# Estimación de parámetros en ecuaciones diferenciales estocásticas aplicadas a finanzas

John Freddy Moreno Trujillo*

*jhon.moreno@uexternado.edu.co*

* Docente investigador. Facultad de Finanzas, Gobierno y Relaciones Internacionales. Universidad Externado de Colombia. Matemático de la Universidad Nacional de Colombia y Magister en matemática aplicada de la misma universidad. Candidato a PHD en economía de la Universidad del Rosario.

# 1. Introducción

Las ecuaciones diferenciales estocásticas (EDE) son un modelo natural para describir sistemas con evolución dinámica, que están sujetos a algún grado de aleatoriedad. Como modelo general consideramos la siguiente EDE paramétrica de Itô:

$$dX_t = \beta(\theta, t, X_t)dt + \sigma(\gamma, t, X_t)dW_t, \quad t \geq 0, \quad X_0 = \xi \tag{1}$$

donde $\{W_t, t \geq 0\}$ es un movimiento browniano estándar, $\beta : \Theta \times [0, T] \times \mathbb{R}$ es llamado coeficiente de tendencia, y $\sigma : \Gamma \times [0, T] \times \mathbb{R}^+$ que es el coeficiente de difusión. Por lo general, estos dos coeficientes son funciones conocidas de los parámetros (o vectores de parámetros) desconocidos $\theta$ y $\gamma$, lo cual hace indispensable desarrollar metodologías para su estimación.

Los siguientes son algunos ejemplos de EDE utilizadas para describir el comportamiento de variables de carácter financiero. En cada caso es posible apreciar explícitamente la forma funcional de cada coeficiente y su dependencia de los parámetros.

- Bachelier:

$$dX_t = \beta dt + \sigma dW_t$$

- Black-Scholes:

$$dX_t = \beta X_t dt + \sigma X_t dW_t$$

- Ornstein-Uhlenbeck:

$$dY_t = \beta X_t dt + \sigma dW_t$$

- Ornstein-Uhlenbeck Radial:

$$dX_t = \left(\alpha X_t^{-1} - X_t\right)dt + \sigma dW_t$$

- Cox-Ingersoll-Ross:

$$dX_t = \left(\alpha + \beta X_t\right)dt + \sigma \sqrt{X_t} dW_t$$

- Difusión hiperbólica:

$$dX_t = \alpha \frac{X_t}{\sqrt{1+X_t^2}} dt + \sigma dW_t$$

- Vasicek:

$$dX_t = \left(\alpha - \beta X_t\right) dt + \sigma dW_t$$

- Dothan:

$$dX_t = \left(\alpha + \beta X_t\right) dt + \sigma X_t dW_t$$

- Black-Derman-Toy:

$$dX_t = \beta\left(t\right) X_t dt + \sigma\left(t\right) X_t dW_t$$

- Black-Karasinksi:

$$d \ln\left(X_t\right) = \ \alpha\left(t\right) - \beta\left(t\right) \ln\left(X_t\right) \ dt + \sigma\left(t\right) dW_t$$

- Ho-Lee:

$$dX_t = \alpha\left(t\right) dt + \sigma dW_t^H$$

- Hull-White (extensión Vasicek):

$$dX_t = \left(\alpha\left(t\right) + \beta\left(t\right) X_t\right) dt + \sigma\left(t\right) dW_t$$

- Hull-White (extensión CIR):

$$dX_t = \left(\alpha\left(t\right) + \beta\left(t\right) X_t\right) dt + \sigma\left(t\right) \sqrt{X_t} dW_t$$

- Cox-Ingersoll-Ross (1.5):

$$dX_t = \sigma X_t^{3/2} dW_t$$

- Ahn-Gao:

$$dX_t = \beta\left(\mu \ - X_t\right) X_t dt + \sigma X_t^{3/2} dW_t$$

- Ait-Sahalia:

$$dX_t = \left(\alpha + \beta X_t + \gamma X_t^{-1} + \delta X_t^2\right)dt + \sigma X_t^{\gamma} dW_t$$

Uno de los procedimientos más utilizados para la estimación de parámetros en EDE es el de máxima verosimilitud, el cual, solo en casos muy particulares, puede ser aplicado de forma directa a partir de conocer la función de densidad de probabilidad del proceso subyacente en la ecuación, mientras que en la mayoría de los casos se hace necesario el uso de métodos numéricos de optimización.

En la siguiente sección se muestran dos ejemplos de la aplicación directa del método de máxima verosimilitud para la estimación de los parámetros en los modelos de Black-Scholes y Vasicek.

## 2. Estimación de parámetros por máxima verosimilitud

El objetivo de la estimación máximo verosímil es establecer los valores de los parámetros que hacen que la función de verosimilitud del proceso se maximice.

Para lo anterior se parte de considerar un conjunto de observaciones del proceso $(X_0, X_1, .., X_n)$, en los instantes de $t = 0, 1, ..., n$, con base en las cuales la función de densidad del proceso puede verse como una función del parámetro o parámetros desconocidos $f_\theta$ y es denominada función de verosímilitud. Entonces, se plantea y resuelve el problema de determinar el valor del parámetro que hace máximo al logaritmo de la función de verosimilitud[1].

La posibilidad de poder determinar la función de verosimilitud condiciona la solución directa del problema de maximización o la aplicación de métodos numéricos. Que el proceso $X_t$ sea de Markov es una condición suficiente para poder escribir la función de verosimilitud $f_\theta$ con base en una serie temporal de observaciones, ya que:

$$f_\theta = f_{X_0,X_1,\ldots,X_n;\theta} = f_{X_n|X_{n-1};\theta} \cdot f_{X_{n-1}|\ X_{n-2};\theta} \cdots f_{X_1|\ X_0;\theta} \cdot f_{X_0;\theta}$$

Si ademas de lo anterior se tiene que las observaciones son independientes e idénticamente distribuidas entonces:

---

[1] Se toma el logaritmo de la función de verosimilitud para simplificar la expresión, y los máximos se mantienen ya que el logaritmo es una función monótona.

$$f_{X_i|X_{i-1};\theta} = f_{X_i;\theta}$$

con lo cual la función de verosimilitud puede ser escrita como:

$$f_\theta = \prod_{i=0}^{n} f_{X_i;\theta}$$

y al tomar logaritmo natural se tiene la función de logverosimiltud:

$$L(\theta) = \ln\left(f_\theta\right) = \sum_{i=0}^{n} \ln\left(f_\theta\right) \quad (2)$$

sobre la cual se resuelve el problema de maximización igualando a cero las derivadas parciales respecto a los parámetros y solucionando las ecuaciones resultantes.

### 2.1. El modelo Black-Scholes

El modelo Black-Scholes está representado por EDE:

$$dX_t = \beta X_t dt + \sigma X_t dW_t \quad (3)$$

expresión que también se conoce como movimiento browniano geométrico, en la cual $\beta$ y $\sigma$ son parámetros constantes desconocidos.

En general este modelo es aplicado para describir el comportamiento temporal del precio de un activo riesgoso. La idea detrás de esta aproximación es básicamente considerar que el precio del activo sigue una recta de tendencia de pendiente $\beta$, con fluctuaciones aleatorias alrededor de dicha recta, ponderadas por el coeficiente $\sigma$.

Asumir que los precios siguen este comportamiento implica asumir que siguen una distribución lognormal, o lo que es equivalente, asumir que los retornos logaritmicos del activo siguen una distribución normal. Esto se justifica por la aplicación de la formula de Itô, y permite determinar, de forma exacta, la función de verosimilitud del proceso.

Al aplicar la formula de Itô[2] a los retornos del activo se tiene que:

$$d\ln\left(X_t\right)= \quad \beta-\frac{1}{2}\sigma^2 \quad dt+\sigma dW_t \tag{4}$$

de donde,

$$\ln \quad \frac{X_{t+\Delta t}}{X_t} \quad \sim N\left(m,v\right) \tag{5}$$

con

$$m= \quad \beta-\frac{1}{2}\sigma^2 \quad \Delta t \quad ; \quad v=\sigma^2\Delta t$$

Si se aplica el método de máxima versimilitud a los retornos logarítmicos, partiendo de una serie de observaciones del precio del activo $x_0, x_1, .., x_n$, y denotando a $Y_i = \ln \quad \frac{X_i}{X_{i-1}}$ de forma que $Y_i \sim N(m, v)$, entonces:

$$f_{Y;m,v}=\prod_{i=1}^{n}\frac{1}{\sqrt{2\pi v}}e^{-\frac{(y_i-m)^2}{2v}} = \quad \frac{1}{\sqrt{2\pi v}} \quad ^{n} e^{-\frac{1}{2v}\sum_{i=1}^{n}(y_i-m)^2} = \left(2\pi v\right)^{-n/2} e^{-\frac{1}{2v}\sum_{i=1}^{n}(y_i-m)^2} \tag{6}$$

---

2 Dado un proceso de Itô $X_t$ y una función $C^2$, la fórmula de Itô unidimensional establece que:

$$df\left(X_t\right)=f\ \left(X_t\right)dX_t+\frac{1}{2}f\ \ \left(X_t\right)\left(dX_t\right)^2$$

Para el caso del movimiento browniano geométrico y tomando $f(x) = \ln(x)$, se tiene que:

$$d\ln\left(X_t\right)=\frac{1}{X_t}\left(\beta X_t dt+\sigma X_t dW_t\right)-\frac{1}{2X_t^2}\left(\beta X_t dt+\sigma X_t dW_t\right)^2= \quad \beta-\frac{1}{2}\sigma^2 \quad dt+\sigma dW_t$$

y dado que el movimiento browniano es un proceso gaussiano se tiene que:

$$d\ln\left(X_t\right)\approx\ln\left(X_{t+\Delta t}\right)-\ln\left(X_t\right)=\ln \quad \frac{X_{t+\Delta t}}{X_t} \quad \sim N \quad \beta-\frac{1}{2}\sigma^2 \quad \left(\Delta t\right);\sigma^2\Delta t$$

y la función de logverosimilitud está determinada por:

$$L(m,v)=\ln\left(f_{Y;m,v}\right)=-\frac{n}{2}\ln(2\pi v)-\frac{1}{2v}\sum_{i=1}^{n}\left(y_i-m\right)^2 \tag{7}$$

Si se toma la derivada parcial de la función de logverosimilitud respecto a $m$ e igualándola a cero se tiene:

$$\begin{aligned}\frac{\partial}{\partial m}L(\theta)&=\frac{1}{2v}\sum_{i=1}^{n}2\left(y_i-m\right)\\&=\sum_{i=1}^{n}\left(y_i-m\right)=0\end{aligned}$$

y entonces:

$$\hat{m}=\frac{1}{n}\sum_{i=1}^{n}y_i=\overline{y} \tag{8}$$

y tomando la derivada parcial respecto a $v$ se tiene que:

$$\begin{aligned}\frac{\partial}{\partial v}L(\theta)&=-\frac{n}{2}\frac{1}{2\pi v}2\pi+\sum_{i=1}^{n}\left(y_i-m\right)^2\frac{1}{2v^2}\\&=-\frac{n}{2v}+\frac{1}{2v^2}\sum_{i=1}^{n}\left(y_i-m\right)^2=0\end{aligned}$$

y entonces:

$$\hat{v}=\frac{1}{n}\sum_{i=1}^{n}\left(y_i-\overline{y}\right)^2 \tag{9}$$

expresiones de las cuales se obtienen los siguientes estimadores de $\beta$ y $\sigma^2$.

$$\hat{\beta} = \frac{\overline{y}}{\Delta t} + \frac{1}{2}\hat{\sigma}^2 \tag{10}$$

$$\hat{\sigma}^2 = \frac{1}{n\Delta t}\sum_{i=1}^{n}\left(y_i - \overline{y}\right)^2 \tag{11}$$

Sobre el estimador para la media *m* del proceso de retornos logarítmicos *Y*, determinado por la expresión (8) es importante notar que:

$$\begin{aligned}\hat{m} = \overline{y} = \sum_{i=1}^{n}\frac{Y_i}{n} &= \frac{1}{n}\sum_{i=1}^{n}\ \ln\left(X_i\right) - \ln\left(X_{i-1}\right)\\ &= \frac{1}{n}\ \ln\left(X_n\right) - \ln\left(X_0\right)\end{aligned}$$

luego el estimador solo utiliza el primer y último dato del precio para estimar su tendencia, y es muestra de la dificultad inherente a la estimación del comportamiento que sigue el mercado.

Por otro lado, podemos observar que el estimador de la varianza de los retornos dado por la expresión (9) resulta ser un estimador sesgado, pero esto puede resolverse multiplicando la expresión por $\frac{n-1}{n}$ si el tamaño de muestra es suficientemente grande.

El método de máxima verosimilitud es lo que en estadística se conoce como un método de estimación puntual, ya que la expresión que se obtiene de su aplicación permite calcular un unico valor estimado del parámetro, pero también resulta de utilidad conocer un intervalo dentro del cual podría estar el valor del parámetro con un determinado nivel de confianza.

Para el caso del movimiento browniano geométrico estos intervalos pueden ser determinados de forma cerrada, tanto para la media como para la varianza, basados en que la suma de variables aleatorias gaussianas independientes es de nuevo una variable gaussiana.

Se tiene, entonces, que un intervalo del $100(1-\alpha)$ % de confianza para la media *m* está dado por:

$$\overline{y} - z_{\alpha/2}\frac{\sqrt{v}}{\sqrt{n}} \leq m \leq \overline{y} + z_{\alpha/2}\frac{\sqrt{v}}{\sqrt{n}} \qquad (12)$$

Para el caso de la varianza el intervalo de confianza está dado por la expresión:

$$\frac{n}{Q_u^x}\hat{v}\,''\; v''\;\frac{n}{Q_l^x}\hat{v} \qquad (13)$$

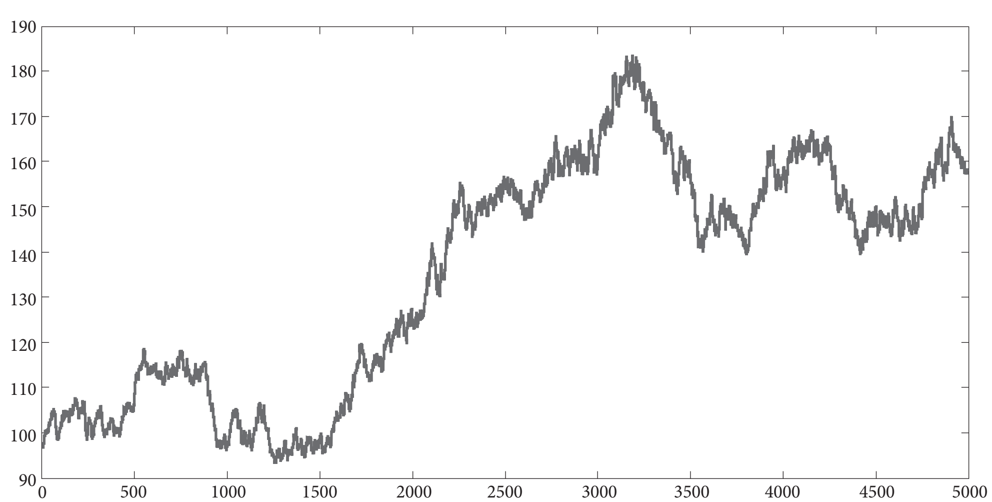
donde $Q_l^x$ y $Q_u^x$ son los cuantiles de la distribución $\chi^2$ con $n$ grados de libertad correspondientes al nivel de confianza deseado.

Como ejemplo ilustrativo del uso de los estimadores, consideremos la siguiente serie de observaciones generadas a partir del movimiento browniano geométrico descrito por la ecuación:

$$dX_t = 0{,}83X_t dt + 0{,}42X_t dW_t \qquad (14)$$

Se realiza la simulación de una trayectoria del proceso con 5000 pasos, utilizando el programa MATLAB, con un valor inicial del proceso de precio de 100, y se determina el valor de los estimadores para el coeficiente de tendencia y de difusión. La Figura 1, muestra la gráfica de la trayectoria simulada.

Figura 1. Trayectoria del movimiento browniano geométrico



El valor de los estimadores para este conjunto de observaciones es:

$$\hat{\beta} = 0,5621 \;\; ; \;\; \hat{\sigma}^2 = 0,4118$$

lo que confirma las observaciones realizadas sobre las bondades y defectos de los estimadores.

## 2.2. El modelo Vasicek

El modelo Vasicek es descrito por la EDE:

$$dW_t = \left(\alpha - \beta X_t\right) dt + \sigma dW_t \tag{15}$$

el cual, por lo general, es utilizado para modelar el comportamiento de la tasa corta de interés y, en general, de procesos que presentan la característica de revertir a la media. Esto se debe a que dependiendo del valor del proceso y el de las constantes $\alpha$ y $\beta$, el valor de la tendencia puede ser positivo o negativo.

Por la aplicación de la fórmula de Itô multidimencional al modelo Vasicek se tiene que[3]:

---

[3] Se puede verificar que esta expresión satisface la EDE del modelo Vasicek considerando:

$$f(t,x) = e^{-\beta t} X_s + \frac{\alpha}{\beta}\left(1 - e^{-\beta t}\right) + \sigma e^{-\beta t} x$$

y

$$R_t = \int_s^t e^{\beta u} dW_u$$

como:

$$\frac{\partial f(t,x)}{\partial t} = -\beta e^{-\beta} X_s + \alpha e^{-\beta t} - \sigma \beta e^{-\beta t} x$$

y

$$f_x(t,x) = \sigma e^{-\beta t} \;\; : \;\; f_{xx}(t,x) = 0$$

entonces:

$$df\left(f, R_t\right) = \alpha - \beta f\left(t, R_t\right) dt + \sigma e^{-\beta t} dR_t$$

de donde:

$$dX_t = \left(\alpha - \beta X_t\right) dt + \sigma dW_t$$

$$X_t = e^{-\beta(t-s)} X_s + \frac{\alpha}{\beta}\left(1 - e^{-\beta(t-s)}\right) + \sigma e^{-\beta t} \int_s^t e^{\beta u}\, dW_u \tag{16}$$

para $0 < s < t$. Se tiene entonces que $X_t$ condicionada a $F_s$ tiene una distribución normal con media y varianza dadas por:

$$E\left[X_t \middle| F_s\right] = \frac{\alpha}{\beta}\left(1 - e^{-\beta(t-s)}\right) + X_s e^{-\beta(t-s)} \tag{17}$$

$$V\left[X_t \middle| F_s\right] = \frac{\sigma^2}{2\beta}\left(1 - e^{-2\beta(t-s)}\right) \tag{18}$$

El modelo discreto correspondiente a la ecuación (16) sobre una partición $0 = t_0 < t_1 < t_2 < \ldots$ con tamaño de paso constante $\Delta t = t_i - t_{i-1}$ es desarrollado por Phillips en 1972 y corresponde a la expresión:

$$X_{t_i} = \frac{\alpha}{\beta}\left(1 - e^{-\beta\Delta t}\right) + e^{-\beta\Delta t} X_{t_{i-1}} + \sigma\sqrt{\left(1 - e^{-2\beta\Delta t}\right)/\, 2\beta} Z \tag{19}$$

con $Z \sim N(0, 1)$. Esta ecuación se puede denotar por:

$$X_{t_i} = c + bX_{t_{i-1}} + \delta Z \tag{20}$$

donde,

$$c = \frac{\alpha}{\beta}\left(1 - e^{-\beta\Delta t}\right)\ ;\ b = e^{-\beta\Delta t}\ ;\ \delta = \sigma\sqrt{\left(1 - e^{-2\beta\Delta t}\right)/\, 2\beta}$$

Si se denota $\theta = \frac{\alpha}{\beta}$ y se utiliza el método de máxima verosimilitud, aplicable en forma cerrada, dado que la función de transición del proceso es normal, se tienen los siguientes estimadores, donde $n$ corresponde al número de observaciones:

$$\hat{b} = \frac{n\sum_{i=1}^{n} x_i x_{i-1} - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_{i-1}}{n\sum_{i=1}^{n} x_{i-1}^2 - \left(\sum_{i=1}^{n} x_{i-1}\right)} \tag{21}$$

$$\hat{\theta} = \frac{\sum_{i=1}^{n}\left(x_i \hat{b} x_{i-1}\right)}{n\left(1-\hat{b}\right)} \tag{22}$$

$$\hat{\delta}^2 = \frac{1}{n}\sum_{i=1}^{n} \; x_i - \hat{b}x_{i-1} - \hat{\theta}\left(1-\hat{b}\right)^{\;2} \tag{23}$$

Como se tiene que:

$$\beta = -\frac{\text{In}(b)}{\Delta t} \;\; ; \;\; \alpha = -\frac{\theta \text{In}(b)}{\Delta t} \;\; ; \;\; \sigma = \frac{\delta}{\sqrt{\left(1-b^2\right)/2\beta}} \tag{23}$$

las expresiones (21), (22) y (23) permiten tener estimadores para los parámetros del proceso.

## 3. Conclusiones

Si bien los casos considerados para la estimación de parámetros en este escrito son elementales y de aplicación directa dado que se conoce la función de densidad del proceso, permiten describir el procedimiento de máxima verosimilitud aplicado a la estimación de parámetros en EDE que modelan variables de carácter financiero, y permiten reconocer las extensiones necesarias para la calibración de procesos en otras condiciones.

Se abre, entonces, la pregunta acerca de cómo establecer estimadores de los parámetros en EDE cuando las distribuciones asociadas al proceso no se conocen. A este respecto, si bien la literatura propone una serie de acercamientos, muestra también un amplio campo de estudio para el desarrollo de trabajos alrededor de la calibración de EDE.

## Referencias

[1] Black, F. and Scholes, M. (1976). "The Pricing of Options and Corporate Liabilities". *Journal of Political Economy* 81, 637-659.

[2] Brigo, D. and Mercurio, F. (2006). "Interest Rate Models: Theory and Practice". Springer Verlang.

[3] Brigo, D., Dalessandro, A., Neugebauer M. and Triki, F. (2007). "A Stochastic processes toolkit for Risk Mangement". *Working Paper*.

[4] Klebaner, Fima C. (2005). "Introduction to Stochastic Calculus with Applications", segunda edición. Imperial College Press.

[5] Korn, R. y Korn, E. (2001). "Option Pricing and Portfolio Optimization. Modern Methods of Financial Mathematics". A.M.S.

[6] Phillips, P.C.B. (1972). "The Structural Estimation of a Stochastic Differential Equatiuon System". *Econometrica*, 40, pp. 1021-1041.

[7] Phillips, P.C.B. and Yu, J. (2007). *Maximum likelihood and Gaussian estimation of continuous time models in finance*. Cowles Foundation for Research in Economics, Yale University.

[8] Shreve, Steven E. (2004). *Stochastic Calculus for Finance II*. Springer Finance.

[9] Vasicek, O.A. (1977). "An equilibrium characterization of the term structure". *Journal of Financial Economics* 5, pp. 177-188.