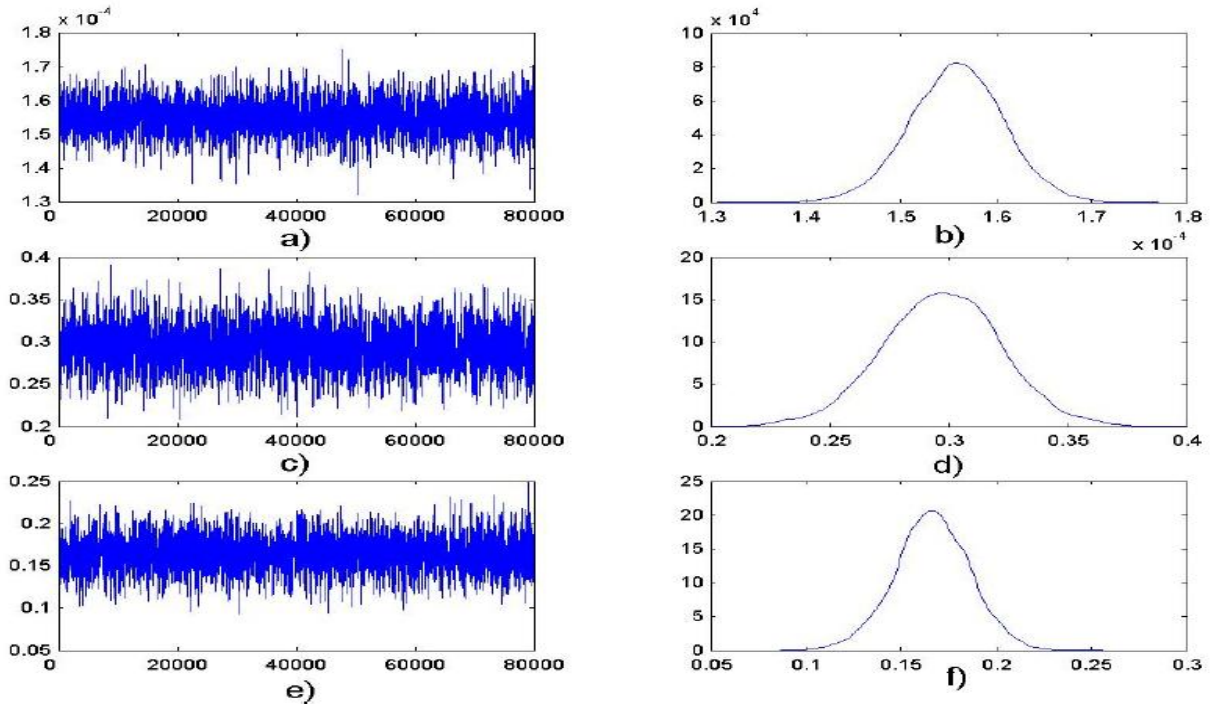


Figura 4. a) Iteraciones α_0 b) Distribución posterior α_0 c) Iteraciones α_1
d) Distribución posterior α_1 e) Iteraciones α_2 f) Distribución posterior α_2

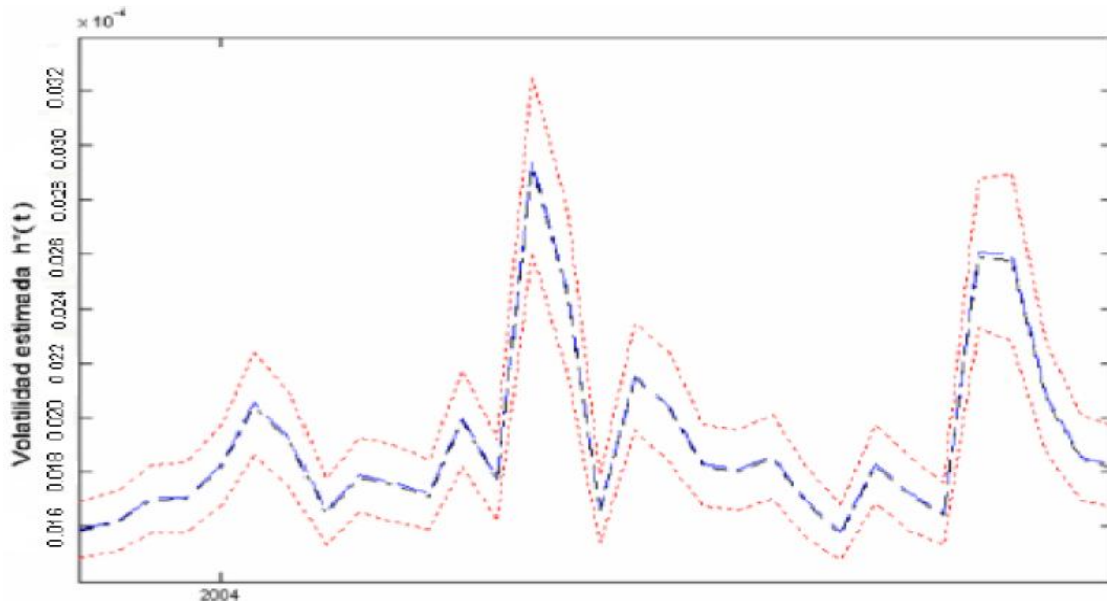


Figura 5. Estimación máxima verosímil (línea punteada negra), media posterior (línea punteada azul) e intervalos de credibilidad al 95% (líneas punteadas rojas)

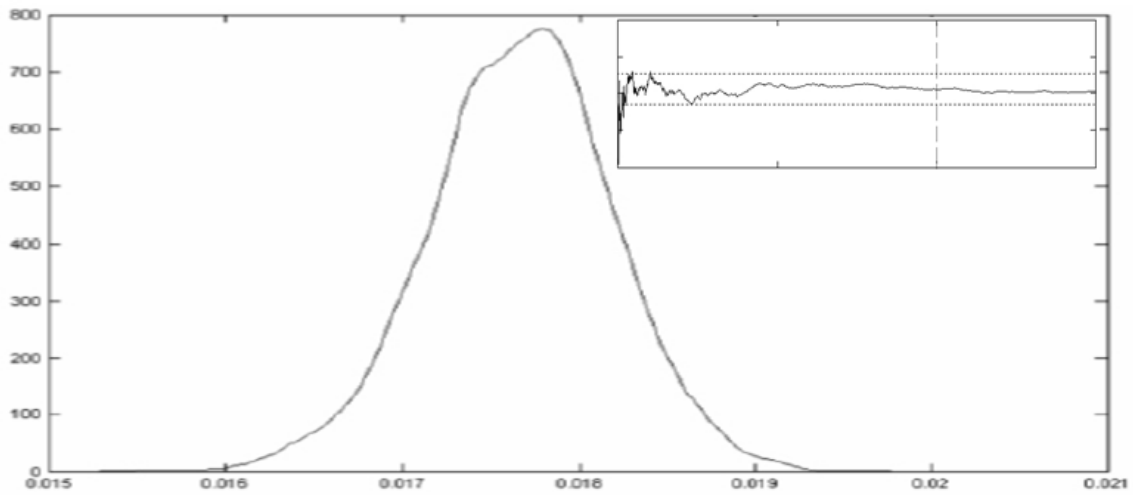


Figura 6. Distribución posterior de la volatilidad al 30 de diciembre de 2004. El gráfico del extremo superior derecho se denomina monitor informal de convergencia, la línea vertical punteada determina el límite del periodo de calentamiento.

C) Implementaciones financieras

1) Implementación en el modelo Black & Scholes

A continuación se muestra una propuesta diferente de inferencia bayesiana en el modelo Black & Scholes. Se trata de reunir aspectos importantes de investigaciones como Johannes (2003), quién 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 errónea y es por esto que es indispensable definirla mediante un modelo de series de tiempo para volatilidad, a manera de ejemplo se implementa las muestras posteriores del ARCH 2 conseguidas en la sección anterior.

El modelo Black & Scholes ya mundialmente conocido está determinado por

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

Los datos utilizados para el cálculo del modelo Black & Scholes se encuentran la figura (7).

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

Figura 7. Datos reales para el cálculo del precio de compra y precio de venta de una opción para el IPC

Los resultados obtenidos se observan en las figuras (8) y (9)

Para análisis de los resultados, análisis del break even y las pruebas out-of-sample generadas, ver Rodríguez-Caballero (2005).

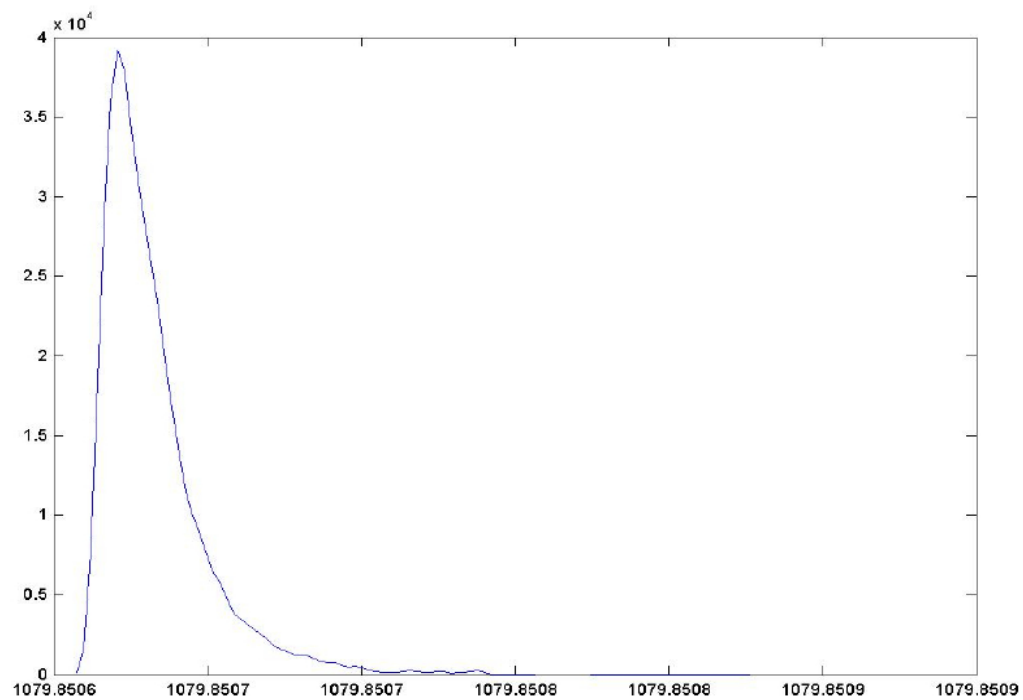


Figura 8. Distribución posterior para una opción de compra para el IPC contratada al 31 de diciembre de 2004 vía el modelo Black & Scholes.

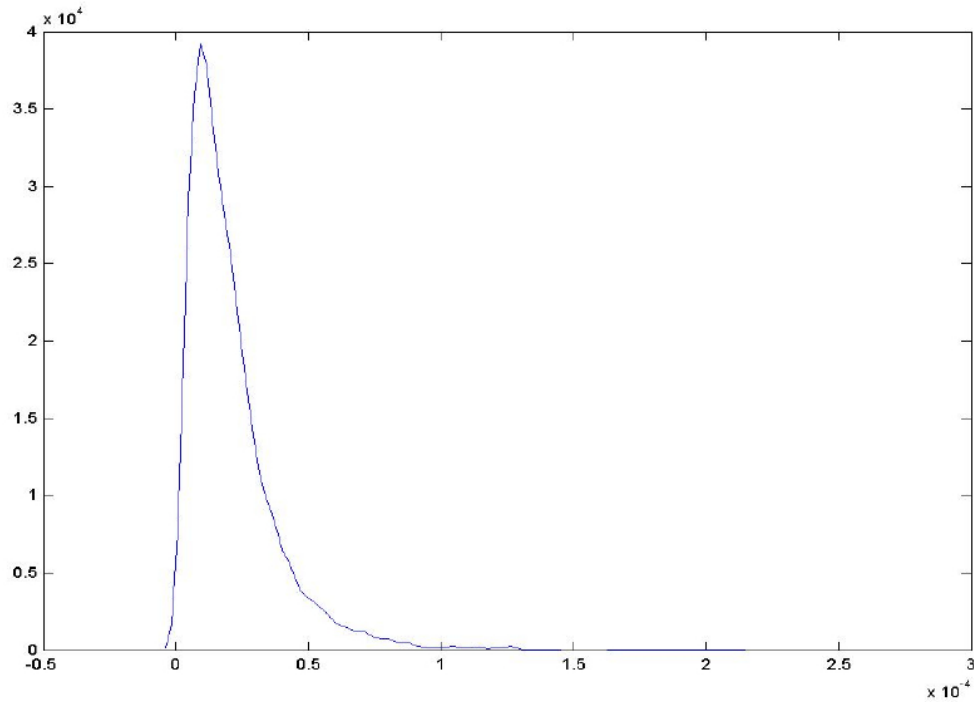


Figura 9. Distribución posterior para una opción de venta para el IPC contratada al 31 de diciembre de 2004 vía el modelo Black & Scholes.

2) Implementación en el modelo de valor en riesgo VaR

El valor en riesgo es una sencilla medida estadística de las posibles pérdidas del portafolio de crédito. Específicamente, en VaR es una medida de pérdida debido a las fluctuaciones “normales” del mercado. Se supone que pérdidas más allá del límite VaR son sufridas con probabilidades muy pequeñas. El VaR, sujeto a sus múltiples suposiciones, agrega todos los riesgos involucrados en la administración integral del riesgo en un sencillo resultado el cual es reportado tanto a los reguladores crediticios como al consejo de administración de las instituciones financieras.

La especificación del VaR se observa en la figura 10.

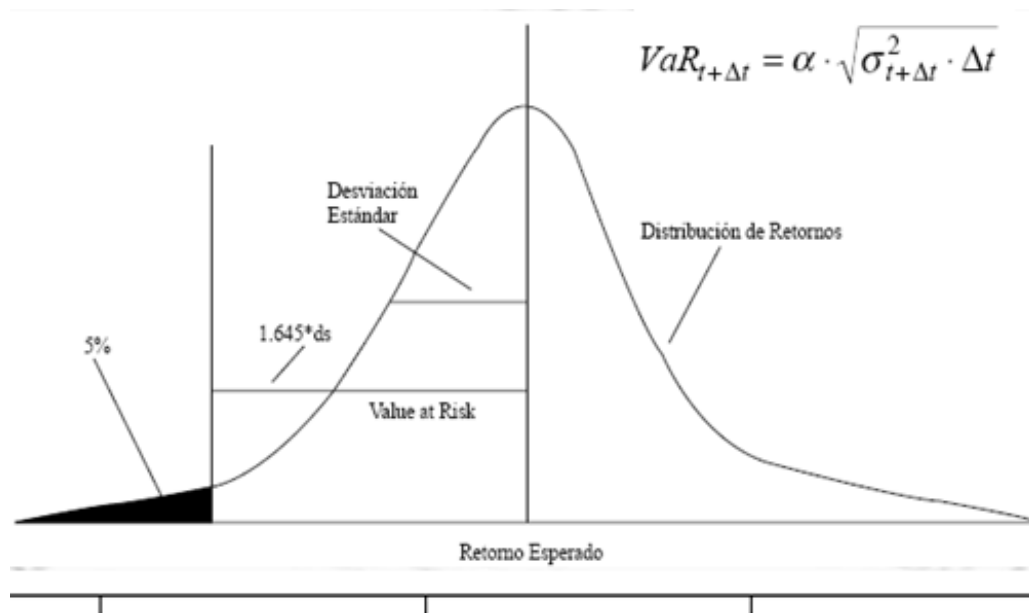


Figura 10. Valor en Riesgo de un portafolio crediticio hipotético al 5%.

El modelo VaR al igual que el modelo Black & Scholes supone una volatilidad constante σ , en la propuesta del presente trabajo se implementa las muestras de la distribución posterior de la volatilidad encontradas en el modelo ARCH(2) considerando, a vil manera de ejemplo, que el portafolio IPC pudiera ser considerado como el portafolio de crédito (de inversión) de una institución financiera.

Para una muestra pequeña de 250 observaciones, se tiene un ejemplo de cómo quedarían definidos los intervalos de credibilidad para el histórico del VaR. La figura (11) lo muestra.

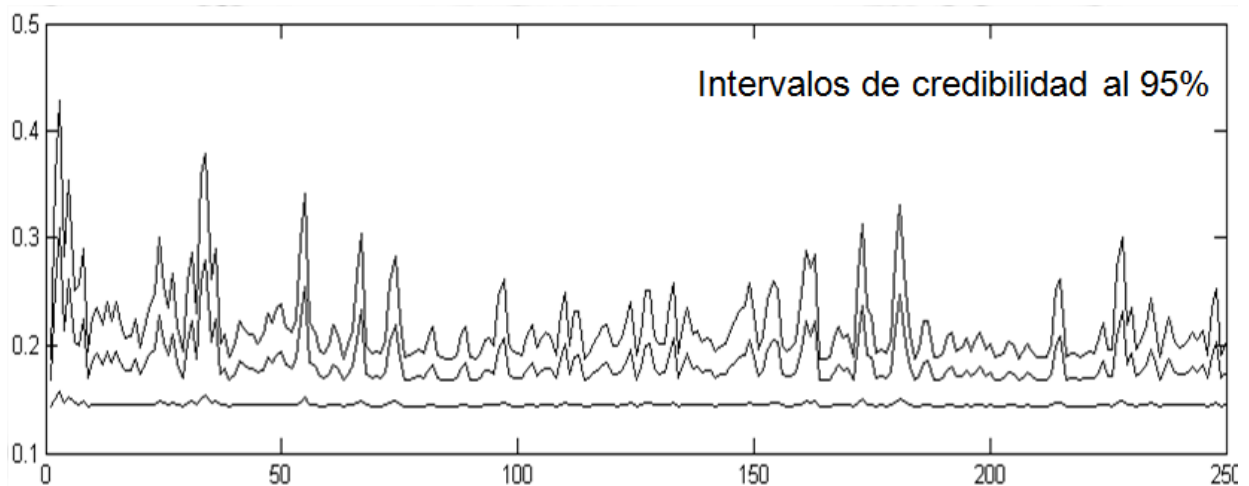


Figura 11. Intervalos de credibilidad al 95%. La línea central es el VaR calculado de manera convencional.

Para el ejercicio actual, la figura (12) muestra la distribución posterior del VaR.

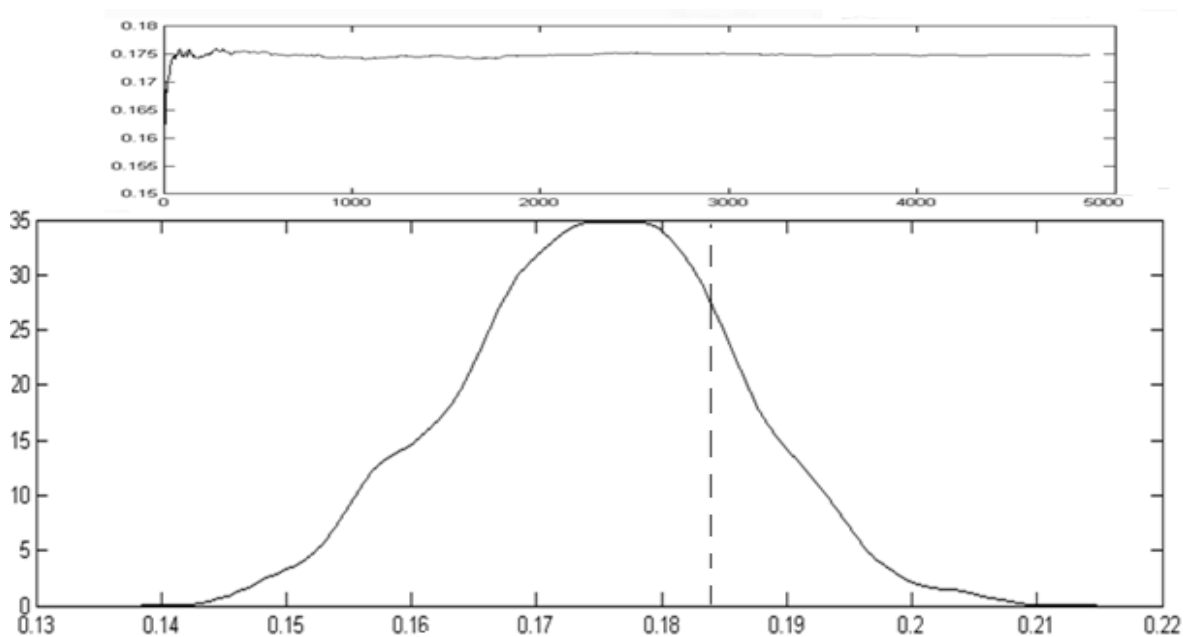


Figura 12. Monitor informal de convergencia (gráfico superior) y distribución posterior del VaR

(gráfico inferior). La línea vertical punteada representa la estimación del VaR con método convencional.

IV. Scoring de crédito bayesiano

Una parte también central en la administración de riesgos es la correspondiente al proceso de generación de crédito, la información obtenida en la solicitud de crédito permite

mediante árboles de decisión y a través de un modelo de regresión logística, calibrar un scoring o scorecard de originación de crédito, el objetivo de este modelo es asignar un puntaje a cada cliente de acuerdo a sus características sociodemográficas y crediticias tal que este determine la calidad de los nuevos solicitantes.

Es objetivo del desarrollo poder identificar grupos o segmentos de mercado dentro de la población de clientes, a fin de desarrollar modelos que permitan realizar un mejor ordenamiento de los solicitantes y, en particular, para generar herramientas que permitan calificar adecuadamente a un segmento de la población a la que se le desee enfocar la promoción de un cierto producto de crédito. Los modelos de scorecard son ya muy usados en cada institución financiera y las metodologías con las que se desarrollan son relativamente ya muy específicas entre las instituciones, de manera general en las instituciones financieras del país suelen usarse árboles de decisión estadísticas tipo CHAID o regresiones tanto lineales como logísticas y combinación de estas.

Existe mucha bibliografía ya disponible al respecto, en Ramírez (2008) se presenta un resumen bastante bien elaborado del modelaje de scoring de crédito. En Rodríguez-Caballero y Espín-García (2011), dos versiones, se muestra a detalle todo el procedimiento de modelaje y sus análisis posteriores conjuntamente con propuestas para la administración de riesgos en un banco en particular.

La clasificación común en la población está determinada en el número de incumplimientos en un cierto periodo de tiempo, de manera estándar se define una población buena como aquellos clientes que tienen un máximo de 1 pago vencido en una ventana de tiempo comúnmente de entre 12 y 18 meses, malos clientes serán aquellos que hubieran estado al menos en una ocasión en 3 incumplimientos (pagos vencidos) en la misma ventana de tiempo. Comúnmente clientes indeterminados se les denomina a los que

han caído como máximo en 2 pagos vencidos, a veces pudieran ser considerados como clientes buenos o malos de acuerdo a las metodologías de crédito de cada banco.

Sin gran detalle y a manera de resumen, los modelos estándares de scoring cuentan con las siguientes desventajas:

- Las solicitudes no representan una muestra aleatoria de la población y éstas inevitablemente serán sesgadas.
- Existencia de valores missing no aleatorios.
- Cambios en condiciones poblacionales crean modelos poco predictivos.
- No hay forma de estar seguro de que la información sea fidedigna.
- Los malos pagadores terminan siendo a menudo los clientes más rentables y los modelos disponibles no incorporan esta información lo cual acarrea dejar de estar percibiendo fuertes ingresos para la institución financiera correspondiente.

Es por ello que se ha buscado modernizar estos tipos de modelos venciendo las limitantes que presentan los scorecard estándares y poder encontrar estimadores más robustos acerca de las características crediticias involucradas en el comportamiento de las poblaciones.

A) Scoring de crédito vía mezclas finitas

Karlis-Rahmouni (2007) adoptaron un enfoque distinto para modelar el scoring de crédito de una institución financiera, en vez de tratar de predecir la clase poblacional (bueno/malo) de los solicitantes, predijeron el número de defaults (incumplimientos) en el futuro cercano. Fijaron mezclas finitas de poissones permitiendo la heterogeneidad entre los grupos para tratar el número de componentes desconocidas las cuales fueron estimadas a través de los propios datos del solicitante. Usando covariables en todos los componentes revelaron el impacto de varias variables sociodemográficas en crear diferentes grupos de clientes y no

solo eso sino que fueron capaces de predecir para cada cliente el grupo al cual deberían pertenecer así como su número esperado de defaults en cierto tiempo.

Lo que muestra el artículo seminal de Karlis-Rahmouni es en verdad formidable ya que abre el camino de conocer la pérdida esperada de cada cliente, algo por demás satisfactorio ya que permitiría estimar la tasa de interés óptima por cliente y así por fin poder maximizar la rentabilidad individual. Dilema existencial del sector bancario y aún no resuelto.

El enfoque de un scoring bayesiano basado en las propuestas originales de Karlis-Rahmouni(2007) debería de contemplar los siguientes puntos:

- Modelar vía mezclas finitas de poissones
- Desconocer el número de componentes (número de grupos de población)
- Emplear el paradigma bayesiano para modelar los dos puntos anteriores.

3) *Mezclas finitas de poissones*

Supóngase que la densidad de probabilidad de una variable aleatoria x , $f(x)$ puede ser expresada en la forma

$$f(x) = \sum_{j=1}^k p_j f(x|\theta_j)$$

Para algunas probabilidades $p_j > 0$, $j = 1, \dots, k$ con $\sum_{j=1}^k p_j = 1$, y θ_j es o un vector de parámetros o un escalar referente al j -ésimo componente de la mezcla.

Ahora bien la probabilidad de la k -ésima mezcla poisson está dada por

$$f(x) = \sum_{j=1}^k p_j \frac{e^{-\lambda_j} \lambda_j^x}{x!} \quad (18)$$

Asumimos que $0 \leq \lambda_1 < \lambda_2 < \dots < \lambda_k$ para asegurar la identificabilidad de (18)

Supongamos que los clientes forman varias subpoblaciones k obviamente desconocidas. Entonces se quiere encontrar el número de grupos (k) así como sus características. Aquí es

donde radica la principal diferencia con los métodos convencionales ya que aquí es posible movernos entre varias subpoblaciones, por ejemplo rentables y no rentables, rentables tipo 1, tipo 2 ,..., tipo k, y no solo entre poblaciones buenas y malas que desde su definición es realmente dudosa.

Entonces el problema “se reduce” a estimar tanto k como los parámetros de las subpoblaciones. No es que sea fácil su obtención sino que al menos ya se tiene identificado el problema a resolver.

4) Simulación RJMCMC

Los métodos comunes de estimación de la verosimilitud tipo algoritmos EM son básicamente imposibles de usarlos al desconocer k. Debido a la forma que presenta la distribución en (18) no es posible encontrar una forma conjugada y por tanto si se quiere estimar con mecanismos bayesianos $\lambda_1, \lambda_2, \dots, \lambda_k, p_1, p_2, \dots, p_{k-1}$ deberán usarse una vez más algoritmos MCMC.

Diebolt-Robert (1994) desarrollaron un algoritmo capaz de resolver este problema usando un algoritmo MCMC sin embargo usando el hecho de que k fuera conocido.

En el problema planteado k es desconocido por lo que no es posible usar su algoritmo planteado.

En la segunda sección del presente artículo se popularizó el método RJMCMC en donde se comentaba que permitía a uno moverse a través de modelos aún si fueran de distintas dimensiones con alguna probabilidad de aceptación de quedarse o moverse entre modelos mediante la corrida exploratoria de un estado en la iteración MCMC.

Los pasos del algoritmo RJMCMC para el problema planteado son:

- 1) Decidir si intentamos ir al componente $k - 1$ o $k + 1$ (componentes split o combine)

- 2) Hacer las transformaciones requeridas para emparejar la diferencia de dimensiones entre los dos modelos.
- 3) Aceptar o rechazar el movimiento con alguna probabilidad
- 4) Hacer una simple iteración MCMC para el modelo actual
- 5) Regresar al paso 1

Para el caso particular de la mezcla finita de poissones usó la siguiente transformación para el caso de combinar componentes, a este paso se le denomina *combine step*.

Sea el par de componentes a analizar (j_1, j_2) , se tiene que transformar el vector actual de parámetros $(p_{j_1}, p_{j_2}, \lambda_{j_1}, \lambda_{j_2})$ al nuevo vector de parámetros (p_*, λ_*) así como la asignación de variables latentes (z_{j_1}, z_{j_2}) a la nueva asignación de la variable latente z_* .

Puede usarse la siguiente transformación

$$p_* = p_{j_1} + p_{j_2}, \quad \lambda_* = \frac{\lambda_{j_1} p_{j_1} + \lambda_{j_2} p_{j_2}}{p_*}, \quad z_* = z_{j_1} + z_{j_2}$$

Ahora bien para que la cadena se mueva ahora a un modelo con más parámetros se deben de generar dos variables aleatorias desde una distribución Beta (2,2). Entonces los nuevos parámetros son calculados como

$$p_{j_1} = p_{j_*} u_1, \quad p_{j_2} = p_{j_*} (1 - u_1)$$

$$\lambda_{j_1} = \lambda_{j_*} u_2, \quad \lambda_{j_2} = \lambda_{j_*} \left(\frac{1 - u_1 u_2}{1 - u_1} \right)$$

Finalmente, si $z_{ij_*} = 1$ para algún i en particular, entonces $z_{ij_1} = 1$ con probabilidad proporcional a $w_{ij_1} / (w_{ij_1} + w_{ij_2})$. Y la matriz jacobiana de la transformación está dada por $|J| = \frac{p_{j_*} \lambda_{j_*}}{1 - u_1}$. A este paso se le denomina *split step*.

Las probabilidades de aceptación para el split step están dadas por $\min(1, A)$, mientras que para el combine step será $\min(1, B)$, donde

$$A = \frac{L(k+1)}{L(k)} \frac{\pi(k+1)}{\pi(k)} \frac{\prod_{i=1}^{k+1} \pi(\lambda_i) \pi(p_1, p_2, \dots, p_{k+1})}{\prod_{i=1}^k \pi(\lambda_i) \pi(p_1, p_2, \dots, p_k)} (k+1) [g(u_1, u_2)]^{-1} |J| \frac{C_{k+1}}{S_k P_{asig}},$$

$$B = \frac{L(k-1)}{L(k)} \frac{\pi(k-1)}{\pi(k)} \frac{\prod_{i=1}^{k-1} \pi(\lambda_i) \pi(p_1, p_2, \dots, p_{k-1})}{\prod_{i=1}^k \pi(\lambda_i) \pi(p_1, p_2, \dots, p_k)} \frac{1}{k} [g(u_1, u_2)] \frac{1}{|J|} \frac{S_{k-1} P_{asig}}{C_k}.$$

Donde, $L(k)$ es la verosimilitud del modelo con k componentes calculados como el producto of $f(x_i|\lambda_j, z_{ij} = 1)$ y $\pi(k)$, $\pi(\lambda)$ y $\pi(p_1, p_2, \dots, p_k)$ son las aprioris para k, λ y p , respectivamente. $g(u_1, u_2)$ es la densidad propuesta desde la cual u_1 y u_2 son generadas. Los factores $(k+1)$ para el split y $(k-1)$ para el combine son derivados desde los ratios de los factoriales en las densidades para garantizar en significado de mezclas.

5) Ejemplo de aplicación

Karlis (1994) analizó los defaults de una institución financiera española en 1990, seguido de ellos implementó un modelo basado en mezclas finitas de poissones para identificar los grupos que hemos venido hablando.

El máximo número de incumplimientos fueron 25, y por tanto se toma este hecho para definir que el máximo número de componentes para el modelo será 25.

Se toma una a priori uniforme a través de los valores de k desde 1 hasta 25, para las λ_j se elige una apriori bastante plana $\Gamma(0.01, 0.01)$, para facilidad de manejo. Finalmente las aprioris para los parámetros proporcionales de las mezclas p_j se toman de una Dirichlet(1, ..., 1).

Como es de esperarse el algoritmo termina siendo insensible a los valores iniciales de la distribución posterior conjunta, más o menos en mil iteraciones. El burn in period se

establece en 5 mil iteraciones y para evitar que exista la posibilidad de que haya autocorrelaciones sesgadas entre la población se toma un punto cada 100 iteraciones.

En el artículo de Karlis (1994) se muestran los monitores informales de convergencia para ciertos grupos en particular. A su vez, se muestran las probabilidades de movimiento para cada k y la distribución posterior para k .³

Para finalizar es importante comentar que debe realizarse un punto de corte para determinar hasta que componente k consideraremos como población rentable o buena. Entonces a partir del cut-off mencionado y usando las distribuciones posteriores mostradas será posible definir las políticas de crédito adecuadas al riesgo y rendimiento de cada cliente. Estas podrían ser fácilmente cambiadas día a día.

V. Conclusiones

Se discute el cimiento y el funcionamiento de la inferencia bayesiana y los métodos MCMC así como el aprovechamiento de estas herramientas estadísticas en algunos ejemplos específicos concernientes a problemas en la administración de riesgos. Se muestra cómo superar la volatilidad constante en modelos como el Black & Scholes o el VaR, hallando muestras de la distribución posterior del modelo de volatilidad propuesto.

El enfoque bayesiano en mezclas finitas de poissones pueden ser aplicados a más problemas por el estilo, no solo en el scorecard mencionado. El enfoque mencionado brinda información natural acerca de la ponderación de la información existente con la información a priori la cual típicamente existe para desarrollar un scoring de crédito.

³ Este esbozo preliminar del artículo no muestra los resultados empiricos obtenidos al seguir la metodología explicada. No obstante en el artículo publicado se encuentran. Aquel lector interesado en los códigos de la simulación podrá pedirlos al autor.

En la práctica del scoring de crédito usado por los bancos son mucho más empíricos y mucho menos automáticos. En el presente trabajo no se toman en cuenta las restricciones financieras existentes en el sector financiero, pero pudieran ser implementadas en modelos más exigentes.

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 usados constantemente en la práctica, implementando métodos MCMC o RJMCMC que si bien no han sido tan explotados en la investigación de riesgos 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 y cálculo de VaR. Si bien se propone el aprovechamiento de un modelo ARCH para explicar la volatilidad de los subyacentes financieros o volatilidades en portafolios de crédito, 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 ϵ_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, sobre todo los RJMCMC 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.

V. Bibliografía

Bauwens, L., (1998). "Bayesian inference on GARCH models using the Gibbs sampler". *Econometrics Journal*, Vol.1. P C23-C46. 1998

Berger, J. (1998). "Decision Theory and Bayesian Analysis". *Springer Series in Statistic*, Nueva York 1998.

Box & Jenkins (1970). "Bayesian inference in statistical analysis". *Wiley Classics Library*. EUA 1970.

Casella, G. (1992). "Explaining the Gibbs Sampler". *The American Statistician*, Vol.46. P.167. 1992

Casella, G. (1999). "Monte Carlo Statistical Methods". *Springer Series in Statistic*, Nueva York. 1999

Chib (1995). "Understanding the Metropolis-Hastings Algorithm". *The American Statistician*, Vol.49. P. 327-335. 1995

Chib, S., (2001). "Markov chain Monte Carlo methods: Computation and inference". *Handbook of Econometrics*, Volume 5, pp. 3569-3649. Amsterdam: North-Holland. 2001

Diebolt J. and Robert C. (1994). "Estimation of finite mixture distributions through bayesian sampling". *Journal of the Royal Statistical Society*, B 56, 363–375. 1994

Florentini, G., (2002). "Likelihood-based estimation of latent generalized ARCH structures". Unpublished paper. Universidad de Alicante, Italia. 2002

Geweke, J. (1989). "Bayesian Inference in Econometric Models Using Monte Carlo Integration". *Econometrica*, Vol.57. P. 1317-1339. 1989

Green, P. J. (1995). "Reversible jump Markov chain Monte Carlo computation and Bayesian model determination". *Biometrika*, 82, 711–732.

- Green, P. J. (2003). "Trans-dimensional Markov chain Monte Carlo". *Oxford Statistical Science Series*, No. 27, pp. 179–198. *Oxford University Press*, Oxford. 2003
- Hastie, D. (2005). "Towards Automatic Reversible Jump Markov Chain Monte Carlo" *PhD dissertation in the Faculty of Science. University of Bristol*. Marzo 2005
- Johannes, M. (2003). "MCMC Methods for Continuous-Time Financial Econometrics". *Handbook of Financial Econometrics*, 2003.
- Karlis, D. (1994). "Credit Scoring based on Bayesian analysis of finite Poisson mixtures". *Department of statistics. Athens University of Economics. Bergen*. Mayo 2004. Technical report.
- Karlis, D. Rahmouni, M. (2007). "Analysis of defaulters' behavior using the Poisson-mixture approach". *Journal of Management Mathematics*, 18, 297–311. 2007
- Lee, B. (1989). "Bayesian Statistic. An Introduction". *Oxford University Press*, Reino Unido, 1989.
- Pedersen, J., (2004). "Markovkæde Monte Carlo (MCMC) med Metropolis-Hastings algoritmen". *Stockholm School of Economics, Technical Report*. 2004
- Ramírez, A. (2008). "Técnicas de minería de datos aplicadas a la construcción de modelos de score crediticio: estado del arte" *Universidad Nacional de Colombia. Artículo sin publicar*. Bogotá. 2008
- Rodríguez-Caballero, C.V. (2005). "Inferencia bayesiana para la volatilidad en el modelo Black & Scholes". *Tesis de licenciatura en actuaria. Facultad de Ciencias. UNAM*. 2005
- Rodríguez-Caballero, C.V y Espín-García, O. (2011). "Metodología para desarrollar un scoring para clientes sin referencias crediticias. Documento de trabajo. [En arbitraje]
- Rodríguez-Caballero, C.V y Espín-García, O. (2011). "A generic scorecard methodology for credit cards". Documento de trabajo. [En arbitraje]
- Ross, S. (2000). "Probability Models". *Academic Press*, San Diego, Cal. 2000.
- Gamerman, D., (2006). "Markov Chain Monte Carlo. Stochastic simulation for Bayesian inference". *Chapman & Hall*, Londres. 2006
- Tierney, L. (1998). "A note on Metropolis-Hastings kernels for general state spaces". *Annals of Applied Probability*, 8, 1–9. 1998
- Tsay, R., (2002). "Analysis of Financial Time Series". *Financial Econometrics. John Wiley & Sons*, E.U.A. 2002
- 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. 2003