

Diversas razões dificultam a estimação de difusões com saltos. A primeira, e mais óbvia, é a existência de amostras discretas dos dados, enquanto o modelo é formulado em tempo contínuo. Isso leva, via de regra, ao uso de aproximações discretas do processo estocástico, tais como a aproximação de Euler ou de Bernoulli, como em Ball & Torous (1983). Por se tratarem de aproximações, resultam em modelos mal especificados<sup>8</sup>.

Dentre os diversos métodos para estimação de difusões com saltos tais como o Método Generalizado dos Momentos (GMM), Métodos de Monte Carlo baseados em Cadeias de Markov (MCMC) e Estimação por Máxima Verossimilhança (MLE), adota-se a estimação por Máxima Verossimilhança (MLE). Tal escolha deu-se, principalmente, pela possibilidade do uso da função característica no lugar da função de distribuição, sem perda de qualidade nos resultados<sup>9</sup>. Há, na literatura, diversas soluções analíticas para a função característica de processos estocásticos, frente a poucas soluções que resultem na expressão da função de densidade.

#### a. Função característica

Uma vez que o interesse é determinar a distribuição futura de  $r(t)$  dada uma condição inicial  $r(0)$ , deve-se obter a função característica  $F(r, T; s)$ , onde  $s$  é o parâmetro da função característica. Para tanto, é necessário obter a expressão da variação esperada em  $F(r, T; s)$ , que pela fórmula de Itô é dada por:

$$dF = \frac{\partial F}{\partial r} dr + \frac{\partial F}{\partial T} dT + \frac{1}{2} \frac{\partial^2 F}{\partial r^2} dr^2$$

<sup>8</sup> Decorrente da má especificação pode surgir viés assintótico, arbitrariamente grande, nos estimadores.

<sup>9</sup> Garantida pela relação unívoca entre a função de distribuição e a função característica.

Fazendo-se a substituição da equação 1 na igualdade acima:

$$dF = \frac{\partial F}{\partial r} [k(\theta - r)dT + v dW + J d\pi(h)] + \frac{\partial F}{\partial T} dT \\ + \frac{1}{2} \frac{\partial^2 F}{\partial r^2} [k(\theta - r)dT + v dW + J d\pi(h)]^2$$

Conforme Dixit & Pindyck (1994), cabe destacar o fato de que variações na taxa impactam a função característica de três maneiras – deterministicamente em função da reversão à média e estocasticamente em função da variância da taxa e do processo de Poisson. Como a probabilidade da ocorrência de um evento de Poisson em  $dT$  é igual a  $h \cdot dT$ , tem-se que:

$$E \left[ \frac{\partial F}{\partial r} \cdot J d\pi(h) \right] = h \cdot E[F(r + J) - F(r)] dT$$

A equação reversa de Kolmogorov, usando o lema de Itô, será:

$$\frac{\partial F}{\partial T} = \frac{\partial F}{\partial r} k(\theta - r) + \frac{1}{2} \frac{\partial^2 F}{\partial r^2} v^2 + h \cdot E[F(r + J) - F(r)] \quad (3)$$

Sujeita a condição de contorno:

$$F(r, 0; s) = e^{i.s.r} ; \quad i = \sqrt{-1}$$

Conforme sugerido por Duffie & Kan (1996), a solução toma forma funcional do tipo<sup>10</sup>:

$$F(r, T; s) = e^{[A(T; s) + B(T; s) \cdot r]} \quad (4)$$

Cujas derivadas parciais são:

$$F_r = \frac{\partial F}{\partial r} \Rightarrow$$

$$\frac{\partial F}{\partial r} = e^{[A(T; s) + B(T; s) \cdot r]} \cdot \frac{\partial [A(T; s) + B(T; s) \cdot r]}{\partial r} \Rightarrow$$

$$\frac{\partial F}{\partial r} = e^{[A(T; s) + B(T; s) \cdot r]} \cdot B(T; s) \Rightarrow$$

$$F_r = B \cdot F$$

$$F_{rr} = \frac{\partial^2 F}{\partial r^2} \Rightarrow$$

$$\frac{\partial^2 F}{\partial r^2} = \frac{\partial}{\partial r} \left( \frac{\partial F}{\partial r} \right) \Rightarrow$$

$$\frac{\partial^2 F}{\partial r^2} = \frac{\partial}{\partial r} \left( e^{[A(T; s) + B(T; s) \cdot r]} \cdot B(T; s) \right) \Rightarrow$$

$$\frac{\partial^2 F}{\partial r^2} = \left( e^{[A(T; s) + B(T; s) \cdot r]} \cdot B(T; s) \right) \cdot \frac{\partial [A(T; s) + B(T; s) \cdot r]}{\partial r} \Rightarrow$$

$$\frac{\partial^2 F}{\partial r^2} = \left[ e^{[A(T; s) + B(T; s) \cdot r]} \cdot B(T; s) \right] \cdot B(T; s) \Rightarrow$$

$$F_{rr} = B^2 F$$

---

<sup>10</sup> Para que a solução tome a forma exponencial afim, basta que o processo estocástico da taxa de juros de curto prazo seja afim, conforme Duffie & Kan (1996). Mais ainda, Duffie et al. (2000) demonstram que o processo de Ornstein-Uhlenbeck com saltos é um processo afim.

$$\begin{aligned}
F_T &= \frac{\partial F}{\partial T} \Rightarrow \\
\frac{\partial F}{\partial T} &= [e^{[A(T;s)+B(T;s) \cdot r]}] \cdot \frac{\partial [A(T;s) + B(T;s) \cdot r]}{\partial T} \Rightarrow \\
\frac{\partial F}{\partial T} &= [e^{[A(T;s)+B(T;s) \cdot r]}] \cdot \left[ \frac{\partial A}{\partial T} + \frac{\partial B}{\partial T} \cdot r \right] \Rightarrow \\
F_T &= F(A_T + B_T r); \quad \frac{\partial A}{\partial T} \equiv A_T \wedge \frac{\partial B}{\partial T} \equiv B_T
\end{aligned}$$

Substituindo tais derivadas parciais e a solução proposta por Duffie & Kan (1996) na equação 3, tem-se:

$$0 = r(-kB - B_T) + \left( k\theta B + \frac{1}{2} v^2 B^2 - A_T + hE[e^{JB} - 1] \right) \quad (5)$$

Da equação 5, resultam duas equações diferenciais ordinárias:

$$\begin{aligned}
B_T &= -k \cdot B \\
A_T &= k\theta B + \frac{1}{2} v^2 B^2 + hE[e^{JB} - 1]
\end{aligned}$$

Tais equações podem ser resolvidas com as condições de contorno apropriadas. Da condição de contorno para a equação reversa de Kolmogorov e a solução proposta na equação 4, tem-se que:

$$A(0; s) = 0$$

$$B(0; s) = i \cdot s$$

Condições suficientes para resolver as equações diferenciais ordinárias resultantes da equação 5:

$$B(T; s) = i \cdot s \cdot e^{-k \cdot T} \quad (6.1)$$

$$A(T; s) = \int \left( k \theta B(T; s) + \frac{1}{2} v^2 B(T; s)^2 + h E[e^{JB(T; s)} - 1] \right) dT \quad (6.2)$$

Retomando a equação 4, a função característica será, então:

$$F(r, T; s) = e^{\left[ \int \left( k \theta B(T; s) + \frac{1}{2} v^2 B(T; s)^2 + h E[e^{JB(T; s)} - 1] \right) dT + i \cdot s \cdot e^{-k \cdot T} \cdot r \right]}$$

## b. Momentos

Uma vez determinada a função característica, os momentos podem ser obtidos por meio de sucessivas diferenciações com respeito a  $s$ , conforme Magalhães (2006):

$$\left. \frac{\partial^n F}{\partial s^n} \right|_{s=0} \equiv F_n = i^n \cdot \mu_n ; n = 1, 2, \dots \quad (7)$$

Onde  $\mu_n$  representa o  $n$ -ésimo momento central. Como  $dF/ds = dA/ds + r \cdot dB/ds$ , deve-se proceder a diferenciação de  $A$  e  $B$  com respeito a  $s$ . Definindo  $A_n \equiv d^n A/ds^n$  e  $B_n \equiv d^n B/ds^n$ , tem-se<sup>11</sup>:

$$\frac{dA}{ds} = \int \left( k \cdot \theta \cdot i \cdot e^{-k \cdot T} - v^2 \cdot s \cdot e^{-2 \cdot k \cdot T} + h \cdot i \cdot e^{-k \cdot T} \cdot E \left[ J \cdot e^{J \cdot s \cdot i \cdot e^{-k \cdot T}} \right] \right) dT$$

$$\left. \frac{dA}{ds} \right|_{s=0} = \int i (k \cdot \theta \cdot e^{-k \cdot T} + h \cdot e^{-k \cdot T} \cdot E[J]) dT =$$

<sup>11</sup> Fazendo-se uso, ainda, que a integral de  $A(T; s)$  e  $E[e^{JB(T; s)} - 1]$  são limitadas.

$$= i. \left( -\theta e^{-kT} - \frac{h}{k} E[J] e^{-kT} \right) + c_1$$

Pela condição de contorno  $A(0; s) = 0$ , pode-se obter  $c_1$ :

$$0 = i. \left( -\theta - \frac{h}{k} E[J] \right) + c_1 \Leftrightarrow$$

$$\Leftrightarrow c_1 = i. \left( \theta + \frac{h}{k} E[J] \right)$$

Que, substituindo, resulta em:

$$\begin{aligned} \frac{dA}{ds} \Big|_{s=0} &= i. \left( -\theta e^{-kT} - \frac{h}{k} E[J] e^{-kT} \right) + i. \left( \theta + \frac{h}{k} E[J] \right) = \\ &= i. \left[ (1 - e^{-kT}). \left( \theta + \frac{h}{k} E[J] \right) \right] \end{aligned}$$

De forma semelhante, pode-se obter a segunda derivada:

$$\begin{aligned} \frac{d^2 A}{ds^2} &= \int \left( -v^2. e^{-2.k.T} - h. e^{-2.k.T}. E \left[ J^2. e^{J.s.i.e^{-kT}} \right] \right) dT \\ \frac{d^2 A}{ds^2} \Big|_{s=0} &= \int \left( -v^2. e^{-2.k.T} - h. e^{-2.k.T}. E[J^2] \right) dT = \\ &= \frac{v^2}{2.k}. e^{-2.k.T} + \frac{h}{2.k}. e^{-2.k.T}. E[J^2] + c_2 \end{aligned}$$

Que pode novamente pela condição de contorno  $A(0; s) = 0$ , ser computada:

$$0 = \frac{v^2}{2.k} \cdot e^{-2.k.T} + \frac{h}{2.k} \cdot e^{-2.k.T} \cdot E[J^2] + c_2 \Leftrightarrow$$

$$\Leftrightarrow c_2 = -\frac{v^2 + hE[J^2]}{2.k}$$

Logo, tem-se que:

$$\left. \frac{d^2 A}{ds^2} \right|_{s=0} = \frac{v^2}{2.k} \cdot e^{-2.k.T} + \frac{h}{2.k} \cdot e^{-2.k.T} \cdot E[J^2] - \frac{v^2 + hE[J^2]}{2.k} =$$

$$= -(1 - e^{-2.k.T}) \cdot \left( \frac{v^2 + hE[J^2]}{2.k} \right)$$

A terceira derivada:

$$\frac{d^3 A}{ds^3} = \int \left( -i.h.e^{-3.k.T} \cdot E[J^3 \cdot e^{J.s.i.e^{-kT}}] \right) dT$$

$$\left. \frac{d^3 A}{ds^3} \right|_{s=0} = \int (-i.h.e^{-3.k.T} \cdot E[J^3]) dT =$$

$$= i.\frac{h}{3.k} \cdot e^{-3.k.T} \cdot E[J^3] + c_3$$

Onde  $c_3 = -i.(h/3k).E[J^3]$ . Substituindo-se na equação acima, tem-se:

$$\left. \frac{d^3 A}{ds^3} \right|_{s=0} = -i.h.E[J^3] \cdot \left( \frac{1 - e^{-3.k.T}}{3.k} \right)$$

E, por último, a quarta derivada:

$$\begin{aligned}\frac{d^4 A}{ds^4} &= \int \left( -h \cdot e^{-4.k.T} \cdot E \left[ J^4 \cdot e^{J.s.i.e^{-kT}} \right] \right) dT \\ \left. \frac{d^4 A}{ds^4} \right|_{s=0} &= \int (-h \cdot e^{-4.k.T} \cdot E[J^4]) dT = \\ &= -\frac{h}{4.k} \cdot e^{-4.k.T} \cdot E[J^4] + c_4\end{aligned}$$

Com  $c_4 = (h/4k) \cdot E[J^4]$  . Analogamente:

$$\left. \frac{d^4 A}{ds^4} \right|_{s=0} = h \cdot E[J^4] \cdot \left( \frac{1 - e^{-4.k.T}}{4.k} \right)$$

As derivadas de  $B$  com respeito a  $s$  são:

$$\begin{aligned}\frac{dB}{ds} &= i \cdot e^{-kT} \\ \frac{d^2 B}{ds^2} &= \frac{d^3 B}{ds^3} = \frac{d^4 B}{ds^4} = 0\end{aligned}$$

Retomando o fato que  $dF/ds = dA/ds + r \cdot dB/ds$  na equação 7, tem-se:

$$\begin{aligned}\mu_1 &= \frac{1}{i} \cdot \left( \frac{dA}{ds} + r \cdot \frac{dB}{ds} \right) \Big|_{s=0} = \\ &= \left( \theta + h \cdot \frac{E[J]}{k} \right) \cdot (1 - e^{-kT}) + r \cdot e^{-kT}\end{aligned}$$

De forma análoga é possível obter os demais momentos:

$$\begin{aligned}
\mu_2 &= \frac{1}{i^2} \cdot \left( e^{A+B \cdot r} \cdot \left( \frac{d^2 A}{ds^2} \cdot \left( \frac{dA}{ds} + r \cdot \frac{dB}{ds} \right)^2 \right) \right) \Big|_{s=0} = \\
&= \frac{v^2 + h \cdot E[J^2]}{2 \cdot k} \cdot (1 - e^{-2kT}) + \mu_1^2 \\
\mu_3 &= \frac{1}{i^3} \cdot \left( \frac{d^3 A}{ds^3} + 3 \cdot \frac{d^2 A}{ds^2} \cdot \left( \frac{dA}{ds} + r \cdot \frac{dB}{ds} \right) + \left( \frac{dA}{ds} + r \cdot \frac{dB}{ds} \right)^3 \right) \Big|_{s=0} = \\
&= h \cdot E[J^3] \cdot \left( \frac{1 - e^{-3kT}}{3 \cdot k} \right) + 3 \cdot \mu_1 \cdot (v^2 + h \cdot E[J^2]) \cdot \left( \frac{1 - e^{-2kT}}{2 \cdot k} \right) + \mu_1^3 \\
\mu_4 &= \frac{1}{i^4} \cdot \left( \frac{d^4 A}{ds^4} + 3 \cdot \left( \frac{d^2 A}{ds^2} \right)^2 + 4 \cdot \frac{d^3 A}{ds^3} \cdot \left( \frac{dA}{ds} + r \cdot \frac{dB}{ds} \right) + 6 \cdot \frac{d^2 A}{ds^2} \cdot \left( \frac{dA}{ds} + r \cdot \frac{dB}{ds} \right)^2 \right. \\
&\quad \left. + \left( \frac{dA}{ds} + r \cdot \frac{dB}{ds} \right)^4 \right) \Big|_{s=0} = \\
&= h \cdot E[J^4] \cdot \left( \frac{1 - e^{-4kT}}{4 \cdot k} \right) + 3 \cdot \left( (v^2 + h \cdot E[J^2]) \cdot \left( \frac{1 - e^{-2kT}}{2 \cdot k} \right) \right)^2 \\
&\quad + 4 \cdot \mu_1 \cdot h \cdot E[J^3] \cdot \left( \frac{1 - e^{-3kT}}{3 \cdot k} \right) \\
&\quad + 6 \cdot \mu_1^2 \cdot \left( (v^2 + h \cdot E[J^2]) \cdot \left( \frac{1 - e^{-2kT}}{2 \cdot k} \right) \right) + \mu_1^4
\end{aligned}$$

A partir dos momentos calculados acima, pode-se computar coeficiente de assimetria e curtose (teóricos) do processo, conforme Das (2002):

$$\begin{aligned}
\text{Coeficiente de Assimetria} \equiv \gamma_1 &= \frac{E(J - \mu_1)^3}{(\mu_2 - \mu_1^2)^{\frac{3}{2}}} \Rightarrow \\
\gamma_1 &= \frac{2 \cdot \sqrt{2 \cdot k} \cdot e^{-kT} \cdot (1 + e^{kT} + e^{2kT}) \cdot h \cdot E[J^3]}{3 \cdot (1 + e^{kT}) \cdot (v^2 + h \cdot E[J^2]) \cdot \sqrt{(1 - e^{-2kT}) \cdot (v^2 + h \cdot E[J^2])}}
\end{aligned}$$

Na ausência de saltos, ou seja  $h = 0$ , tem-se que:

$$\gamma_1 = \frac{2 \cdot \sqrt{2 \cdot k} \cdot e^{-kT} \cdot (1 + e^{kT} + e^{2kT}) \cdot 0 \cdot E[J^3]}{3 \cdot (1 + e^{kT}) \cdot (v^2 + 0 \cdot E[J^2]) \cdot \sqrt{(1 - e^{-2kT}) \cdot (v^2 + 0 \cdot E[J^2])}}$$

$$\gamma_1 = 0$$

$$\text{Curtose} \equiv \gamma_2 = \frac{E(J - \mu_1)^4}{(\mu_2 - \mu_1^2)^2} \Rightarrow$$

$$\gamma_2 = \frac{(e^{2kT} - 1) \cdot (3h^2 \cdot E[J^2]^2 + 6h \cdot v^2 \cdot E[J^2] + 3v^4) + kh \cdot E[J^4] \cdot (e^{2kT} + 1)}{(e^{2kT} - 1) \cdot (v^2 + h \cdot E[J^2])^2}$$

Novamente, supondo a ausência de saltos tem-se que:

$$\gamma_2 = \frac{(e^{2kT} - 1) \cdot (3 \cdot 0 \cdot E[J^2]^2 + 6 \cdot 0 \cdot v^2 \cdot E[J^2] + 3v^4) + k \cdot 0 \cdot E[J^4] \cdot (e^{2kT} + 1)}{(e^{2kT} - 1) \cdot (v^2 + 0 \cdot E[J^2])^2} \Rightarrow$$

$$\gamma_2 = \frac{(e^{2kT} - 1) \cdot (3v^4)}{(e^{2kT} - 1) \cdot (v^2)^2} \Rightarrow$$

$$\gamma_2 = 3$$

A derivação dos momentos constitui-se uma prática amplamente utilizada na literatura, até mesmo para uma conferência da qualidade das estimativas desenvolvidas. No presente estudo, por exemplo, foram obtidos resultados consistentes (i.e. assimetria e curtose) condizentes com um modelo puramente gaussiano na ausência de saltos. Os momentos aqui obtidos podem, ainda, ser utilizados em trabalhos subseqüentes para estimação dos parâmetros pelo Método Generalizado dos Momentos.

### c. Verossimilhança

Para proceder à estimação de um processo estocástico contínuo com observações em períodos discretos, será utilizada uma aproximação discreta do processo descrito pela equação 1. Para estimar difusões com saltos, segundo Chan (2005), há duas grandes alternativas: o uso de misturas discretas de distribuição (1) Poisson-Normal e (2) Normal-Normal. As limitações decorrentes do truncamento da distribuição de Poisson, apontadas por Vlaar & Palm (1993), justificam a opção pelo arcabouço da mistura de Normal-Normal obtido pela aproximação de Bernoulli. Tal aproximação segue a estrutura proposta originalmente por Ball & Torous (1983) e utilizada também em Vlaar & Palm (1993) e Das (2002).

Para a aproximação de Bernoulli, segundo Ball & Torous (1983), supõe-se que a cada instante de tempo o salto na taxa de juros ocorre ou não. Essa suposição é razoável em intervalos de tempo pequenos. Em um dia, por exemplo, não há a ocorrência de um salto ou então apenas um salto ocorre.

Seja  $M(T)$  o número de saltos que ocorrem no intervalo de tempo de tamanho  $T$ . Dividindo-se, ainda, o intervalo de tempo  $(0, T)$  em  $N$  intervalos menores de tamanho  $\Delta t$ , ou seja,  $\Delta t = T/N$  e sendo  $Y_i$  o número de saltos no subintervalo  $i$ , temos que<sup>12</sup>:

$$M(T) = \sum_{i=1}^N Y_i$$

Portanto,  $M(T)$  é a soma de  $N$  variáveis aleatórias independentes e identicamente distribuídas (i.i.d.) com as seguintes propriedades:

- a.  $M(0) = 0$
- b.  $Prob[Y_i = 0] = 1 - h \cdot \Delta t + o(\Delta t)$
- c.  $Prob[Y_i = 1] = h \cdot \Delta t + o(\Delta t)$
- d.  $Prob[Y_i > 1] = o(\Delta t)$

Onde uma função  $f$  é do tipo  $o(k)$ , ou de “ordem  $k$ ”, se:

---

<sup>12</sup> Utilizando a suposição de incrementos independentes.

$$\lim_{k \rightarrow 0} \frac{f(k)}{k} = 0$$

Para  $N$  grande, a probabilidade de um salto no intervalo  $i$  tende a ser desprezível, de forma que cada  $Y_i$  é aproximadamente uma Bernoulli com parâmetro  $h.T/N$ . Sendo  $M(T)$  uma soma de variáveis aleatórias Bernoulli i.i.d. (*Bernoulli*( $q$ ), onde  $q = \text{Prob}[Y_i = 1] = h.\Delta t + o(\Delta t)$ ),  $M(T)$  está distribuída segundo a distribuição binomial, de forma<sup>13</sup>:

$$\begin{aligned} \text{Prob}[M = k] &\cong \binom{N}{k} \cdot \left(\frac{h.T}{n}\right)^k \cdot \left(1 - \frac{h.T}{n}\right)^{N-k} \\ \lim_{N \rightarrow \infty} \binom{N}{k} \cdot \left(\frac{h.T}{n}\right)^k \cdot \left(1 - \frac{h.T}{n}\right)^{N-k} &= \frac{e^{-h.T} \cdot (hT)^k}{k!} \end{aligned}$$

Resultado que demonstra como a aproximação de Bernoulli converge para a densidade da distribuição de Poisson.

Segundo Das (2002), a aproximação da função de densidade de probabilidade de um processo estocástico de salto-difusão Poisson pode ser obtida usando a mistura Bernoulli Normal-Normal. A versão discreta proposta pelo autor para a equação 1 é:

$$\Delta r = k.(\theta - r).\Delta t + v.\Delta W + J(\mu, \gamma^2).\Delta \pi(q)$$

Onde  $\Delta W \sim N(0,1)$ ,  $J \sim N(\mu, \gamma^2)$ ,  $\Delta \pi(q)$  é a aproximação de Bernoulli com parâmetro  $q = \text{Prob}[Y_i = 1] = h.\Delta t + o(\Delta t)$ . Dessa forma, então, tem-se que:

$$\begin{aligned} f[r(s)|r(t)] &= q \cdot \frac{1}{\sqrt{2.\pi.(v^2.\Delta t + \gamma^2)}} \cdot e^{\left[-\frac{(r(s)-r(t)-k.(\theta-r).\Delta t-\mu)^2}{2.(v^2.\Delta t+\gamma^2)}\right]} + \\ &+ (1-q) \cdot \frac{1}{\sqrt{2.\pi.v^2.\Delta t}} \cdot e^{\left[-\frac{(r(s)-r(t)-k.(\theta-r).\Delta t)^2}{2.v^2.\Delta t}\right]} \end{aligned} \quad (8)$$

<sup>13</sup> Para  $k$  realizações, onde  $k = 0, 1, 2, \dots, N$ .

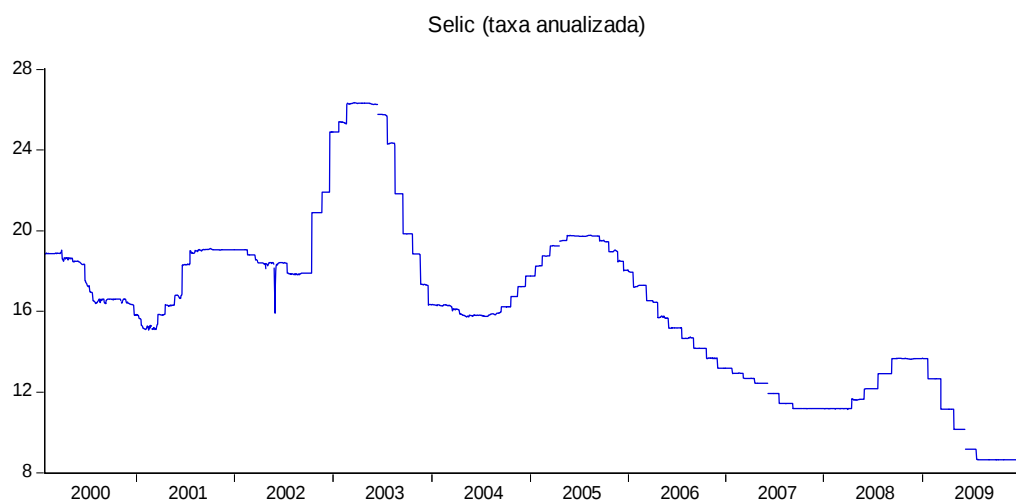
A função de verossimilhança pode ser escrita, então, como:

$$L = \prod_{t=1}^T f[r(t + \Delta t)|r(t)]$$

#### d. Dados

O conjunto de dados utilizado nas estimações é referente à taxa de juros Selic coletada da base histórica do Banco Central do Brasil (2010). A amostra empregada contém 2512 observações diárias da taxa, realizada entre 24/01/2000 e 22/01/2010.

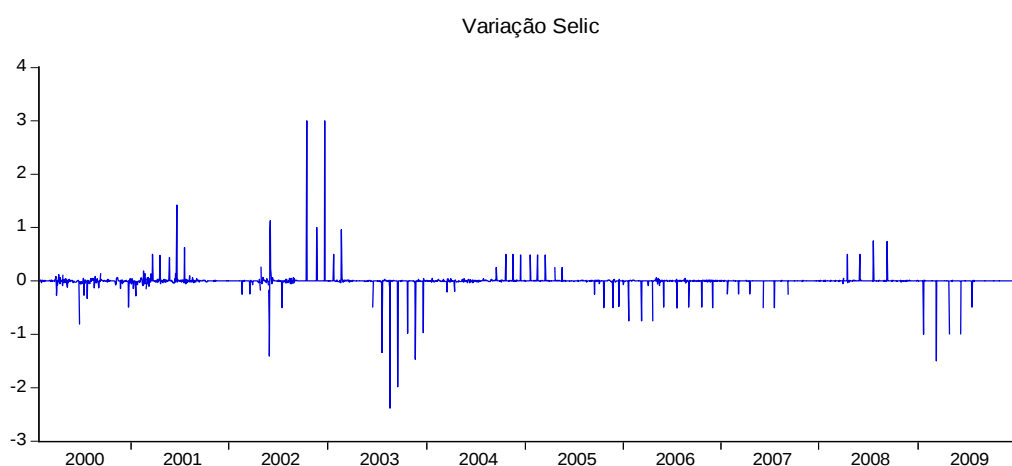
No período supracitado, a taxa Selic anualizada oscilou entre 8,64%, em final de 2009, e 26,35%, no começo do primeiro mandato do presidente Luiz Inácio Lula da Silva, em 2003. Pode-se observar a trajetória da taxa, com destaque aos pontos de “descontinuidade”<sup>14</sup>, no gráfico abaixo:



**Figura 3**

<sup>14</sup> Em 2003 e 2009, mais acentuadamente, o que pode sugerir a ocorrência de saltos.

Podem-se observar, no gráfico abaixo, os movimentos de variação<sup>15</sup> na taxa. Os períodos de 2002 e 2003 mostram maiores oscilações, relacionadas, provavelmente, com a incerteza acerca da política monetária no período eleitoral e começo do primeiro mandato Lula.



**Figura 4**

Abaixo, podem-se observar as estatísticas descritivas da taxa Selic e sua variação:

**Tabela 1**

|                      | Selic   | Var. Selic |
|----------------------|---------|------------|
| <b>Média</b>         | 16,1963 | -0,0041    |
| <b>Mediana</b>       | 16,3400 | 0,0000     |
| <b>Máximo</b>        | 26,3500 | 3,0000     |
| <b>Mínimo</b>        | 8,6400  | -2,3800    |
| <b>Desvio-Padrão</b> | 4,1284  | 0,1542     |
| <b>Assimetria</b>    | 0,3087  | 2,4218     |
| <b>Curtose</b>       | 3,0409  | 172,3494   |
|                      |         |            |
| <b>Observações</b>   | 2512    | 2511       |

O excesso de curtose na variação da taxa, segundo Das (2002), é um indicativo do uso de modelos com saltos. Mais ainda, em Becker (1991), fica

<sup>15</sup> Aqui definida como  $Selc_{t+1} - Selc_t$ , ou seja, a primeira diferença da variável.

evidente a falha de modelos gaussianos puros para explicar a taxa de juros de curto prazo.

Procedendo-se ao teste de Jarque-Bera, cuja hipótese nula é de normalidade nos dados, temos:

**Tabela 2**

|                             | <b>Selic</b> | <b>Var. Selic</b> |
|-----------------------------|--------------|-------------------|
| <b>Estatística do teste</b> | 40,0717      | 300301            |
| <b>P-valor</b>              | 0,0000       | 0,0000            |

Fica claro, pelos elevados valores da estatística do teste e o p-valor nulo (em 9 casas decimais para a taxa e absoluto para sua primeira diferença), a rejeição da hipótese nula tanto para a taxa quanto para sua variação.

Outro teste realizado, ainda, foi o teste Augmented Dickey-Fuller (ADF) com intercepto e cuja hipótese nula é de raiz unitária:

**Tabela 3**

|                             | <b>Selic</b> | <b>Var. Selic</b> |
|-----------------------------|--------------|-------------------|
| <b>Estatística do teste</b> | -0,6932      | -8,1549           |
| <b>P-valor</b>              | 0,8465       | 0,0000            |

Frente aos resultados, falha-se em rejeitar a hipótese nula na taxa Selic. Já com relação à variação, ou primeiro diferencial da taxa, rejeita-se a hipótese nula de raiz unitária.

## **e. Estimativas**

Uma vez definida a função de verossimilhança para os processos estocásticos, a estimação dos parâmetros, então, pode ser obtida pela maximização dessa função (máxima verossimilhança). Honore (1998) aponta que a função de verossimilhança para processos de salto-difusões pode possuir várias

singularidades<sup>16</sup>. Tal constatação resulta, então, na baixa confiabilidade de estimações feitas por máxima verossimilhança em salto-difusões.

Ainda, segundo Honore (1998), diversos artigos seminais<sup>17</sup> realizam estimações por máxima verossimilhança de salto-difusões sem se preocupar com tais problemas.

Para contornar o problema das singularidades, Honore (1998) propõe que seja imposta uma restrição no espaço de otimização – que as variâncias da mistura de normais sejam proporcionais entre si. Deste modo, caso tenhamos a mistura de duas normais  $N_1(\mu_1, \sigma_1)$  e  $N_2(\mu_2, \sigma_2)$  do tipo:

$$L = w \cdot f_{N_1(\mu_1, \sigma_1)} + (1 - w) \cdot f_{N_2(\mu_2, \sigma_2)}$$

A solução de Honore (1998) é impor que  $\sigma_1^2 = \alpha \cdot \sigma_2^2$ , com  $\alpha$  sendo uma constante não-nula e finita. Essa restrição impede a ocorrência de singularidades geradas por  $\sigma_1 = 0$  e  $\sigma_2 \neq 0$  (ou vice-versa). Mais ainda, respeitada tal condição é possível obter estimativas consistente e assintoticamente normais.

No caso, a função descrita na equação (8) é uma mistura de normais (ponderadas por  $q$ ) onde as variâncias são  $\sigma_1^2 = v^2 \Delta t + \gamma^2$  e  $\sigma_2^2 = v^2 \Delta t$ . Logo, tem-se que  $\sigma_1^2 = \sigma_2^2 \cdot (1 + \gamma^2)$ . Para satisfazer a solução proposta por Honore (1998), basta que  $(1 + \gamma^2)$  seja não-nulo (sempre satisfeita, dado que  $\gamma^2 \geq 0$ ) e finito<sup>18</sup>.

Para a estimação dos parâmetros, foram empregados<sup>19</sup> os algoritmos de otimização (Berndt-Hall-Hall-Hausman, Broyden-Fletcher-Goldfarb-Shanno e Simplex) do software Estima Rats®.

Como procedimento, realizou-se, para todos os processos, duas etapas de estimação: a primeira sem restrições nos parâmetros, com o algoritmo Simplex, para calibração dos valores iniciais da segunda estimação, essa com as devidas restrições no espaço de otimização e utilizando como condições iniciais os valores encontrados na etapa anterior. Uma particularidade relevante do algoritmo Simplex do software utilizado é que, via de regra, não é possível determinar o

<sup>16</sup> Quando uma das variâncias da mistura de normais tende a zero e a outra não, por exemplo.

<sup>17</sup> Tais como Jorion (1988) e Ball & Torous (1983).

<sup>18</sup> Satisfeita por hipótese (seção de verossimilhança).

<sup>19</sup> Os códigos estão disponíveis no apêndice.

desvio-padrão das estimativas. Os resultados de significância dos coeficientes foram possíveis, então, apenas para a segunda etapa de estimação

Para fins de comparação entre os modelos utilizaram-se os critérios de informação de Akaike e Schwarz. Segundo Ramezani & Zeng (2007), as grandes vantagens na utilização dos critérios de informação frente a testes como o de razão de verossimilhança, são: a possibilidade de comparação simultânea dos diversos modelos (versus comparação aos pares) e não necessidade de que os modelos sejam aninhados, isto é, que um modelo não seja um caso particular do outro.

Primeiramente otimizou-se da função de verossimilhança para o processo de Ornstein–Uhlenbeck, sem restrições nos parâmetros e os resultados encontram-se abaixo:

**Tabela 4**

| <b>Parâmetro</b>           | <b>Valor</b> |
|----------------------------|--------------|
| $k$                        | 0,029039     |
| $\theta$                   | -0,191356    |
| $v^2$                      | 0,000599     |
|                            |              |
| <b>Log-Verossimilhança</b> | 12690,69     |

Pode-se observar inconsistência econômica na estimativa (negativa) da média de longo prazo para a taxa ( $\theta$ ). Tal resultado indicaria, portanto, que a taxa Selic reverte à um valor negativo no longo prazo. Restrições relevantes são, então, que a média de longo prazo do processo ( $\theta$ ) seja não-negativa, bem como a taxa de reversão à essa média ( $k$ ).

A partir das restrições levantadas<sup>20</sup>, então, realizou-se nova rodada de otimização cujos resultados encontram-se abaixo:

<sup>20</sup> As restrições ativas para otimização foram  $\theta \geq 0$  e  $k \geq 0$ .

Tabela 5

| Parâmetro                        | Valor     | Desv. Pad. | Estat. t  | Sig.     |
|----------------------------------|-----------|------------|-----------|----------|
| $k$                              | -0,000001 | 0,000000   | -427,5930 | 0,000000 |
| $\theta$                         | 0,167200  | 0,019300   | 8,653250  | 0,000000 |
| $v^2$                            | 0,000599  | 0,000017   | 35,436230 | 0,000000 |
|                                  |           |            |           |          |
| Log-Verossimilhança              |           | 12689,80   |           |          |
| Critério de Informação de Akaike |           | -10,11     |           |          |
| Critério de Schwarz              |           | -10,10     |           |          |

De forma análoga, em uma primeira tentativa de otimização sem restrições para o processo de Ornstein–Uhlenbeck com a variância ARCH(1), temos os seguintes resultados:

Tabela 6

| Parâmetro           | Valor        |
|---------------------|--------------|
| $k$                 | -0,470999    |
| $\theta$            | 0,086037     |
| $b_0$               | 0,000000     |
| $b_1$               | 25424,115759 |
|                     |              |
| Log-Verossimilhança | 13320,32     |

Os estimadores para os parâmetros mostram-se inconsistentes novamente, com o parâmetro  $k$  violando a restrição de ser não-negativo<sup>21</sup> e  $b_1$  atingindo valor de alta magnitude. Realizando-se a otimização com as restrições levantadas anteriormente, foram obtidos os seguintes resultados:

<sup>21</sup> Conforme a definição de um processo de Ornstein-Uhlenbeck

Tabela 7

| Parâmetro                               | Valor      | Desv. Pad. | Estat. t | Sig.     |
|-----------------------------------------|------------|------------|----------|----------|
| $k$                                     | 0,000000   | 0,000000   | 0,000000 | 0,000000 |
| $\theta$                                | 0,056922   | 0,035567   | 1,600410 | 0,109506 |
| $b_0$                                   | 0,000481   | 0,000014   | 33,46524 | 0,000000 |
| $b_1$                                   | 856,299130 | 152,924483 | 5,599490 | 0,000000 |
|                                         |            |            |          |          |
| <b>Log-Verossimilhança</b>              |            | 12767,85   |          |          |
| <b>Critério de Informação de Akaike</b> |            | -10,17     |          |          |
| <b>Critério de Schwarz</b>              |            | -10,17     |          |          |

Já os resultados da estimação irrestrita dos parâmetros do processo de Ornstein–Uhlenbeck com saltos são:

Tabela 8

| Parâmetro                  | Valor     |
|----------------------------|-----------|
| $k$                        | 0,008491  |
| $\theta$                   | 0,065500  |
| $v^2$                      | 0,000001  |
| $q$                        | 0,163000  |
| $\mu$                      | -0,000230 |
| $\gamma^2$                 | 0,000015  |
|                            |           |
| <b>Log-Verossimilhança</b> | 18969,51  |

Os resultados com as restrições já descritas anteriormente encontram-se na tabela abaixo:

Tabela 9

| Parâmetro                               | Valor     | Desv. Pad. | Estat. t  | Sig.     |
|-----------------------------------------|-----------|------------|-----------|----------|
| $k$                                     | 0,008497  | 0,006464   | 1,314540  | 0,188664 |
| $\theta$                                | 0,065700  | 0,078200   | 0,839710  | 0,401069 |
| $v^2$                                   | 0,000001  | 0,000000   | 26,37221  | 0,000000 |
| $q$                                     | 0,163000  | 0,007452   | 21,87171  | 0,000000 |
| $\mu$                                   | -0,000230 | 0,000192   | -1,198410 | 0,230759 |
| $\gamma^2$                              | 0,000015  | 0,000001   | 14,61815  | 0,000000 |
|                                         |           |            |           |          |
| <b>Log-Verossimilhança</b>              |           | 18969,51   |           |          |
| <b>Critério de Informação de Akaike</b> |           | -15,13     |           |          |
| <b>Critério de Schwarz</b>              |           | -15,11     |           |          |

Por fim, a estimação irrestrita dos parâmetros do processo de Ornstein-Uhlenbeck com saltos e variância ARCH(1) é dada por:

Tabela 10

| Parâmetro                  | Valor     |
|----------------------------|-----------|
| $k$                        | -0,000001 |
| $\theta$                   | 0,064645  |
| $b_0$                      | 0,000000  |
| $b_1$                      | 75,238683 |
| $q$                        | 0,192436  |
| $\mu$                      | 0,000804  |
| $\gamma^2$                 | 0,000017  |
|                            |           |
| <b>Log-Verossimilhança</b> | 31847,59  |

Já os resultados da otimização com restrições nos parâmetros são:

Tabela 11

| Parâmetro                               | Valor      | Desv. Pad. | Estat. t  | Sig.     |
|-----------------------------------------|------------|------------|-----------|----------|
| $k$                                     | 0,005113   | 0,004815   | 1,061980  | 0,288246 |
| $\theta$                                | 0,019268   | 0,129831   | 0,148410  | 0,882021 |
| $b_0$                                   | 0,000000   | 0,000000   | 34,82040  | 0,000000 |
| $b_1$                                   | 210,722922 | 10,446016  | 20,17256  | 0,000000 |
| $q$                                     | 0,074213   | 0,005882   | 12,61591  | 0,000000 |
| $\mu$                                   | -0,000548  | 0,000427   | -1,282740 | 0,199584 |
| $\gamma^2$                              | 0,000030   | 0,000001   | 20,13248  | 0,000000 |
|                                         |            |            |           |          |
| <b>Log-Verossimilhança</b>              |            | 19299,52   |           |          |
| <b>Critério de Informação de Akaike</b> |            | -15,40     |           |          |
| <b>Critério de Schwarz</b>              |            | -15,38     |           |          |

Como podem ser observados nas tabelas, os critérios de informação adotados indicam vantagem ao emprego dos processos com saltos, tanto puro quanto com volatilidade estocástica. Outra questão relevante é que a introdução dos saltos gera uma redução no parâmetro de variância ( $\nu^2$ ), o que indica que os saltos representam boa parte da volatilidade na taxa Selic.

No processo de Ornstein-Uhlenbeck com saltos, o parâmetro que representa a probabilidade de ocorrência de um salto ( $q$ ) é estimado em 16,3%. Isso indica que, em média, ocorre um salto a cada seis dias (aproximadamente um salto por semana). Na estimação do processo de Ornstein-Uhlenbeck com saltos e variância ARCH(1) esse parâmetro é reduzido, indicando que parte da variação da taxa é absorvida pela volatilidade estocástica em detrimento dos saltos<sup>22</sup>.

No trabalho de Das (2002), para uma amostra da taxa *fed funds* entre 1988 e 1997 modelada segundo o processo de Ornstein-Uhlenbeck com saltos, o resultado obtido para a probabilidade de ocorrência de um salto ( $q$ ) é maior (21,6%). Isso indica, portanto, que a taxa Selic no período estudado apresentou menor probabilidade de ocorrência saltos. Tal fato é consistente com a política monetária e o sistema de metas de inflação vigentes na última década no Brasil.

Outro ponto interessante são os resultados encontrados para o parâmetro da tendência de longo prazo para a taxa de juros ( $\theta$ ). No artigo supracitado, o

<sup>22</sup> Tais resultados encontram-se bastante alinhados com os obtidos por Das (2002) e Johannes (2004).

valor do parâmetro foi consistentemente mais baixo do que no caso brasileiro. Isso mostra que apesar de apresentar movimentos menos bruscos na taxa curta, a taxa básica brasileira, em conformidade com o risco soberano dos países em questão, tende a operar com valores mais elevados.