

Michelle Ferreira Miranda

*Estimação dos coeficientes de
um processo de difusão*

Belo Horizonte – MG

Março / 2007

Michelle Ferreira Miranda
michellemiranda_est@hotmail.com

*Estimação dos coeficientes
de um processo de difusão*

Dissertação apresentada ao Programa de
Pós-Graduação do Departamento de Estatística da
Universidade Federal de Minas Gerais.

Orientador: Professor Gregorio Saravia Atuncar

MESTRADO EM ESTATÍSTICA
DEPARTAMENTO DE ESTATÍSTICA
INSTITUTO DE CIÊNCIAS EXATAS
UNIVERSIDADE FEDERAL MINAS GERAIS

Belo Horizonte – MG

Março / 2007

Agradecimentos

Gostaria de agradecer às muitas pessoas que contribuíram para a realização deste trabalho, das quais muitas vezes tive que me privar e não pude lhes dar a atenção merecida.

Em primeiro lugar, agradeço ao professor, orientador e amigo Gregorio Saravia Atuncar por sua sábia orientação.

Agradeço à Deus e aos meus pais Artêmio Miranda e Luzia Madalena Ferreira Miranda e à minha irmã Izabela Ferreira Miranda pela estrutura familiar, pelo amor e pela força. E que eu possa retribuir pelo menos um pouco do orgulho que sinto de vocês.

Agradeço também a todos os meus familiares que sempre me apoiaram.

Agradeço aos professores Adrian Pablo Hinojosa Luna, Aniura Milanés Barrietos e Chang Chung Yu Dorea por terem acompanhado o trabalho e pelas críticas e sugestões de grande contribuição.

Agradeço ao colega Elias Krainski pela dicas sobre o R e por sua boa vontade. Também não posso deixar de agradecer os amigos que fiz nestes anos, pelo suporte emocional e pela união. Especialmente, agradeço aos bons e velhos amigos, pela paciência e compreensão da minha ausência.

“A auto-satisfação é inimiga do estudo. Se queremos realmente aprender alguma coisa, devemos começar por libertar-nos disso. Em relação a nós próprios devemos ser insaciáveis na aprendizagem e em relação aos outros, insaciáveis no ensino ”
(Mao Tse-Tung)

Resumo

Os processos de difusão podem ser usados para a modelagem estocástica com aplicações em física, ciências biológicas, médicas e mais recentemente na economia. Entretanto, todos os modelos envolvem parâmetros desconhecidos ou funções desconhecidas que precisamos estimar utilizando as observações do processo. Em nosso trabalho estudamos uma forma de estimar as funções desconhecidas do modelo utilizando o método não-paramétrico de regressão polinomial local, de acordo com a proposta de Fan e Gijbels (1995). Além disso, comparamos o método não-paramétrico com o paramétrico proposto por Cleur e Manfredi (1999) que utiliza a função de máxima verossimilhança. Os resultados obtidos fazem com que acreditemos que o método não-paramétrico funciona tão bem quanto o paramétrico para estimar as funções desconhecidas de um modelo de difusão. Para finalizar, aplicamos o método a três conjuntos de dados e estimamos as funções do modelo.

Sumário

| | |
|---|--------|
| Lista de Figuras | p. vii |
| Lista de Tabelas | p. x |
| 1 Introdução e Revisão da Literatura | p. 1 |
| 2 Definições e Conceitos Básicos | p. 6 |
| 2.1 Definições | p. 6 |
| 2.2 Construção da Integral de Itô | p. 8 |
| 2.3 Processos de Difusão | p. 10 |
| 2.4 Esquemas de Discretização | p. 12 |
| 2.4.1 O Esquema de Euler | p. 12 |
| 2.4.2 O Esquema de Milstein | p. 12 |
| 2.5 Núcleo-Estimador | p. 13 |
| 2.5.1 Critérios de Avaliação das Estimativas | p. 13 |
| 2.5.2 A Função Núcleo | p. 14 |
| 3 Aproximações para as funções $\mu(\cdot)$ e $b^2(\cdot)$ | p. 15 |
| 3.1 Aproximações de Primeira Ordem | p. 16 |
| 3.2 Aproximações de Ordem Superior | p. 17 |
| 4 Estimação Não-Paramétrica dos coeficientes | p. 20 |
| 4.1 Estimadores da Função de Regressão | p. 20 |
| 4.1.1 O Estimador de Nadaraya-Watson | p. 20 |

| | | |
|----------|---|--------------|
| 4.1.2 | Regressão Polinomial Local | p. 22 |
| 4.2 | Estimação do Parâmetro de Suavização | p. 24 |
| 4.2.1 | Método de Dois Estágios para Regressão Polinomial Local | p. 25 |
| 5 | Validação do Método de Escolha da Janela | p. 30 |
| 5.1 | Caso I - $m(x) = \sin(2x) + 2e^{-16x^2}$ | p. 31 |
| 5.2 | Caso II - $m(x) = x + 2e^{-16x^2}$ | p. 33 |
| 5.3 | Caso III - $m(x) = \sin(2\pi x)$ | p. 35 |
| 5.4 | Caso IV - $m(x) = 0.4x + 1$ | p. 37 |
| 5.5 | Comparação para Diferentes Tamanhos Amostrais-Caso III | p. 39 |
| 5.6 | Conclusões do Capítulo | p. 40 |
| 6 | Simulações | p. 41 |
| 6.1 | Estimação Paramétrica do Coeficiente de Tendência | p. 41 |
| 6.2 | Legenda das Tabelas e Gráficos da Seção | p. 44 |
| 6.3 | Modelo Cox-Ingersoll-Ross (CIR) | p. 46 |
| 6.3.1 | $\kappa = 0,3$ e $T = 30$ | p. 47 |
| 6.3.2 | $\kappa = 0,8$ e $T = 30$ | p. 48 |
| 6.3.3 | $\kappa = 0,3$ e $T = 48$ | p. 49 |
| 6.3.4 | $\kappa = 0,8$ e $T = 48$ | p. 50 |
| 6.3.5 | $\kappa = 0,3$ e $T = 60$ | p. 51 |
| 6.3.6 | $\kappa = 0,8$ e $T = 60$ | p. 52 |
| 6.3.7 | Comentários | p. 53 |
| 6.3.8 | Tabelas dos Erros-Modelo CIR | p. 54 |
| 6.3.9 | Uma Comparação entre o Esquema de Euler e o de Milstein | p. 57 |
| 6.4 | Modelo de Brennan-Schwartz | p. 58 |
| 6.4.1 | $\kappa = 0,3$ e $T = 30$ | p. 59 |

| | | |
|----------|--|-------|
| 6.4.2 | $\kappa = 0,8$ e $T = 30$ | p. 60 |
| 6.4.3 | $\kappa = 0,3$ e $T = 48$ | p. 61 |
| 6.4.4 | $\kappa = 0,8$ e $T = 48$ | p. 62 |
| 6.4.5 | $\kappa = 0,3$ e $T = 60$ | p. 63 |
| 6.4.6 | $\kappa = 0,8$ e $T = 60$ | p. 64 |
| 6.4.7 | Comentários | p. 65 |
| 6.4.8 | Tabelas dos Erros- Modelo Brennan-Schwartz | p. 66 |
| 6.5 | Modelo Heuristically Derived Logistic | p. 69 |
| 6.5.1 | $\kappa = 0,3$ e $T = 30$ | p. 70 |
| 6.5.2 | $\kappa = 0,8$ e $T = 30$ | p. 71 |
| 6.5.3 | $\kappa = 0,3$ e $T = 48$ | p. 72 |
| 6.5.4 | $\kappa = 0,8$ e $T = 48$ | p. 73 |
| 6.5.5 | $\kappa = 0,3$ e $T = 60$ | p. 74 |
| 6.5.6 | $\kappa = 0,8$ e $T = 60$ | p. 75 |
| 6.5.7 | Comentários | p. 76 |
| 6.5.8 | Tabelas dos Erros- Modelo HD Logistic | p. 77 |
| 7 | Aplicações | p. 80 |
| 7.1 | Depósito Interbancário | p. 80 |
| 7.1.1 | Aproximação de Primeira Ordem | p. 82 |
| 7.1.2 | Aproximação de Segunda Ordem | p. 85 |
| 7.1.3 | Aproximação de Terceira Ordem | p. 88 |
| 7.2 | C-Bond spread | p. 91 |
| 7.3 | Overnight/Selic | p. 94 |
| 8 | Considerações Finais | p. 97 |
| | Referências Bibliográficas | p. 98 |

Lista de Figuras

| | | |
|------|---|-------|
| 5.1 | Exemplo Caso I. | p. 31 |
| 5.2 | Janelas e ISE Caso I. | p. 32 |
| 5.3 | Exemplo Caso II. | p. 33 |
| 5.4 | Janelas e ISE Caso II. | p. 34 |
| 5.5 | Exemplo Caso III. | p. 35 |
| 5.6 | Janelas e ISE Caso III. | p. 36 |
| 5.7 | Exemplo Caso IV. | p. 37 |
| 5.8 | Janelas e ISE Caso IV. | p. 38 |
| 5.9 | Comparação ISE para Diferentes Tamanhos Amostrais. | p. 39 |
| 6.1 | Exemplo CIR $\kappa = 0,3$ e $T = 30$ | p. 47 |
| 6.2 | Exemplo CIR $\kappa = 0,8$ e $T = 30$ | p. 48 |
| 6.3 | Exemplo CIR $\kappa = 0,3$ e $T = 48$ | p. 49 |
| 6.4 | Exemplo CIR $\kappa = 0,8$ e $T = 48$ | p. 50 |
| 6.5 | Exemplo CIR $\kappa = 0,3$ e $T = 60$ | p. 51 |
| 6.6 | Exemplo CIR $\kappa = 0,8$ e $T = 60$ | p. 52 |
| 6.7 | Exemplo CIR Utilizando os Dois Esquemas de Discretização. | p. 57 |
| 6.8 | Exemplo Brennan-Schwartz $\kappa = 0,3$ e $T = 30$ | p. 59 |
| 6.9 | Exemplo Brennan-Schwartz $\kappa = 0,8$ e $T = 30$ | p. 60 |
| 6.10 | Exemplo Brennan-Schwartz $\kappa = 0,3$ e $T = 48$ | p. 61 |
| 6.11 | Exemplo Brennan-Schwartz $\kappa = 0,8$ e $T = 48$ | p. 62 |
| 6.12 | Exemplo Brennan-Schwartz $\kappa = 0,3$ e $T = 60$ | p. 63 |
| 6.13 | Exemplo Brennan-Schwartz $\kappa = 0,8$ e $T = 60$ | p. 64 |

| | | |
|------|--|-------|
| 6.14 | Exemplo HD Logistic $\kappa = 0,3$ e $T = 30$ | p. 70 |
| 6.15 | Exemplo HD Logistic $\kappa = 0,8$ e $T = 30$ | p. 71 |
| 6.16 | Exemplo HD Logistic $\kappa = 0,3$ e $T = 48$ | p. 72 |
| 6.17 | Exemplo HD Logistic $\kappa = 0,8$ e $T = 48$ | p. 73 |
| 6.18 | Exemplo HD Logistic $\kappa = 0,3$ e $T = 60$ | p. 74 |
| 6.19 | Exemplo HD Logistic $\kappa = 0,8$ e $T = 60$ | p. 75 |
| 7.1 | Série de Depósito Interbancário-DI. | p. 80 |
| 7.2 | Gráfico de Dispersão e Estimativa do Coef. de Tendência-DI. | p. 82 |
| 7.3 | Estimativa do Coef. de Tendência-DI. | p. 83 |
| 7.4 | Estimativa do Coef. de Difusão-DI. | p. 83 |
| 7.5 | Séries Reconstruídas Usando as Estimativa do Coef. de Tendência. . . . | p. 84 |
| 7.6 | Gráfico de Dispersão e Estimativa do Coef. de Tendência-DI- Ordem 2. | p. 85 |
| 7.7 | Estimativa do Coef. de Tendência-DI Ordem2. | p. 86 |
| 7.8 | Estimativa do Coef. de Difusão-DI Ordem2. | p. 86 |
| 7.9 | Séries Reconstruídas Usando as Estimativa do Coef. de Tendência - Ordem2. | p. 87 |
| 7.10 | Gráfico de Dispersão e Estimativa do Coef. de Tendência-DI- Ordem 3. | p. 88 |
| 7.11 | Estimativa do Coef. de Tendência-DI Ordem3. | p. 89 |
| 7.12 | Estimativa do Coef. de Difusão-DI Ordem3. | p. 89 |
| 7.13 | Séries Reconstruídas Usando as Estimativa do Coef. de Tendência - Ordem3. | p. 90 |
| 7.14 | Série C-Bond spread. | p. 91 |
| 7.15 | Gráfico de Dispersão e Estimativa do Coef. de Tendência-CBond. | p. 92 |
| 7.16 | Estimativa do Coef. de Tendência-CBond. | p. 92 |
| 7.17 | Estimativa do Coef. de Difusão-CBond. | p. 93 |
| 7.18 | Séries Reconstruídas Usando as Estimativa do Coef. de Tendência. | p. 93 |
| 7.19 | Série Over Selic. | p. 94 |

| | | |
|------|---|-------|
| 7.20 | Gráfico de Dispersão e Estimativa do Coef. de Tendência-Over/Selic. . . | p. 95 |
| 7.21 | Estimativa do Coef. de Tendência-OverSelic. | p. 95 |
| 7.22 | Estimativa do Coef. de Difusão-OverSelic. | p. 96 |
| 7.23 | Séries Reconstruídas Usando as Estimativa do Coef. de Tendência. . . . | p. 96 |

Lista de Tabelas

| | | |
|------|---|-------|
| 2.1 | Exemplos de Funções Núcleo. | p. 14 |
| 5.1 | Casos considerados para validação do método de estimação da janela ótima. | p. 30 |
| 5.2 | Caso I-Janela e ISE nos 2 Estágios. | p. 32 |
| 5.3 | Caso II-Janela e ISE nos 2 Estágios. | p. 34 |
| 5.4 | Caso III-Janela e ISE nos 2 Estágios. | p. 36 |
| 5.5 | Caso IV-Janela e ISE nos 2 Estágios. | p. 37 |
| 6.1 | Integrais envolvidas no CTMLE para os 3 modelos. | p. 42 |
| 6.2 | Integrais não-estocásticas utilizando a regra do Trapézio. | p. 43 |
| 6.3 | Integrais estocásticas utilizando a fórmula de Itô. | p. 43 |
| 6.4 | Erros $\kappa = 0,3$ e $T = 30$ - Modelo CIR. | p. 54 |
| 6.5 | Erros $\kappa = 0,8$ e $T = 30$ - Modelo CIR. | p. 54 |
| 6.6 | Erros $\kappa = 0,3$ e $T = 48$ - Modelo CIR. | p. 55 |
| 6.7 | Erros $\kappa = 0,8$ e $T = 48$ - Modelo CIR. | p. 55 |
| 6.8 | Erros $\kappa = 0,3$ e $T = 60$ - Modelo CIR. | p. 55 |
| 6.9 | Erros $\kappa = 0,8$ e $T = 60$ - Modelo CIR. | p. 56 |
| 6.10 | Erros $\kappa = 0,3$ e $T = 30$ - Modelo Brennan-Schwartz. | p. 66 |
| 6.11 | Erros $\kappa = 0,8$ e $T = 30$ - Modelo Brennan-Schwartz. | p. 66 |
| 6.12 | Erros $\kappa = 0,3$ e $T = 48$ - Modelo Brennan-Schwartz. | p. 67 |
| 6.13 | Erros $\kappa = 0,8$ e $T = 48$ - Modelo Brennan-Schwartz. | p. 67 |
| 6.14 | Erros $\kappa = 0,3$ e $T = 60$ - Modelo Brennan-Schwartz. | p. 67 |
| 6.15 | Erros $\kappa = 0,8$ e $T = 60$ - Modelo Brennan-Schwartz. | p. 68 |
| 6.16 | Erros $\kappa = 0,3$ e $T = 30$ - HD Logistic. | p. 77 |

| | |
|---|-------|
| 6.17 Erros $\kappa = 0,8$ e $T = 30$ - HD Logistic. | p. 77 |
| 6.18 Erros $\kappa = 0,3$ e $T = 48$ - HD Logistic. | p. 78 |
| 6.19 Erros $\kappa = 0,8$ e $T = 48$ - HD Logistic. | p. 78 |
| 6.20 Erros $\kappa = 0,3$ e $T = 60$ - HD Logistic. | p. 78 |
| 6.21 Erros $\kappa = 0,8$ e $T = 60$ - HD Logistic. | p. 79 |
| 7.1 Valores de $V_1(k)$ e $V_2(k)$ | p. 81 |

1 *Introdução e Revisão da Literatura*

Considere o problema de estimar o coeficiente de tendência $\mu(\cdot)$, e o coeficiente de difusão, $b(\cdot)$, de um processo de difusão em tempo contínuo $\{X_t, a \leq t \leq T\}$ que é observado em tempos discretos e é definido pela seguinte equação diferencial estocástica:

$$dX_t = \mu(X_t)dt + b(X_t)dB_t \quad (1.1)$$

onde $\{B_t, a \leq t \leq T\}$ é um movimento Browniano padrão unidimensional.

Este modelo de difusão em tempo contínuo tem sido muito usado para descrever a dinâmica estocástica de muitos modelos econômicos, como por exemplo preços de ações e retornos de investimentos.

A estimação dos coeficientes μ e b segue duas linhas principais, a estimação paramétrica e a não-paramétrica. Rao (1999) descreve métodos paramétricos para estimação de μ e b . Dentre eles, o método de mínimos quadrados e o de máxima verossimilhança. O método dos momentos generalizados é proposto por Hansen (1982).

Os métodos paramétricos são utilizados quando as formas funcionais dos coeficientes são conhecidas e visam basicamente estimar os parâmetros envolvidos nas formas funcionais. Porém, frequentemente não conhecemos as formas dos coeficientes e neste caso especificar uma forma para μ e b pode levar a conclusões errôneas quando usadas com o propósito de inferência.

Apesar da teoria nos fornecer informações sobre o que fazer quando temos o modelo para a variável desejada, há muito pouca informação de como escolher o modelo correto. Dessa maneira, para evitar as muitas suposições para o modelo e reduzir possíveis vícios de modelagem, técnicas de regressão não-paramétrica têm sido estudadas nessa área.

Pham (1981) e Rao (1985) propuseram estimadores não-paramétricos para o coeficiente de tendência. Arfi (1995) mostrou que o estimador de Nadaraya-Watson para o coeficiente de tendência é fortemente e uniformemente consistente sob condições ergódicas e Arfi (1998) chegou à mesma conclusão para o coeficiente de difusão.

Fan e Yao (1998) usaram regressão linear local para estimar $b^2(\cdot)$ e mostraram que a abordagem proposta é eficiente. Aït-Sahalia (1996a) propôs um procedimento semi-paramétrico para estimar o coeficiente de difusão, sob condições paramétricas do coeficiente de tendência. Jiang e Knight (1997) desenvolveram o procedimento não-paramétrico usando núcleo estimador para o coeficiente de difusão e então encontraram um estimador não-paramétrico consistente para o coeficiente de tendência.

Stanton (1997), através do gerador infinitesimal e expansão em série de Taylor, construiu aproximações de primeira, segunda e terceira ordem para $\mu(\cdot)$ e $b(\cdot)$. Estas fórmulas contêm esperanças condicionais estimadas pelo método de regressão de Nadaraya-Watson. Na abordagem feita por Stanton (1997), a estimação do coeficiente de difusão $b(\cdot)$ é feita de forma separada, sem conhecer a estimativa de $\mu(\cdot)$ a priori. Esta característica torna o método simples e atrativo.

Fan e Zhang (2003) generalizaram a abordagem de Stanton e obtiveram as aproximações para todas as ordens, inclusive superiores a terceira. Um dos objetivos de Fan e Zhang (2003) foi examinar o impacto das aproximações de ordens superiores na estimativa dos coeficientes. Neste trabalho eles mostram que os estimadores de ordem superior reduzem o erro de aproximação no vício assintótico das estimativas, porém as variâncias assintóticas crescem exponencialmente com a ordem das aproximações.

O objetivo principal do nosso trabalho é utilizar a abordagem feita por Stanton (1997) para construir as esperanças condicionais para as aproximações de primeira e outras ordens, seguindo a generalização proposta por Fan e Zhang (2003) e então utilizar o procedimento de regressão não paramétrica para estimar as funções $\mu(\cdot)$ e $b(\cdot)$.

O modelo não paramétrico remove as restrições paramétricas para as funções que desejamos estimar e permite que estruturas de várias formas sejam utilizadas. Essa estrutura é determinada somente pelos dados observados, porém depende de um parâmetro, conhecido como janela (bandwidth), que deve ser estimado.

É conhecido da literatura sobre núcleo estimador que a escolha da função núcleo (que definiremos no capítulo 2) não é muito relevante para estimar a função desejada. Porém, a escolha da janela é de crucial importância para o desempenho dos métodos que utilizam núcleo estimadores. Por esse motivo, os métodos automáticos para a seleção da janela têm sido assunto para muitas pesquisas.

No contexto de regressão não-paramétrica, a estimativa da janela ótima consiste basicamente em minimizar o erro de predição. Esse procedimento apresenta variações de acordo com a decisão de qual é a medida de erro que devemos utilizar. Härdle e Marron (1986) mostraram que a escolha do critério dos resíduos quadráticos não é importante do ponto de vista assintótico. De acordo com Simonoff (1996), a seleção da janela na regressão que utiliza núcleo estimador segue três categorias principais.

A primeira categoria é conhecida como "leave-one-out", ou seja, deixar um de fora, que é a forma de estimação da validação cruzada que foi desenvolvida por Stone (1974) e foi primeiramente aplicada por Clark (1975), no contexto de regressão.

A segunda abordagem é basear o critério de ajuste na ressubstituição da estimativa do erro de predição, $p(h) = 1/n \sum_{i=1}^n (y_i - \hat{m}(x_i))^2$, modificado para ser um estimador não-viciado do $ASE(h)$ -Averaged Squared Error. O vício assintótico de $p(h)$ é

$$-\frac{2}{n^2 h} \sum_{i=1}^n \frac{K\left(\frac{x-x_i}{h}\right) \sigma^2(x_i)}{\hat{f}_X(x_i)}.$$

Seja $z_i = (nh)^{-1} K(0) / \hat{f}_X(x_i)$. Essa abordagem escolhe uma função de pesos tal que

$$G(h) = \frac{1}{n} \sum_{i=1}^n [y_i - \hat{m}(x_i)]^2 \Xi(z_i)$$

seja um estimador não-viciado do ASE.

Propostas que consideram esta abordagem incluem o método de validação cruzada generalizado proposto por Craven e Wahba (1979), o método de seleção de Shibata proposto por Shibata (1981), o AIC por Akaike (1970), o método de erro de predição finito proposto por Akaike (1974) e o método T de Rice proposto por Rice (1984).

A terceira e mais comum abordagem para a seleção da janela permite uma melhora significativa ao mudar o alvo de \hat{h}_0 pela janela teórica h_0 . Os métodos desta categoria se baseiam em estimar as quantidades desconhecidas da expressão da janela teórica. Maiores detalhes podem ser vistos em Simonoff (1996).

Dentro dessa abordagem, Chiu (1991) propôs uma forma de estimar a janela para o estimador de Priestley-Chao e Gasser *et al.* (1991) propuseram o método do *plug-in* para o estimador de Gasser-Müller. Cada um destes estimadores pode atingir uma taxa de convergência de $O(n^{-1/2})$, assumindo que a função de regressão m tenha derivadas suficientes. Härdle *et al.* (1992) propuseram uma janela para o método de Nadaraya-Watson, que utiliza dois núcleos para a suavização dos dados, que eles chamaram de Dupla Suavização. Essa abordagem é relativa ao método de validação cruzada para suavização da densidade proposto por Hall *et al.* (1992) e também pode atingir uma taxa de convergência de $O(n^{-1/2})$ para a janela ótima.

Fan e Gijbels (1995) propuseram um estimador do tipo *plug-in* para uma variância que usa uma versão ponderada de $p(h)$. Eles também encontraram um critério para a estimação da janela. Este critério, abordaremos no capítulo 4.

No Capítulo 2 apresentamos algumas definições necessárias para o entendimento do problema. No Capítulo 3 construímos as aproximações para os coeficientes μ e b da equação (1.1). No Capítulo 4, apresentamos a técnica não-paramétrica utilizada para estimar os coeficientes a partir das aproximações construídas no Capítulo 3.

A parte de simulação foi dividida em duas etapas. A primeira, no capítulo 5, é referente ao método de regressão não-paramétrica. Consideramos diversas situações que expressam diferentes formas para a função que desejamos estimar. O principal objetivo desta etapa de simulação é avaliar o comportamento do método de escolha da janela ótima para regressão polinomial local.

No capítulo 6 encontra-se a segunda etapa do processo de simulação, que está relacionada com a escolha de formas para as funções $\mu(X_t)$ e $b(X_t)$. Nesta etapa partimos do pressuposto que temos um modelo conhecido para gerar as observações através de um dos esquemas de discretização mencionados no capítulo 2. Após a geração dos dados, estimamos as funções $\mu(X_t)$ e $b(X_t)$ e comparamos com os verdadeiros valores, que foram estabelecidos em cada situação.

No Capítulo 7 o objetivo principal foi utilizar a metodologia para dados reais. Procuramos utilizar séries de natureza variada para englobar situações que encontramos no mundo real.

No Capítulo 8 se encontram as conclusões finais e propostas para futuros trabalhos.

2 Definições e Conceitos Básicos

Neste capítulo apresentamos alguns conceitos e definições necessárias para construir o principal objeto de estudo neste trabalho, o processo de difusão. Além disso, definimos os critérios para avaliar as estimativas no estudo de simulação e falamos um pouco sobre a função núcleo que utilizamos no trabalho. As duas primeiras seções do capítulo não são necessárias para entender a metodologia que será apresentada nos capítulos seguintes, mas são de grande relevância para a construção do processo de difusão na seção 2.3. As definições foram adaptadas de Oksendal (2003).

2.1 Definições

Definição 2.1.1 *Um processo estocástico é uma coleção de variáveis aleatórias*

$$\{X_t, t \in T\},$$

definida em um espaço de probabilidade (Ω, \mathcal{F}, P) , que assumem valores em \mathfrak{R} .

Definição 2.1.2 (Movimento Browniano) *Chamaremos de Movimento Browniano Padrão um processo estocástico em tempo contínuo $B = \{B_t, t \geq 0\}$ que satisfaz as seguintes propriedades:*

1. $P(u : B_0(u) = 0) = 1$, ou seja o movimento Browniano começa em 0.
2. B é um processo Gaussiano, isto é, para todo $0 \leq t_1, \dots, \leq t_k$, a variável aleatória $Z = (B_{t_1}, \dots, B_{t_k}) \in \mathfrak{R}^k$ tem distribuição normal multivariada. Isto significa que existe um vetor $M \in \mathfrak{R}^k$ e uma matriz não-negativa definida $C = [c_{jm}] \in \mathfrak{R}^{k \times k}$ tais que

$$E \left[\exp \left(i \sum_{j=1}^k u_j Z_j \right) \right] = \exp \left(-\frac{1}{2} \sum_{j,m} u_j c_{jm} u_m + i \sum_j u_j M_j \right), \quad (2.1)$$

para todo $u = (u_1, \dots, u_{nk}) \in \mathfrak{R}^k$ e E denota a esperança com respeito a P .

Além disso, se (2.1) vale, então

$$M = E[Z] \quad (2.2)$$

e

$$c_{jm} = E[(Z_j - M_j)(Z_m - M_m)] \quad (2.3)$$

é a matriz de covariância de Z .

Como temos um Movimento Browniano Padrão,

$$M = E[Z] = (0, 0, \dots, 0) \in \mathfrak{R}^k. \quad (2.4)$$

Então

$$E[B_t] = 0 \quad \forall t \geq 0. \quad (2.5)$$

Considerando (2.3), temos

$$C = \begin{pmatrix} t_1 & t_1 & \dots & t_1 \\ t_1 & t_2 & \dots & t_2 \\ \vdots & \vdots & & \vdots \\ t_1 & t_2 & \dots & t_k \end{pmatrix}. \quad (2.6)$$

Dessa maneira,

$$E[(B_t - E(B_t))(B_s - E(B_s))] = \min(s, t). \quad (2.7)$$

Além disso,

$$E[(B_t - B_s)^2] = t - s \quad \text{se } t \geq s. \quad (2.8)$$

3. B tem incrementos independentes, isto é:

$$B_{t_1}, B_{t_2} - B_{t_1}, \dots, B_{t_k} - B_{t_{k-1}}$$

são independentes para todo $0 \leq t_1 < t_2 < \dots < t_k$.

4. As trajetórias são contínuas.

Definição 2.1.3 Consideremos B um Movimento Browniano. Definimos \mathcal{F}_t como a σ -álgebra gerada pelas variáveis aleatórias $\{B_s\}_{0 \leq s \leq t}$.

Definição 2.1.4 Seja $\{\mathcal{N}_t\}_{t \geq 0}$ uma família crescente de σ -álgebras de subconjuntos de Ω . Um processo $g(t, \omega) : [0, \infty) \times \Omega \rightarrow \mathfrak{R}^n$ é chamado \mathcal{N}_t -adaptado se para todo $t \geq 0$ a função

$$\omega \rightarrow g(t, \omega)$$

é \mathcal{N}_t -mensurável.

Definição 2.1.5 Seja $\mathcal{V} = \mathcal{V}(S, T)$ a classe de funções

$$f(t, \omega) : [0, \infty) \times \Omega \rightarrow \mathfrak{R}$$

tal que:

1. $(t, \omega) \rightarrow f(t, \omega)$ é $\mathcal{B} \times \mathcal{F}$ -mensurável, onde \mathcal{B} denota a σ -álgebra de Borel em $[0, \infty)$.
2. $f(t, \omega)$ é \mathcal{F}_t -adaptada.
3. $E \left[\int_S^T f(t, \omega)^2 dt \right] < \infty$.

Definição 2.1.6 Uma função $\phi \in \mathcal{V}$ é chamada elementar se ela tem a forma

$$\phi(t, \omega) = \sum_j e_j(\omega) \mathcal{X}_{[t_j, t_{j+1})}(t), \quad (2.9)$$

onde \mathcal{X} é função indicadora. Note que como $\phi \in \mathcal{V}$ cada função e_j é \mathcal{F}_t -mensurável.

2.2 Construção da Integral de Itô

Consideremos o seguinte modelo

$$dX_t = \mu(t, X_t)dt + b(t, X_t)dB_t. \quad (2.10)$$

Uma versão discretizada deste modelo é

$$X_{k+1} - X_k = \mu(t_k, X_k)\Delta t_k + b(t_k, X_k)(B_{k+1} - B_k), \quad (2.11)$$

onde

$$X_k = X(t_k), B_k = B_{t_k}, \Delta t_k = t_{k+1} - t_k.$$

Dessa forma, obtemos de (2.11)

$$X_k = X_0 + \sum_{j=0}^{k-1} \mu(t_j, X_j) \Delta t_j + \sum_{j=0}^{k-1} b(t_j, X_j) \Delta B_j. \quad (2.12)$$

Se for possível provar que o limite do lado direito da expressão (2.12) existe em algum sentido quando $\Delta t_k \rightarrow 0$, então aplicando a notação usual de integral temos

$$X_t = X_0 + \int_0^t \mu(s, X_s) ds + \text{“} \int_0^t b(s, X_s) dB_s \text{”}. \quad (2.13)$$

Seja $f : [0, \infty) \times \Omega \rightarrow \mathfrak{R}$. Queremos definir $\int_0^t f(s, \omega) dB_s(\omega)$. As variações dos incrementos de B são muito grandes para que possamos definir a integral como a soma de Rieman. Em particular, a variação total dos incrementos é infinita quase certamente (ver em Oksendal (2003)).

Para funções $f \in \mathcal{V}$ iremos agora mostrar como definimos a integral de Itô

$$\mathcal{I}[f](\omega) = \int_S^T f(t, \omega) dB_t(\omega). \quad (2.14)$$

A idéia é a seguinte. Primeiramente definimos $\mathcal{I}[\phi]$ para uma classe de funções elementares ϕ tal que

$$\phi(t, \omega) = \sum_j e_j(\omega) \mathcal{X}_{[j2^{-n}, (j+1)2^{-n})}(t). \quad (2.15)$$

Como $\phi \in \mathcal{V}$, as funções e_j têm que ser escolhidas de modo que sejam \mathcal{F}_t -mensuráveis. Para tais função é razoável definir

$$\int_S^T \phi(t, \omega) dB_t(\omega) = \sum_j e_j(\omega) [B_{t_{j+1}} - B_{t_j}](\omega).$$

Depois, cada $f \in \mathcal{F}$ pode ser aproximada pelas ϕ_n 's e usamos isso para definir $\int f dB$ como o limite de $\int \phi_n dB$. Os detalhes para esta construção são mostrados em Oksendal (2003).

Definição 2.2.1 (A Integral de Itô) *Seja $f \in \mathcal{V}(S, T)$. Então, a integral de Itô de f (de S a T) é definida por:*

$$\int_S^T f(t, \omega) dB_t(\omega) = \lim_{n \rightarrow \infty} \int_S^T \phi_n(t, \omega) dB_t(\omega), \quad (2.16)$$

onde $\{\phi_n\}$ é uma sequência de funções elementares tais que

$$E \left\{ \int_S^T (f(t, \omega) - \phi_n(t, \omega))^2 dt \right\} \rightarrow 0 \text{ quando } n \rightarrow \infty.$$

A integral de Itô $\int f dB$ pode ser definida para uma classe de integrandos f maior do que a classe \mathcal{V} . Primeiramente, a condição de mensurabilidade 2 da definição 2.1.5 pode ser relaxada pela seguinte

2'. *Existe uma família crescente de σ -álgebras \mathcal{H}_t ; $t \geq 0$ tal que:*

a) *B_t é um martingal com respeito a \mathcal{H}_t .*

b) *f_t é \mathcal{H}_t -adaptada.*

O próximo passo é enfraquecer a condição 3 da definição 2.1.5 para

$$3'. P \left[\int_S^T f(s, \omega)^2 < \infty \right] = 1.$$

Definição 2.2.2 $\mathcal{W}_{\mathcal{H}}(S, T)$ *denota a classe de processos $f(t, \omega) \in \mathfrak{R}$ que satisfaz a condição 1 da definição 2.1.5 e as condições 2' e 3' mostradas anteriormente.*

2.3 Processos de Difusão

Os processos de difusão, também conhecidos como processos de Itô são definidos da seguinte forma.

Definição 2.3.1 *Seja B um movimento Browniano padrão unidimensional em (Ω, \mathcal{F}, P) . Um processo de Itô é um processo estocástico $\{X_t, 0 \leq t \leq T\}$ em (Ω, \mathcal{F}, P) na forma*

$$X_t = X_0 + \int_0^t \mu(s, X_s) ds + \int_0^t b(s, X_s) dB_s, \quad (2.17)$$

onde $b \in \mathcal{W}_{\mathcal{H}}$, tal que

$$P \left[\int_0^t b(s, \omega)^2 ds < \infty \quad \text{para todo } t \geq 0 \right] = 1.$$

Além disso,

$$P \left[\int_0^t |b(s, \omega)| ds < \infty \quad \text{para todo } t \geq 0 \right] = 1.$$

Se $\{X_t, 0 \leq t \leq T\}$ é um processo de Itô na forma da equação (2.19), pode ser escrito na forma diferencial

$$dX_t = \mu(t, X_t)dt + b(t, X_t)dB_t. \quad (2.18)$$

Em nosso trabalho consideramos somente o processo homogêneo no tempo, o que significa que os coeficientes μ e b são funções apenas do processo X_t e a equação (2.18) se reduz a

$$dX_t = \mu(X_t)dt + b(X_t)dB_t. \quad (2.19)$$

É importante observar que o processo em (2.18) e (2.19), por serem processos em tempo contínuo, nunca são observados na prática.

Existem alguns esquemas de discretização para os processos dados em (2.18) e (2.19). Esses esquemas de discretização são úteis para os casos de simulação, pois o método para estimação dos coeficientes depende dos valores observados do processo, que são obtidos em tempos discretos. A seguir, mostramos alguns destes esquemas.

2.4 Esquemas de Discretização

Para efeitos de simulação, definiremos os coeficientes da equação (2.19) e aplicaremos o método que será descrito nas próximas seções do trabalho para estimar esses coeficientes. O método para estimação dos coeficientes depende somente dos valores observados do processo, por isso a necessidade dos esquemas descritos a seguir.

2.4.1 O Esquema de Euler

O mais simples esquema de discretização de um processo de difusão é o esquema de Euler. Consideremos o processo de Itô $X = \{X_t, t_0 \leq t \leq T\}$ que satisfaz a equação diferencial estocástica dada em (2.19). Consideremos também que $t_0 \leq t \leq T$ e o valor inicial $X_{t_0} = X_0$. Para uma certa discretização, $t_0 = \tau_0 < \tau_1 < \dots < \tau_n = T$ no intervalo de tempo $[t_0, T]$, a aproximação de Euler de um processo estocástico em tempo contínuo, $X = \{X(t), t_0 \leq t \leq T\}$ satisfaz o seguinte esquema iterativo

$$X_{n+1} = X_n + \mu(X_n)(\tau_{n+1} - \tau_n) + b(X_n)(B_{\tau_{n+1}} - B_{\tau_n}), \quad (2.20)$$

para $n = 0, 1, 2, \dots, N - 1$ com valor inicial X_0 .

2.4.2 O Esquema de Milstein

No esquema proposto por Milstein, acrescentamos ao esquema de Euler o seguinte termo:

$$bb'I_{(1,1)} = \frac{1}{2}bb'\{(\Delta B)^2 - \Delta\} \quad (2.21)$$

Usando expansão em série de Taylor, nós obtemos o seguinte sistema iterativo

$$X_{n+1} = X_n + \mu^*(X_n)(\tau_{n+1} - \tau_n) + b(X_n)(B_{\tau_{n+1}} - B_{\tau_n}) + \frac{1}{2}bb'\{(\Delta B)^2 - \Delta\}, \quad (2.22)$$

onde

$$\mu^* = \mu - \frac{1}{2}bb'$$

$$\Delta = (T - t_0)/N.$$

Existem muitos outros esquemas de discretização. Detalhá-los foge do escopo do nosso trabalho. Maiores detalhes podem ser vistos em Kloeden e Platen (1992).

2.5 Núcleo-Estimador

A suavização utilizando núcleo-estimador refere-se a uma classe geral de técnicas para a estimação não-paramétrica de funções. Uma situação onde o núcleo-estimador é bastante utilizado é para a estimação de funções de densidade mas outra situação bem simples é na estimação não-paramétrica de uma função de regressão.

2.5.1 Critérios de Avaliação das Estimativas

Para analisar a performance do núcleo estimador é necessário primeiramente especificar qual o critério que deve ser utilizado para medir o erro. Existem critérios muito utilizados na literatura de núcleo estimador que definimos a seguir.

Uma medida local é o Erro Quadrático Médio (EQM), também conhecido como MSE (Mean Squared Error), que é definido da seguinte forma

$$EQM(f(\hat{x})) = E(\hat{f}(x) - f(x))^2. \quad (2.23)$$

Outras medidas de erro bastante utilizadas para medir o desempenho global das estimativas são o Erro Quadrático Integrado (EQI), conhecido como ISE (Integrated Squared Error) e o Erro Quadrático Médio Integrado (EQMI), também conhecido como MISE (Mean Integrated Squared Error). Estes erros são definidos como

$$EQI(f(\hat{x})) = \int [\hat{f}(x) - f(x)]^2 dx. \quad (2.24)$$

$$EQMI(f(\hat{x})) = E \left\{ \int [\hat{f}(x) - f(x)]^2 dx \right\}. \quad (2.25)$$

2.5.2 A Função Núcleo

Em nosso trabalho utilizamos a função núcleo K que satisfaz as seguintes condições:

1. $\int K(u)du = 1$
2. $\int uK(u)du = 0$
3. $\int u^2K(u) = \sigma_K^2 > 0$

Alguns exemplos de funções núcleo bastante utilizadas são mostrados na tabela 2.1:

Tabela 2.1: Exemplos de Funções Núcleo.

| Núcleo | | |
|--------------|-----------------------------|------------------------|
| Epanechnikov | $\frac{3}{4}(1 - u^2),$ | $ u \leq 1$ |
| Biponderado | $\frac{15}{16}(1 - u^2)^2,$ | $ u \leq 1$ |
| Gaussiano | $(2\pi)^{-1/2}e^{-u^2/2},$ | $-\infty < u < \infty$ |
| Uniforme | $\frac{1}{2},$ | $ u \leq 1$ |

3 Aproximações para as funções $\mu(\cdot)$ e $b^2(\cdot)$

Na abordagem paramétrica, a abordagem usual é primeiramente especificar formas paramétricas para as funções $\mu(\cdot)$ e $b(\cdot)$, dadas na equação (1.1). Dados μ e b , a densidade de transição do valor x no tempo t , ao valor y no tempo s , $p(s, y \setminus t, x)$, deve satisfazer a equação forward de Kolmogorov,

$$\frac{\partial p(s, y \setminus t, x)}{\partial s} = -\frac{\partial}{\partial y}(\mu(y)p(s, y \setminus t, x)) + \frac{1}{2} \frac{\partial^2}{\partial y^2}(b^2(y)p(s, y \setminus t, x)) \quad (3.1)$$

e a equação backward de Kolmogorov,

$$-\frac{\partial p(s, y \setminus t, x)}{\partial s} = \mu(x) \frac{\partial}{\partial x}(p(s, y \setminus t, x)) + \frac{1}{2} b^2(x) \frac{\partial^2}{\partial x^2}(p(s, y \setminus t, x)). \quad (3.2)$$

Maiores detalhes podem ser vistos em Oksendal (2003). A princípio, para uma dada parametrização de μ e b , podemos resolver a equação (3.1) para uma densidade condicional p como uma função dos parâmetros, então usar máxima verossimilhança para estimar os parâmetros do modelo. Essa abordagem foi usada por Pearson e Sum (1994) para estimar os parâmetros de interesse do modelo CIR (Cox, Ingersoll, Ross) que pode ser visto em Cox *et al.* (1985).

Exceto em alguns casos, a equação (3.1) só pode ser resolvida numericamente, o que faz a implementação da máxima verossimilhança inconveniente. O Método dos Momentos Generalizados (GMM) proposto por Hansen (1982) pode ser usado no lugar da máxima verossimilhança mesmo quando a função de verossimilhança é complicada e demorada para ser calculada computacionalmente ou também quando deseja-se saber somente certas propriedades da distribuição ao invés da função completa de verossimilhança.

Um problema potencialmente sério que ocorre com qualquer modelo paramétrico, quando não existe razão para preferir um modelo do que o outro, é a má especificação. Mesmo quando um modelo ajusta bem uma determinada variável de interesse, por exemplo taxa de juros, isso não necessariamente implica que ajustará bem, por exemplo, aos preços de ações.

Para evitar as restrições paramétricas na especificação de formas funcionais para μ e b , utilizaremos em nosso trabalho, a abordagem não-paramétrica que é descrita na próxima seção.

3.1 Aproximações de Primeira Ordem

Sob condições apropriadas de μ e b , dados pela equação 1.1, considere uma função $f : \mathfrak{R} \rightarrow \mathfrak{R}$ contínua, duas vezes diferenciável, com suporte compacto, e o limite de f existe para todo $x \in \mathfrak{R}$. Podemos expressar a esperança condicional $E_t\{f(X_{t+\Delta}, t + \Delta)\}$ na forma de uma expansão em série de Taylor

$$\begin{aligned} E_t\{f(X_{t+\Delta}, t + \Delta)\} &= f(X_t, t) + \mathcal{L}f(X_t, t)\Delta + \frac{1}{2}\mathcal{L}^2f(X_t, t)\Delta^2 \\ &+ \dots + \frac{1}{n!}\mathcal{L}^n f(X_t, t)\Delta^n + O(\Delta^{n+1}), \end{aligned} \quad (3.3)$$

com o incremento de tempo $\Delta \downarrow 0$. O símbolo E_t denota a esperança condicional dado X_t , e o gerador infinitesimal, \mathcal{L} , do processo $\{X_t\}$ é definido por

$$\begin{aligned} \mathcal{L}(f(x, t)) &= \lim_{\tau \rightarrow t} \frac{E\{f(X_\tau, \tau) | X_t = x\} - f(x, t)}{\tau - t} \\ &= \frac{\partial f(x, t)}{\partial t} + \frac{\partial f(x, t)}{\partial x} \mu(x) \\ &+ \frac{1}{2} \frac{\partial^2 f(x, t)}{\partial^2 x} b^2(x). \end{aligned} \quad (3.4)$$

A fórmula para a aproximação de primeira ordem para a função $\mathcal{L}(f(x, t))$ é dada por

$$\Delta^{-1} E_t\{f(X_{t+\Delta}, t + \Delta) - f(X_t, t)\} = \mathcal{L}f(X_t, t) + O(\Delta). \quad (3.5)$$

Se $f(x, t) = x$, $\mathcal{L}f(x, t) = \mu(x)$.

Se $f(x, t) = (x - X_t)^2$, $\mathcal{L}f(x, t) = 2(x - X_t)\mu(x) + b^2(x)$.

Essas duas funções especiais nos levam a $\mu(X_t)$ e $b^2(X_t)$. Nesse caso, estimar o lado esquerdo de (3.5) pelo método de Nadaraya Watson nos leva às estimativas de Stanton para $\mu(x)$ e $b^2(x)$, baseadas na aproximação de primeira ordem.

3.2 Aproximações de Ordem Superior

Além da aproximação de primeira ordem, Stanton (1997) também propôs aproximações de segunda e terceira ordem. Fan e Zhang (2003) propuseram uma generalização do que foi feito por Stanton (1997). De acordo com Fan e Zhang (2003) as aproximações de ordem superior (ou diferenças) podem ser obtidas através de uma combinação linear do lado esquerdo de (3.5). Mais precisamente, para todo inteiro fixo $k \geq 1$, qualquer sequência de constantes $\{a_{k,j}, j = 1, \dots, k\}$, e quaisquer intervalos de tempo $j\Delta$, para $j = 1, \dots, k$, consideremos a seguinte combinação linear:

$$\begin{aligned}
& \Delta^{-1} \sum_{j=1}^k a_{k,j} E_t \{f(X_{t+j\Delta}, t+j\Delta) - f(X_t, t)\} \\
&= \left\{ \sum_{j=1}^k j a_{k,j} \right\} \mathcal{L}f(X_t, t) + \left\{ \sum_{j=1}^k j^2 a_{k,j} \right\} \frac{\mathcal{L}^2 f(X_t, t)}{2} \Delta \\
&+ \dots + \left\{ \sum_{j=1}^k j^k a_{k,j} \right\} \frac{\mathcal{L}^k f(X_t, t)}{k!} \Delta^{k-1} \\
&+ \left\{ \sum_{j=1}^k j^{k+1} a_{k,j} \right\} \frac{\mathcal{L}^{k+1} f(X_t, t)}{(k+1)!} \Delta^k + O(\Delta^{k+1})
\end{aligned} \tag{3.6}$$

Então, uma aproximação de ordem k ,

$$\Delta^{-1} \sum_{j=1}^k a_{k,j} E_t \{f(X_{t+j\Delta}, t+j\Delta) - f(X_t, t)\} = \mathcal{L}f(X_t, t) + O(\Delta^k), \tag{3.7}$$

é obtida escolhendo os coeficientes $\{a_{k,j}\}_{j=1}^k$ que satisfazem o seguinte sistema de equações

$$\begin{aligned}
\sum_{j=1}^k j a_{k,j} &= 1 \\
\sum_{j=1}^k j^2 a_{k,j} &= 0 \\
&\vdots \\
\sum_{j=1}^k j^k a_{k,j} &= 0.
\end{aligned} \tag{3.8}$$

A forma geral das soluções, $\{a_{k,j}, j = 1, \dots, k\}$, é apresentada no Teorema 3.2.1, cuja demonstração se encontra em Fan e Zhang (2003).

Teorema 3.2.1 *Para todo inteiro fixo $k \geq 1$, a única solução para o sistema (3.8) é dada por*

$$a_{k,j} = (-1)^{j+1} \binom{k}{j} / j, \quad j = 1, \dots, k. \tag{3.9}$$

Além disso, com essa escolhas de $\{a_{k,j}\}_{j=1}^k$ nós temos

$$\sum_{j=1}^k j^k a_{k,j} = (-1)^{k+1} k!.$$

Portanto, usando as soluções únicas $(a_{k,1}, \dots, a_{k,k})$, obtemos para $\mathcal{L}f(X_t, t)$ uma fórmula geral para a aproximação de k-ésima ordem,

$$\Delta^{-1} \sum_{j=1}^k a_{k,j} E_t \{f(X_{t+j\Delta}, t+j\Delta) - f(X_t, t)\}, \tag{3.10}$$

com o termo do erro de aproximação

$$(-1)^{k+1} \frac{\mathcal{L}^{k+1} f(X_t, t)}{(k+1)} \Delta^k + O(\Delta^{k+1}) \tag{3.11}$$

As equações (3.10) e (3.11) implicam que

$$\Delta^{-1} \sum_{j=1}^k a_{k,j} E_t (X_{t+j\Delta} - X_t) = \mu(X_t) + \left\{ (-1)^{k+1} \frac{\mathcal{L}^{k+1} f(X_t, t)}{(k+1)} \Delta^k + O(\Delta^{k+1}) \right\}, \tag{3.12}$$

com a escolha de $f_1(x, t) = x$. E

$$\Delta^{-1} \sum_{j=1}^k a_{k,j} E_t (X_{t+j\Delta} - X_t)^2 = b^2(X_t) + \left\{ (-1)^{k+1} \frac{\mathcal{L}^{k+1} f(X_t, t)}{(k+1)} \Delta^k + O(\Delta^{k+1}) \right\}, \quad (3.13)$$

com a escolha de $f_2(x, t) = (x - X_t)^2$. De (3.13), podemos simplesmente tomar a raiz quadrada para obter a fórmula para a aproximação de k -ésima ordem para a função $b(X_t)$ tal que

$$b(X_t) = \left\{ \Delta^{-1} \sum_{j=1}^k a_{k,j} E_t (X_{t+j\Delta} - X_t)^2 \right\}^{1/2} + O(\Delta^k). \quad (3.14)$$

Com as fórmulas para as aproximações de μ e b disponíveis, precisamos ainda estimar a esperança condicional. Dado um tempo inicial t_0 e os dados de uma série temporal $(X_{t_0+i\Delta}, i = 1, \dots, n)$ observados em tempos igualmente espaçados, o primeiro passo é obter os $(n - k)$ pares de dados,

$$\left(X_{t_0+i\Delta}, \Delta^{-1} \sum_{j=1}^k a_{k,j} \{X_{t_0+(i+j)\Delta} - X_{t_0+i\Delta}\} \right) \equiv (X_{i\Delta}^*, Y_{i\Delta}^*), \quad (3.15)$$

para estimar $\mu(\cdot)$. E

$$\left(X_{t_0+i\Delta}, \Delta^{-1} \sum_{j=1}^k a_{k,j} \{X_{t_0+(i+j)\Delta} - X_{t_0+i\Delta}\}^2 \right) \equiv (X_{i\Delta}^*, Z_{i\Delta}^*), \quad (3.16)$$

para estimar $b^2(\cdot)$. Onde $i = 1, \dots, n - k$.

Nosso segundo passo é estimar, usando regressão não-paramétrica, as esperanças condicionais obtidas das equações (3.12) e (3.13), dadas por:

$$E(Y_{i\Delta}^* | X_{i\Delta}^* = x_0) = \mu(x_0) + O(\Delta^k), \quad (3.17)$$

$$E(Z_{i\Delta}^* | X_{i\Delta}^* = x_0) = b^2(x_0) + O(\Delta^k). \quad (3.18)$$

No Capítulo 4 apresentamos a forma de estimar as esperanças condicionais dadas em (3.17) e (3.18).

4 *Estimação Não-Paramétrica dos coeficientes*

Este capítulo está dividido em duas seções. A seção 4.1 define os estimadores para as esperanças condicionais dadas em (3.17) e (3.18). Estas definições foram adaptadas de Simonoff (1996). Na seção 4.2 apresentamos o método para escolha da janela ótima para a regressão polinomial local apresentada na seção 4.1.

4.1 Estimadores da Função de Regressão

4.1.1 O Estimador de Nadaraya-Watson

Considere o modelo de regressão linear simples:

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i, \quad i = 1, \dots, n. \quad (4.1)$$

Quando temos uma relação que não é linear, ajustar o modelo (4.1) pode levar a resultados ruins, implicando em conclusões errôneas. Uma maneira alternativa a (4.1) é utilizar o modelo de regressão não-paramétrica

$$Y_i = m(x_i) + \epsilon_i, \quad 1 \leq i \leq n. \quad (4.2)$$

A curva de regressão $m(x)$ é a esperança condicional $m(x) = E(Y|X = x)$, com $E(\epsilon|X = x) = 0$ e $V(\epsilon|X = x) = \sigma^2(x)$ não necessariamente constante. O modelo (4.2) remove as restrições paramétricas para $m(x)$ e permite que estruturas de várias formas

sejam consideradas. Por definição

$$\begin{aligned} m(x) &= E(Y|X = x) \\ &= \int y f(y|x) dy \\ &= \int y \frac{f(x, y)}{f_X(x)} dy, \end{aligned} \tag{4.3}$$

onde $f_X(x)$, $f(x, y)$, e $f(y|x)$ são a densidade marginal de X , a densidade conjunta de X e Y , e a densidade condicional de Y dado X , respectivamente.

Um estimador produto para $f(x, y)$ é

$$\hat{f}(x, y) = \frac{1}{nh_x h_y} \sum_{i=1}^n K_x \left(\frac{x - x_i}{h_x} \right) K_y \left(\frac{y - y_i}{h_y} \right),$$

enquanto que um estimador para $f_X(x)$ é

$$\hat{f}_X(x) = \frac{1}{nh_x} \sum_{i=1}^n K_x \left(\frac{x - x_i}{h_x} \right).$$

Considerando o fato de que $\int K_y(u) du = 1$ e $\int u K_y(u) du = 0$ em (4.3), chegamos ao seguinte estimador, que é denominado estimador de Nadaraya-Watson

$$\hat{m}_h(x) = \frac{\sum_i Y_i K[(x - x_i)/h]}{\sum_i K[(x - x_i)/h]}, \tag{4.4}$$

onde K é uma função de densidade chamada núcleo e h é o parâmetro de suavização que chamaremos de janela.

O estimador de Nadaraya-Watson é o mais natural para dados com efeito aleatório na variável X , ou seja, quando X é uma variável aleatória com densidade f_X . Quando X não é aleatória, mas sim um conjunto fixo de números ordenados, a intuição de (4.4) é perdida e uma forma diferente de estimar deve ser considerada. Uma abordagem é utilizar a seguinte estimativa para a função de densidade

$$\hat{f}_X(x_i) = \frac{1}{n(x_i - x_{i-1})}.$$

Ao substituir esta estimativa em (4.3), temos o estimador de Priestley-Chao.

Utilizando de cálculo básico, podemos observar que o estimador de Nadaraya-Watson dado por (4.4) é a solução de um problema de mínimos quadrados ponderados e é o minimizador $\hat{\beta}_0$ de

$$\sum_{i=1}^n (y_i - \beta_0)^2 K\left(\frac{x - x_i}{h}\right). \quad (4.5)$$

Dessa forma, o estimador dado por (4.4) corresponde a aproximar $m(x)$ por uma "constante" β_0 , ponderando os valores de y pelos x'_i s mais próximos de x . É importante observar que β_0 não é constante uma vez que sua estimativa é dada pela equação (4.4), que depende de todos os valores de x_i .

4.1.2 Regressão Polinomial Local

O fato do estimador de Nadaraya-Watson ser equivalente a uma aproximação polinomial local por uma "constante", nos sugere que podemos utilizar uma aproximação polinomial local com uma ordem superior. Tal estimador de regressão polinomial local de p -ésima ordem, é o minimizador de

$$\sum_{i=1}^n (y_i - \beta_0 - \dots - \beta_p(x - x_i)^p)^2 K\left(\frac{x - x_i}{h}\right). \quad (4.6)$$

Seja X a matriz de desenho

$$\begin{pmatrix} 1 & x - x_1 & & (x - x_1)^p \\ \vdots & \vdots & \dots & \vdots \\ 1 & x - x_n & & (x - x_n)^p \end{pmatrix}, \quad (4.7)$$

e seja

$$W = h^{-1} \text{diag} \left[K\left(\frac{x - x_1}{h}\right), \dots, K\left(\frac{x - x_n}{h}\right) \right] \quad (4.8)$$

a matriz de pesos. Então, se $X'WX$ é invertível,

$$\hat{\beta} = (X'WX)^{-1} X'Wy. \quad (4.9)$$

O estimador $\hat{m}_p(x)$ é então o termo do intercepto $\hat{\beta}_0$, ou em notação matricial

$$\hat{m}_p(x) = e_r'(X'WX)^{-1}X'Wy, \quad (4.10)$$

onde e_r é um vetor de tamanho $(p + 1)$ que tem o valor 1 na r -ésima entrada e zero nas outras.

Como podemos observar na expressão (4.10), para encontrar a estimativa de $m_p(x)$ o único problema é determinar a matriz de pesos W cuja única quantidade desconhecida é a janela h .

4.2 Estimação do Parâmetro de Suavização

É bem conhecido na literatura que a precisão do estimador depende crucialmente da janela h . Vários métodos para escolha automática da janela foram exaustivamente estudados, incluindo Validação Cruzada, Validação Cruzada Generalizado e Critério de Informação de Akaike. Härdle *et al.* (1988) demonstram que esses métodos estão sujeitos a uma grande variabilidade, e é por isso que não se tornaram amplamente usados como uma ferramenta para análise dos dados.

Em nosso trabalho, como a variável aleatória X consiste das observações de tempo discreto de uma variável aleatória que segue um processo de difusão dado pela equação (1.1), consideraremos somente os estimadores apropriados para efeitos aleatórios. Como mencionado anteriormente, após a definição de qual estimador utilizar, é necessário estimar a janela ótima. Para o estimador de Nadaraya-Watson, Härdle *et al.* (1992) propuseram uma técnica de dupla suavização para estimação da janela.

Já no segundo caso, de acordo com Fan e Gijbels (1995), a performance dos estimadores de regressão polinomial local depende crucialmente de dois parâmetros: a ordem do ajuste polinomial local e a janela. O ajuste polinomial local pode capturar curvas que são altamente não-homogêneas ao escolher adequadamente a ordem do ajuste e/ou a janela, que pode ser fixa ou variável. Estas duas escolhas reduzem o erro quadrático médio nos casos onde a ordem é fixa e a janela constante não consegue ser uma estimativa satisfatória. Porém, devido a uma alta correlação entre a ordem do ajuste e o tamanho da janela, estimar os dois parâmetros simultaneamente não é uma tarefa fácil. Em Fan e Gijbels (1995) o objetivo foi simplesmente estimar a melhor janela para o ajuste polinomial local. Os autores propuseram uma técnica que consiste em dois estágios. No primeiro estima-se uma janela piloto e no segundo estágio do procedimento estimamos a janela mais refinada.

De acordo com os autores mencionados, este método de suavização tem excelente propriedades minimax. Dessa forma, não há necessidade de modificações para problemas de fronteira, é aplicável para várias formas de funções, é fácil de implementar e ainda é adaptável à estimação das derivadas das funções. Por todas estas características, utilizamos este método para estimação da janela quando utilizamos o estimador de regressão polinomial local.

4.2.1 Método de Dois Estágios para Regressão Polinomial Local

Se a $(p+1)$ -ésima derivada de $m(x)$ no ponto x_0 existe, podemos aproximar localmente $m(x)$ dado em (4.2) por um polinômio de ordem p

$$m(x) \approx m(x_0) + m'(x_0)(x - x_0) + \dots + m^{(p)}(x_0)(x - x_0)^p/p!. \quad (4.11)$$

Através de uma regressão polinomial local temos

$$\sum_{i=1}^n \left\{ Y_i - \sum_{j=0}^p \beta_j (X_i - x_0)^j \right\}^2 K \left(\frac{X_i - x_0}{h} \right). \quad (4.12)$$

Observe que a expressão (4.12) é igual a expressão (4.6).

É mais conveniente escrever o problema de mínimos quadrados em notação matricial. Portanto, consideremos W como uma matriz diagonal com entradas $W_i = K\{(X_i - x_0)/h\}$. Seja X a matriz de desenho cujo (i, j) -ésimo elemento é $(X_i - x_0)^{j-1}$ e seja y um vetor coluna. Então o problema de mínimos quadrados ponderados pode ser escrito como

$$\min_{\beta} \{(y - X\beta)^T W (y - X\beta)\}. \quad (4.13)$$

A teoria de mínimos quadrados nos fornece a solução

$$\hat{\beta} = (X'WX)^{-1} X'W y, \quad (4.14)$$

cuja média e variância condicional são, respectivamente

$$E(\hat{\beta}|X_1, \dots, X_n) = (X'WX)^{-1} X'W m = \beta + (X'WX)^{-1} X'W r, \quad (4.15)$$

$$Var(\hat{\beta}|X_1, \dots, X_n) = (X'WX)^{-1} (X'\Sigma X) (X'WX)^{-1}, \quad (4.16)$$

onde $m = \{m(X_1), \dots, m(X_n)\}'$, $r = m - X\beta$ é o resíduo da aproximação local polinomial e $\Sigma = \text{diag}[K^2\{(X_i - x_0)/h\}\sigma^2(X_i)]$.

Utilizando o método proposto por Fan e Gijbels (1995), a seleção da janela é dividida em dois estágios. Primeiramente, definiremos o procedimento que automaticamente seleciona a janela no primeiro estágio. Chamaremos esta estimativa de janela piloto.

PRIMEIRO ESTÁGIO-SELEÇÃO DA JANELA PILOTO

A princípio é necessário introduzir o critério de quadrados residuais, que a partir de agora será denominado RSC (Residual Squares Criterion). O RSC estima o erro quadrático médio local, que chamaremos de MSEs (Mean Squared Error). Entretanto, este procedimento é un tanto quanto primitivo e apresenta uma taxa muito lenta de convergência. Esta é a principal razão pela qual utilizamos o RSC somente como um procedimento piloto para estimar a curvatura.

O RSC é definido por

$$RSC(x_0, h) = \hat{\sigma}^2(x_0)\{1 + (p + 1)V\}, \quad (4.17)$$

onde

$$\hat{\sigma}^2(x_0) = \frac{1}{tr\{W - (X'WX)^{-1}X'W^2X\}} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 K\left(\frac{x - x_i}{h}\right) \quad (4.18)$$

e V é o primeiro elemento da diagonal da matriz $S_n^{-1}S_n^*S_n^{-1}$, com $S_n = X'WX$ e $S_n^* = X'W^2X$.

A intuição por trás da estatística (4.18) é a seguinte: quando a regressão polinomial local não se ajusta bem, isto é, a janela h é muito grande, o vício é grande e logo a soma dos quadrados residuais $\hat{\sigma}^2(x_0)$ também é. Quando a janela h é muito pequena, o termo da variância V tende a ser maior. Então, a quantidade RSC protege contra as duas escolhas extremas.

Maiores detalhes e justificativas teóricas podem ser encontradas no artigo de Fan e Gijbels (1995) mas é interessante ressaltar que a quantidade RSC pode ser aplicada quando encontramos erros homocedásticos. Na prática, podemos aplicar naturalmente a erros quase homocedásticos. Para ruídos com grande heterocedasticidade é recomendado selecionar a janela utilizando o método mais refinado que é proposto no segundo estágio.

O uso do RSC é feito da seguinte forma. Suponhamos que queremos estimar $m(x)$, onde x se encontra no intervalo $[c, d]$. Então, obtemos \hat{h} que minimiza a versão integrada de RSC

$$IRSC(h) = \int_{[c,d]} RSC(y; h)dy. \quad (4.19)$$

SEGUNDO ESTÁGIO-ACESSO DO VÍCIO E DA VARIÂNCIA DO AJUSTE POLINOMIAL LOCAL

O vício e a variância nas equações (4.15) e (4.16) não são diretamente acessíveis, pois dependem de quantidades desconhecidas, o resíduo r e a matriz diagonal Σ . É desejável ter boas estimativas do vício e da variância para amostras finitas, para garantir boa performance do processo de seleção da janela.

O vício $(X'WX)^{-1}X'Wr$, com $r = m - X\beta$, pode ser aproximado por $(X'WX)^{-1}X'W\tau$, onde o i -ésimo elemento do vetor τ de tamanho $n \times 1$, é igual a:

$$\beta_{p+1}(X_i - x_0)^{p+1} + \dots + \beta_{p+a}(X_i - x_0)^{p+a} \quad (4.20)$$

A escolha $a = 4$ garante que o procedimento de seleção da janela proposto seja \sqrt{n} consistente para o ajuste linear local. Entretanto, a escolha $a = 2$ leva a uma redução do esforço computacional, enquanto ainda apresenta uma regra de seleção que não é muito longe de ser \sqrt{n} consistente. Isso faz esse último caso mais atraente do ponto de vista computacional. Em nosso trabalho, utilizamos $a = 2$.

Seja

$$s_{n,j} = \sum_{i=1}^n (X_i - x_0)^j K \left(\frac{X_i - x_0}{h} \right). \quad (4.21)$$

O vício aproximado é equivalente a

$$S_n^{-1} \begin{pmatrix} \beta_{p+1}s_{n,p+1} + \dots + \beta_{p+a}s_{n,p+a} \\ \vdots \\ \beta_{p+1}s_{n,2p+1} + \dots + \beta_{p+a}s_{n,2p+a} \end{pmatrix}, \quad (4.22)$$

onde $S_n = X'WX$ é uma matriz $(p+1) \times (p+1)$ cujo elemento (i, j) é $s_{n,i+j-2}$.

Então, o vetor de vícios pode ser estimado por

$$S_n^{-1} \begin{pmatrix} \hat{\beta}_{p+1}s_{n,p+1} + \dots + \hat{\beta}_{p+a}s_{n,p+a} \\ \vdots \\ \hat{\beta}_{p+1}s_{n,2p+1} + \dots + \hat{\beta}_{p+a}s_{n,2p+a} \end{pmatrix}, \quad (4.23)$$

onde $\hat{\beta}_{p+1}, \dots, \hat{\beta}_{p+a}$ são os coeficientes de regressão estimados ao ajustar localmente um polinômio de ordem $p+a$, com a janela piloto selecionada, por exemplo, pelo método do primeiro estágio.

A diferença existente entre a abordagem do autor e o método usual do plug-in é que não é substituído $s_{n,j}$ por sua parte assintótica. As quantidades $s_{n,j}$ já são conhecidas e substituí-las pelas correspondentes quantidades assintóticas introduz não somente uma aproximação extra mas também mais parâmetros desconhecidos, por exemplo $f_X(x_0)$. A idéia abordada faz com que seja possível acessar o vício sem que seja necessário aprofundar na parte assintótica.

A matriz de variâncias da equação (4.16) pode ser aproximada por

$$(X'WX)^{-1}(X'W^2X)(X'WX)^{-1}\sigma^2(x_0). \quad (4.24)$$

Substituindo $\sigma^2(x_0)$ por $\hat{\sigma}^2(x_0)$ estimado na equação (4.18) com o ajuste polinomial de ordem $p + a$, temos a seguinte estimativa para a variancia

$$(X'WX)^{-1}(X'W^2X)(X'WX)^{-1}\hat{\sigma}^2(x_0). \quad (4.25)$$

Então, o MSE de $\hat{\beta}_0(x_0)$ é estimado por

$$\widehat{MSE}_{p,0}(x_0, h) = \hat{b}_{p,0}^2(x_0) + \hat{V}_{p,0}(x_0), \quad (4.26)$$

onde $\hat{b}_{p,0}(x_0)$ estima o vício e é dado pelo primeiro elemento do vetor de vícios estimados na expressão (4.23), e $\hat{V}_{p,0}(x_0)$ é o primeiro elemento diagonal da matriz na expressão (4.25). O MSE estimado é utilizado para encontrar a janela final no procedimento de seleção da janela.

AS REGRAS PARA A SELEÇÃO DA JANELA

Para estimar a janela ótima procederemos da seguinte maneira:

1. Escolhemos $p=1$
2. Ajustamos um polinômio de ordem 3 e encontramos o IRSC na equação (4.19) para os valores de h que desejamos pesquisar.
3. Escolhemos o h que minimiza o IRSC na equação (4.19)
4. Com o h escolhido, obtemos as estimativas $\hat{\beta}_2(x_0)$, $\hat{\beta}_3(x_0)$ e $\sigma^2(x_0)$, ainda no primeiro estágio.
5. Com as estimativas encontradas no item anterior, estimamos o vetor de vício e a matriz de variância, nas equações (4.23) e (4.25), respectivamente.
6. Encontramos a estimativa de MSE
7. Finalmente, encontramos o valor de h que minimiza:

$$\hat{h}_{1,0}^R = \arg \min_h \left\{ \int_{[c,d]} \widehat{MSE}_{1,0}(y; h) dy \right\} \quad (4.27)$$

5 Validação do Método de Escolha da Janela

O objetivo deste capítulo é avaliar como se comporta o método de estimação da janela para a regressão polinomial local proposto por Fan e Gijbels (1995). Consideremos a variável aleatória X com distribuição Uniforme(0,1) e os casos que se encontram na tabela 5.1.

Tabela 5.1: Casos considerados para validação do método de estimação da janela ótima.

| Caso | $m(x)$ |
|----------|---------------------------------|
| Caso I | $m(x) = \sin(2x) + 2e^{-16x^2}$ |
| Caso II | $m(x) = x + 2e^{-16x^2}$ |
| Caso III | $m(x) = \sin(2\pi x)$ |
| Caso IV | $m(x) = 0.4x + 1$ |

Para todos os casos apresentamos um exemplo gráfico onde MpH1 é a estimativa da função $m(x)$ com o valor de h obtido no primeiro estágio e MpH2 é a estimativa da função $m(x)$ com o valor de h obtido no segundo estágio. Em todos os casos, as simulações foram feitas para tamanhos de amostra iguais a 50. Como este tamanho de amostra é bem pequeno e o procedimento para estimação tende a melhorar com tamanhos de amostra maiores, as conclusões tiradas podem se estender para casos com tamanhos de amostras maiores. Para confirmar o que foi dito, escolheremos um dos casos para mostrar o aumento do desempenho do método quando aumentamos o tamanho amostral.

É importante ressaltar que com o aumento do tamanho de amostra, o tempo de simulação cresce bastante, uma vez que o procedimento depende de operações de inversão de matrizes.

Um comentário importante a ser feito é que para o cálculo das integrais numéricas apresentadas durante todo o texto, foi utilizado o Método Composto de Simpson. Este método é descrito no teorema a seguir, de acordo com Mathews (1987).

Teorema 5.0.1 *Seja $f, f^{(1)}, f^{(2)}, f^{(3)}$ e $f^{(4)}$ contínuas no intervalo $[a, b]$. Se dividimos o intervalo em $2M$ subintervalos de igual largura $h = (b - a)/2M$ e usarmos os nodos $a = x_0 < x_1 < \dots < x_{2M} = b$, onde $x_k = a + hk$ para $k = 0, 1, \dots, 2M$, a regra de Simpson para $f(x)$ com $2M$ intervalos de largura h é*

$$S(f, h) = \frac{h}{3} \sum_{k=1}^M [f(x_{2k-2}) + 4f(x_{2k-1}) + f(x_{2k})] \quad (5.1)$$

5.1 Caso I - $m(x) = \sin(2x) + 2e^{-16x^2}$

Consideremos o modelo $m(x) = \sin(2x) + 2e^{-16x^2} + \epsilon$, onde X é Uniforme(0,1) e ϵ é Normal(0,0.1). O exemplo para 1 amostra é mostrado na Figura 5.1:

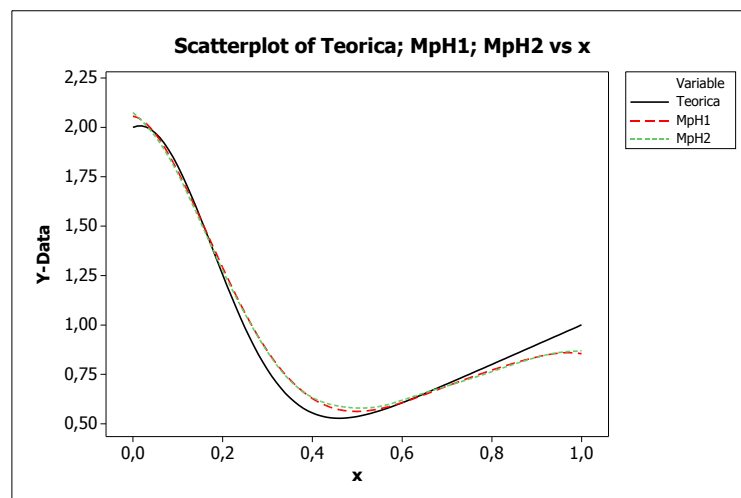


Figura 5.1: Exemplo Caso I.

Como podemos observar não existe muita diferença nas estimativas da função quando utilizamos as janelas 1 e 2. Para as duas janelas as curvas estimadas acompanham a curva teórica. Na Tabela 5.2 podemos observar os valores das janelas e do Erro Quadrático Médio-EQI (neste capítulo chamaremos de ISE) para as duas escolhas de janela após uma simulação com 1000 amostras.

Tabela 5.2: Caso I-Janela e ISE nos 2 Estágios.

| | Média | Desvio-Padrão | Mínimo | Q1 | Mediana | Q3 | Máximo |
|-----------|----------|---------------|----------|----------|----------|----------|----------|
| Janela E1 | 0,741420 | 0,261390 | 0,049680 | 0,508770 | 0,908290 | 0,959310 | 0,989960 |
| Janela E2 | 0,411140 | 0,298200 | 0,049340 | 0,179660 | 0,279370 | 0,576430 | 0,989960 |
| ISE E1 | 0,000977 | 0,000943 | 0,000019 | 0,000390 | 0,000761 | 0,001228 | 0,011440 |
| ISE E2 | 0,000652 | 0,000685 | 0,000001 | 0,000207 | 0,000481 | 0,000854 | 0,009630 |

As informações contidas na Tabela 5.2 podem ser melhor visualizadas na Figura 5.2.

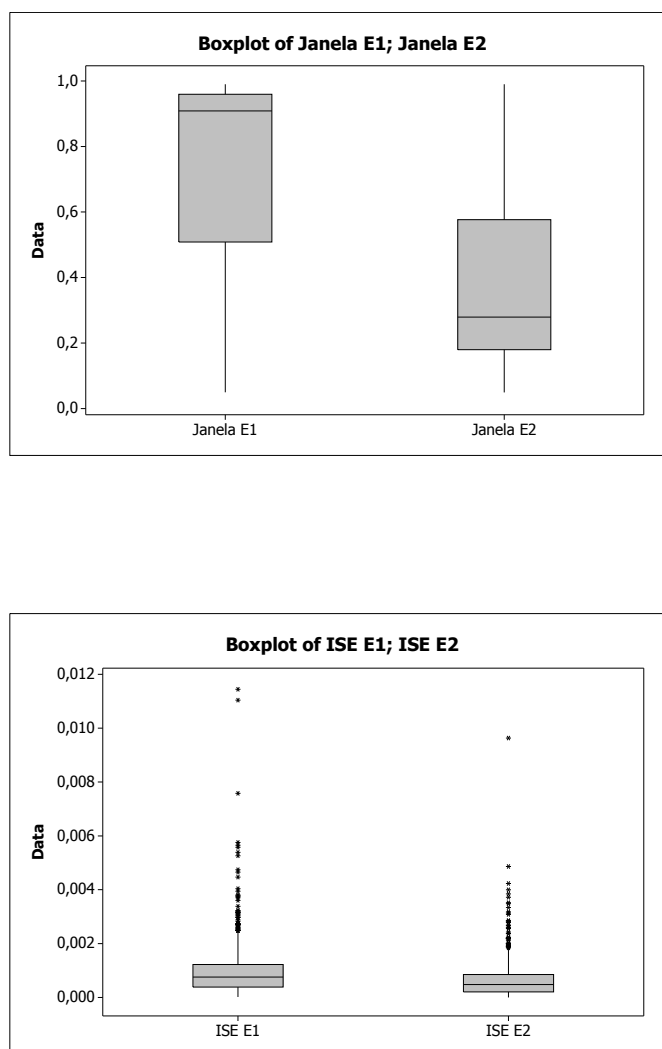


Figura 5.2: Janelas e ISE Caso I.

Como podemos observar na Tabela 5.2 e na Figura 5.2, a janela do Estágio 2 apresenta valores menores comparados às do Estágio 1. Em relação à variabilidade, praticamente não há diferença entre os dois estágios. Porém as janelas do Estágio 2, apresentam uma variabilidade ligeiramente menor. Além disso, as estimativas apresentam um Erro Quadrático Integrado (ISE) menor quando utilizamos as janelas do Estágio 2, como podemos observar no gráfico inferior na Figura 5.2.

5.2 Caso II - $m(x) = x + 2e^{-16x^2}$

Consideremos o modelo $m(x) = x + 2e^{-16x^2} + \epsilon$, onde X é Uniforme(0,1) e ϵ é Normal(0,0.1). O exemplo para 1 amostra é mostrado na Figura 5.3.

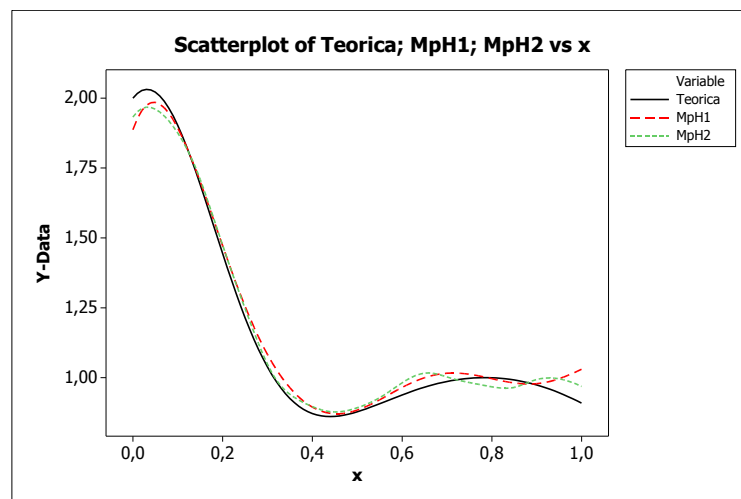


Figura 5.3: Exemplo Caso II.

Também neste caso, como podemos observar, não existe muita diferença nas estimativas da função quando utilizamos as janelas dos estágios 1 e 2. Para as duas janelas, as curvas estimadas acompanham a curva teórica para quase todos os valores de x . Na Tabela 5.3 podemos observar os valores das janelas e do ISE (definido no Capítulo 2) para as duas escolhas após uma simulação com 1000 amostras.

Tabela 5.3: Caso II-Janela e ISE nos 2 Estágios.

| | Média | DP | Mínimo | Q1 | Mediana | Q3 | Máximo |
|-----------|----------|----------|----------|----------|----------|----------|----------|
| Janela E1 | 0,244880 | 0,240930 | 0,048820 | 0,119180 | 0,139270 | 0,199730 | 0,979780 |
| Janela E2 | 0,063532 | 0,009231 | 0,048538 | 0,058994 | 0,059855 | 0,069343 | 0,107351 |
| ISE E1 | 0,003661 | 0,003914 | 0,000108 | 0,001333 | 0,002177 | 0,004443 | 0,032132 |
| ISE E2 | 0,002389 | 0,001500 | 0,000461 | 0,001456 | 0,002100 | 0,002885 | 0,016669 |

Os gráficos abaixo possibilitam uma melhor visualização dos dados da Tabela 5.3.

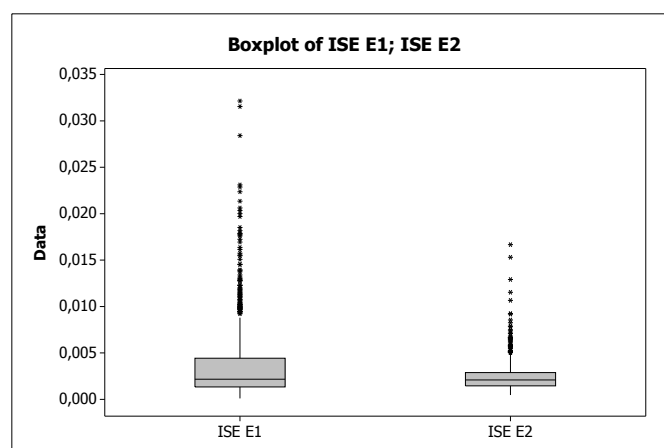
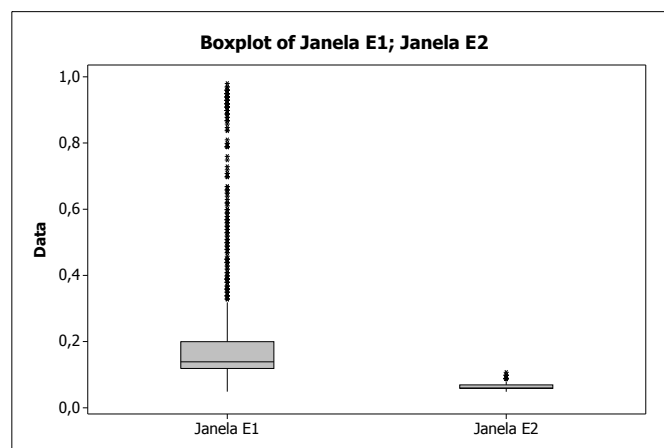


Figura 5.4: Janelas e ISE Caso II.

Como podemos observar a janela do Estágio 2 apresenta valores menores comparados às do Estágio 1. Quando comparamos as janelas dos estágios 1 e 2, observamos que as janelas do Estágio 2 são mais homogêneas e apresentam um Erro Quadrático Integrado (ISE) bem menor. Podemos observar melhor no segundo gráfico da Figura 5.4.

5.3 Caso III - $m(x) = \sin(2\pi x)$

Consideremos o modelo $m(x) = \sin(2\pi x) + \epsilon$, onde X é Uniforme(0,1) e o ϵ é Normal(0,0.1). O exemplo para 1 amostra é mostrado na Figura 5.5:

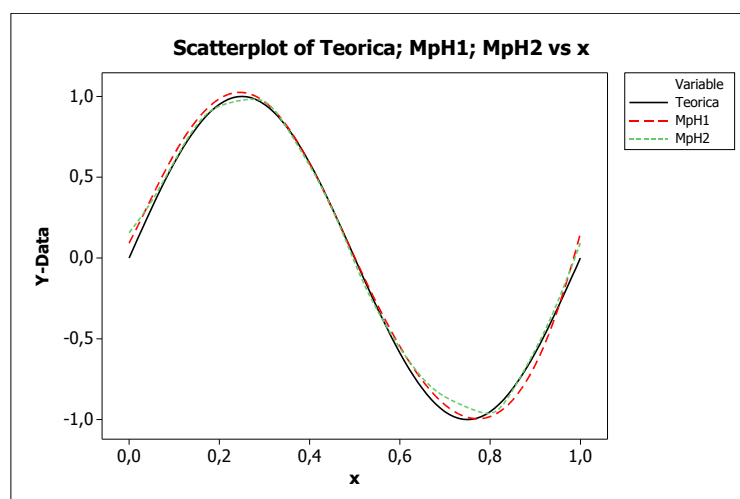


Figura 5.5: Exemplo Caso III.

Como observamos na Figura 5.5, a estimativa da função $m(x)$ utilizando o valor de h piloto, descrito no método, não apresenta grandes diferenças, em relação à proximidade com a função teórica, quando comparada à estimativa utilizando o valor de h obtido no segundo estágio.

Os resultados para as janelas obtidas após simulação com 1000 amostras são apresentados na tabela e nos gráfico abaixo. O ISE foi calculado e o resumo das 1000 amostras se encontra na tabela a seguir.

Os gráficos abaixo possibilitam uma melhor visualização dos dados da Tabela 5.4.

Tabela 5.4: Caso III-Janela e ISE nos 2 Estágios.

| | Média | Desvio-Padrão | Mínimo | Q1 | Mediana | Q2 | Máximo |
|-----------|----------|---------------|----------|----------|----------|----------|----------|
| Janela E1 | 0,268230 | 0,261870 | 0,048860 | 0,119530 | 0,139150 | 0,338840 | 0,989870 |
| Janela E2 | 0,049247 | 0,000927 | 0,046899 | 0,048920 | 0,049330 | 0,049596 | 0,066464 |
| ISE E1 | 0,003458 | 0,003492 | 0,000149 | 0,001243 | 0,002198 | 0,005142 | 0,038681 |
| ISE E2 | 0,003304 | 0,002053 | 0,000482 | 0,002102 | 0,002900 | 0,003928 | 0,023647 |

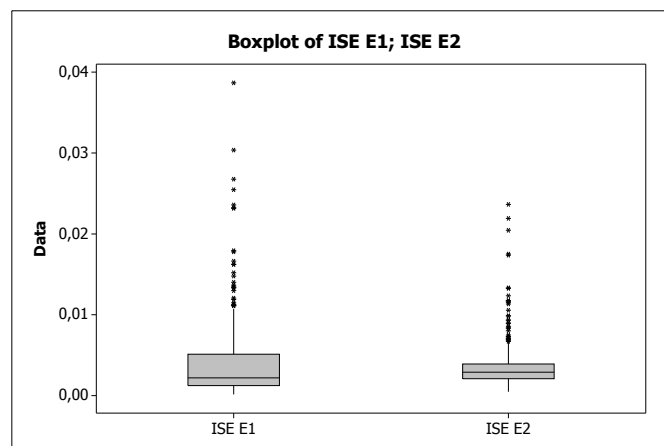
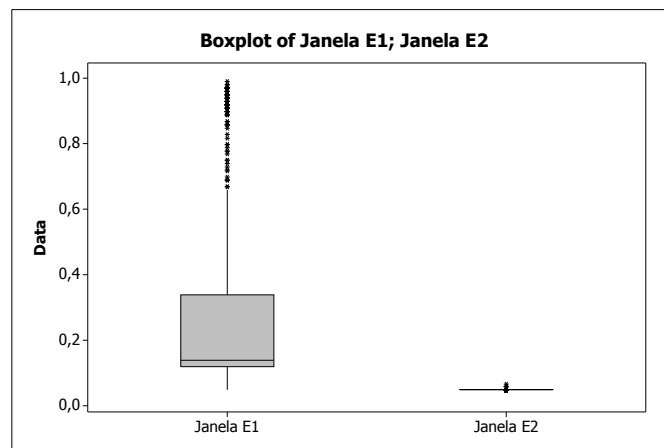


Figura 5.6: Janelas e ISE Caso III.

Observamos que para a função m utilizada, os valores de h obtidos no primeiro estágio são superiores aos valores de h no segundo estágio. Além disso, a variabilidade dos valores de h diminui consideravelmente no segundo estágio. O ISE também foi menor quando utilizamos as janelas do segundo estágio. Podemos visualizar melhor na Figuras 5.6.

5.4 Caso IV - $m(x) = 0.4x + 1$

Consideremos o modelo $m(x) = 0.4x + 1 + \epsilon$, onde X é Uniforme(0,1) e o ϵ é Normal(0,0.1). O exemplo para 1 amostra é mostrado na Figura 5.7:

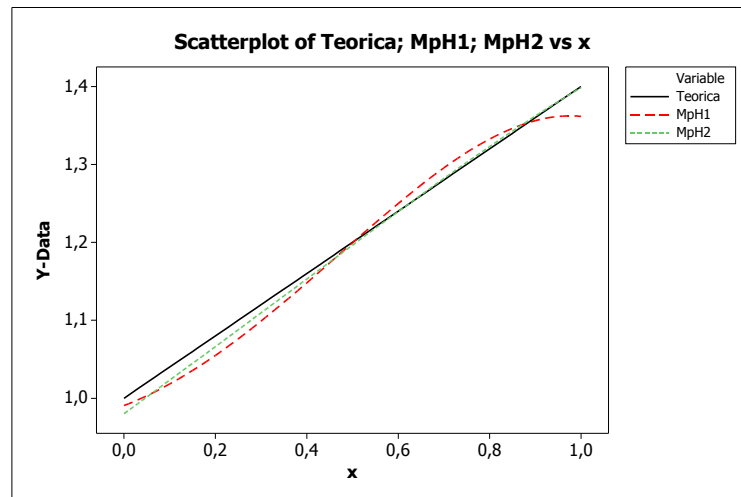


Figura 5.7: Exemplo Caso IV.

Para este último caso considerado, observamos na Figura 5.7 que a estimativa da função m utilizando o valor de h piloto é mais distante da curva teórica, quando comparada à estimativa utilizando o valor de h obtido no segundo estágio.

Este caso é de grande interesse pois alguns modelos supõem que o coeficiente de tendência tem comportamento linear. Os resultados para as janelas obtidas após simulação com 1000 amostras são apresentados na Tabela 5.5 e na Figura 5.8. O ISE também foi calculado e o resumo das 1000 amostras se encontra na Tabela 5.5.

Tabela 5.5: Caso IV-Janela e ISE nos 2 Estágios.

| | Média | DP | Mínimo | Q1 | Mediana | Q3 | Máximo |
|-----------|----------|----------|----------|----------|----------|----------|----------|
| Janela E1 | 0,728840 | 0,258380 | 0,079120 | 0,488770 | 0,887820 | 0,959310 | 0,989980 |
| Janela E2 | 0,408080 | 0,298310 | 0,049120 | 0,179590 | 0,269190 | 0,589540 | 0,989970 |
| ISE E1 | 0,000960 | 0,000783 | 0,000033 | 0,000425 | 0,000748 | 0,001277 | 0,007561 |
| ISE E2 | 0,000647 | 0,000638 | 0,000002 | 0,000212 | 0,000455 | 0,000887 | 0,005636 |

Os gráficos abaixo possibilitam uma melhor visualização dos dados da Tabela 5.5.

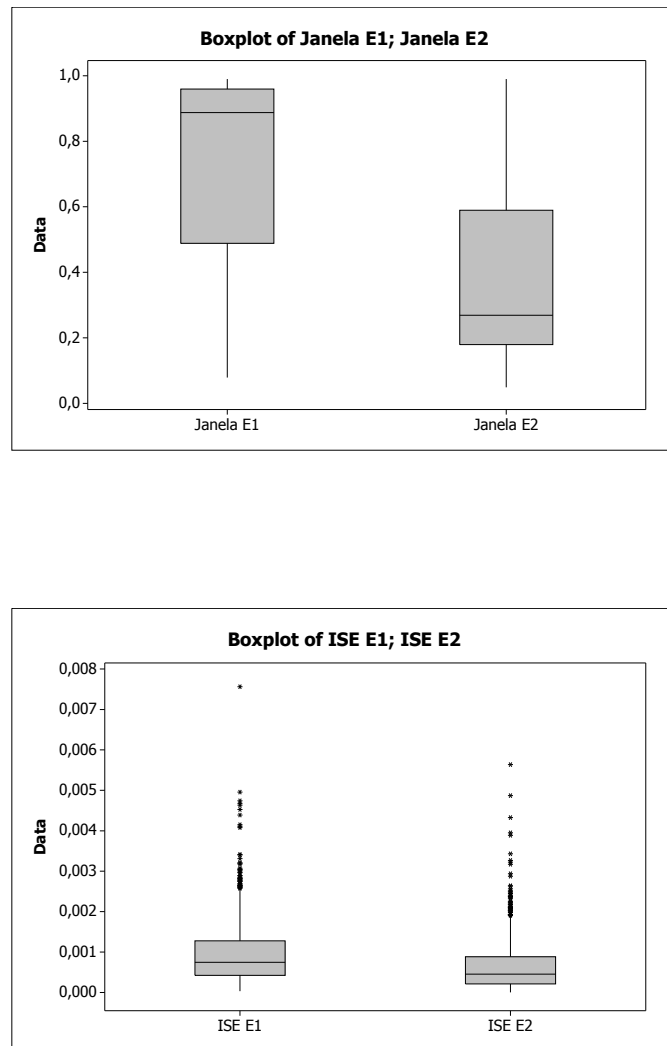


Figura 5.8: Janelas e ISE Caso IV.

Analogamente aos outros casos, para a função m utilizada, os valores de h obtidos no primeiro estágio são superiores aos valores de h no segundo estágio. Além disso, a variabilidade dos valores de h diminui consideravelmente no segundo estágio. O ISE também foi menor quando utilizamos as janelas do segundo estágio. Podemos visualizar melhor na Figuras 5.8.

5.5 Comparação para Diferentes Tamanhos Amostrais-Caso III

Como foi afirmado no início deste capítulo, o método de estimação da janela aumenta sua eficiência quando aumentamos o tamanho amostral. Somente para ilustrar e confirmar, consideramos o modelo do Caso III e geramos amostras de tamanho 100 e 200. Comparamos os casos onde $n = 50, 100$ e 200 . Foi calculado o ISE e o resultado é mostrado na Figura 5.9.

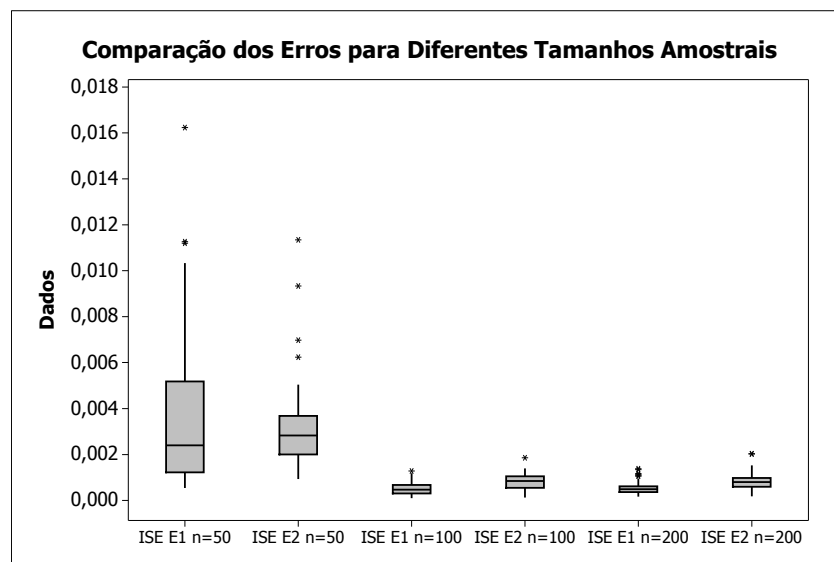


Figura 5.9: Comparação ISE para Diferentes Tamanhos Amostrais.

5.6 Conclusões do Capítulo

Para as diversas funções utilizadas pode ser observado que o método de regressão polinomial local é eficiente para estimar as funções. Como observamos nos exemplos do capítulo, as curvas estimadas se aproximam das curvas teóricas. Este fato, além de ser observado nas Figuras 5.1, 5.3, 5.5 e 5.7, se confirma quando simulamos as 1000 amostras para cada caso considerado.

Um caso interessante para ser analisado é o Caso IV. Como podemos observar, o método de regressão polinomial local forneceu boas estimativas mesmo se tratando de uma função linear. Este caso é de grande utilidade pois na área de economia muitas vezes são utilizados modelos onde o coeficiente de tendência $\mu(\cdot)$ do modelo (1.1) é linear.

Como observamos nas Figuras 5.2, 5.4, 5.6 e 5.8, os valores encontrados para o Erro Quadrático Integrado- ISE (definido no capítulo 2) são ligeiramente menores quando utilizamos a janela estimada no Estágio 2 quando comparamos com a janela estimada no Estágio 1. Além disso, as janelas do Estágio 2 apresentam menor variabilidade.

Como as janelas estimadas no Estágio 2 apresentam menor Erro Quadrático Integrado-ISE e uma menor variabilidade, conclui-se que a melhor opção é utilizar o método não-paramétrico até a obtenção da janela mais refinada e não parar quando estimamos a janela piloto.

6 Simulações

Neste capítulo, nosso objetivo principal é estimar as funções $\mu(X_t)$ e $b(X_t)$. Partimos do pressuposto que temos um modelo conhecido para gerar as observações através de um dos esquemas de discretização mencionados no Capítulo 2.

Comparamos o método não-paramétrico com o método paramétrico descrito pelos autores Cleur e Manfredi (1999).

Os parâmetros para as simulações foram escolhidos pensando em situações reais e foram usados em estudos similares por Bianchi e Cleur (1996), Bianchi *et al.* (1994) e Cleur e Manfredi (1999).

É importante observar que, com os parâmetros utilizados, os esquemas de discretização de Euler e Milstein não apresentam diferenças consideráveis e portanto decidimos utilizar o esquema de Euler por ser mais simples. Para ilustrar este fato, apresentamos no final da seção 6.3 um exemplo comparando os dois esquemas.

É importante ressaltar que as rotinas foram implementadas no software gratuito R e totalmente construídas no decorrer do trabalho.

6.1 Estimação Paramétrica do Coeficiente de Tendência

Seguindo Cleur e Manfredi (1999), consideremos o modelo geral

$$dX_t = \kappa(\nu - X_t)X_t^\alpha dt + \sigma X_t^\beta dB_t. \quad (6.1)$$

Cada um dos modelos que apresentaremos nas próximas seções apresenta um valor para as constantes α e β da seguinte maneira

| Modelo | α | β |
|--------------------------------|----------|---------|
| CIR | 0 | 0.5 |
| Brennan and Schwarz | 0 | 1 |
| Heuristically Derived logistic | 1 | 1 |

As propriedades destes modelos são bem conhecidas na literatura. O objetivo é estimar o coeficiente de tendência através do chamado CTMLE (Continuous Time Maximum Likelihood Estimator). É importante lembrar que em nosso trabalho o enfoque é dado na parte de estimação não-paramétrica. Utilizamos este método de estimação paramétrico somente para comparação. Maiores detalhes podem ser vistos em Cleur e Manfredi (1999).

A estimativa dos parâmetros κ e ν têm a forma

$$\hat{\kappa} = \frac{I_1 I_5 - I_2 I_3}{I_3 I_4 - I_5^2} \quad \hat{\nu} = \frac{I_1 I_4 - I_2 I_5}{I_1 I_5 - I_2 I_3}, \quad (6.2)$$

onde I_j são as integrais reportadas na Tabela 6.1.

Tabela 6.1: Integrais envolvidas no CTMLE para os 3 modelos.

| Modelo | I_1 | I_2 | I_3 | I_4 | I_5 |
|---------------------|-------------------------------|-----------------------------|------------------------------|----------------------|----------------------------|
| CIR | $\int_0^T \frac{dX_s}{X_s}$ | $\int_0^T dX_s$ | $\int_0^T \frac{d_s}{X_s}$ | $\int_0^T X_s d_s$ | $\int_0^T d_s$ |
| Brennan and Schwarz | $\int_0^T \frac{dX_s}{X_s^2}$ | $\int_0^T \frac{dX_s}{X_s}$ | $\int_0^T \frac{d_s}{X_s^2}$ | $\int_0^T d_s$ | $\int_0^T \frac{d_s}{X_s}$ |
| HD Logistic | $\int_0^T \frac{dX_s}{X_s}$ | $\int_0^T dX_s$ | $\int_0^T d_s$ | $\int_0^T X_s^2 d_s$ | $\int_0^T X_s d_s$ |

Para as integrais não estocásticas, o autor sugere utilizar a regra do trapézio e as integrais são avaliadas da forma mostrada na tabela a seguir.

Tabela 6.2: Integrais não-estocásticas utilizando a regra do Trapézio.

| Modelo | I_3 | I_4 | I_5 |
|---------------------|---|---|---|
| CIR | $\Delta \sum \frac{X_{n-1}^{-1} + X_n^{-1}}{2}$ | $\Delta \sum \frac{X_{n-1} + X_n}{2}$ | T |
| Brennan and Schwarz | $\Delta \sum \frac{X_{n-1}^{-2} + X_n^{-2}}{2}$ | T | $\Delta \sum \frac{X_{n-1}^{-1} + X_n^{-1}}{2}$ |
| HD Logistic | T | $\Delta \sum \frac{X_{n-1}^2 + X_n^2}{2}$ | $\Delta \sum \frac{X_{n-1} + X_n}{2}$ |

As integrais estocásticas são reduzidas a uma forma não estocástica aplicando sistematicamente a fórmula de Itô (detalhes em Oksendal (2003)). Desta forma as integrais estocásticas mostradas na Tabela 6.1 são calculadas da forma mostrada na tabela a seguir.

Tabela 6.3: Integrais estocásticas utilizando a fórmula de Itô.

| Modelo | I_1 | I_2 |
|---------------------|---|---|
| CIR | $\log \frac{X_T}{X_0} + \frac{\sigma^2}{2} I_3$ | $X_T - X_0$ |
| Brennan and Schwarz | $\sigma^2 I_5 - \left(\frac{1}{X_T} - \frac{1}{X_0} \right)$ | $\log \frac{X_T}{X_0} + \frac{\sigma^2}{2} T$ |
| HD Logistic | $\log \frac{X_T}{X_0} + \frac{\sigma^2}{2} T$ | $X_T - X_0$ |

6.2 Legenda das Tabelas e Gráficos da Seção

O primeiro gráfico apresentado em cada figura relacionada aos exemplos, têm a seguinte legenda:

mreal Verdadeira $\mu(x)$ do modelo.

mveros Estimativa de $\mu(x)$ utilizando máxima verossimilhança.

Mp1,hat Estimativa de $\mu(x)$ utilizando a janela do Estágio 1.

Mp2,hat Estimativa de $\mu(x)$ utilizando a janela do Estágio 2.

O segundo gráfico apresentado em cada figura relacionada aos exemplos, têm a seguinte legenda:

Dif Real Verdadeiro $b(x)$ do modelo.

Dif1 Estimativa de $b(x)$ utilizando a janela do Estágio 1.

Dif2 Estimativa de $b(x)$ utilizando a janela do Estágio 2.

As tabelas têm a seguinte codificação:

EQMV Erro Quadrático Médio do coeficiente de tendência utilizando o estimador de máxima verossimilhança.

EQM1 Erro Quadrático Médio do coeficiente de tendência utilizando o estimador não-paramétrico com a janela do Estágio 1.

EQM2 Erro Quadrático Médio do coeficiente de tendência utilizando o estimador não-paramétrico com a janela do Estágio 2.

EQM1D Erro Quadrático Médio do coeficiente de difusão utilizando o estimador não-paramétrico com a janela do Estágio 1.

EQM2D Erro Quadrático Médio do coeficiente de difusão utilizando o estimador não-paramétrico com a janela do Estágio 2.

EQIV Erro Quadrático Integrado do coeficiente de tendência utilizando o estimador de máxima verossimilhança.

EQI1 Erro Quadrático Integrado do coeficiente de tendência utilizando o estimador não-paramétrico com a janela do Estágio 1.

EQI2 Erro Quadrático Integrado do coeficiente de tendência utilizando o estimador não-paramétrico com a janela do Estágio 2.

EQI1D Erro Quadrático Integrado do coeficiente de difusão utilizando o estimador não-paramétrico com a janela do Estágio 1.

EQI2D Erro Quadrático Integrado do coeficiente de difusão utilizando o estimador não-paramétrico com a janela do Estágio 2.

6.3 Modelo Cox-Ingersoll-Ross (CIR)

Cox *et al.* (1985) propuseram o seguinte modelo:

$$dX_t = \kappa(\nu - X_t)dt + \sigma\sqrt{X_t}dB_t, \quad \text{para } \kappa, \nu, \sigma > 0. \quad (6.3)$$

Pode-se notar que o modelo acima é o processo de difusão mostrado na equação (1.1) com

$$\begin{aligned} \mu(X_t) &= \kappa(\nu - X_t) \\ b(X_t) &= \sigma\sqrt{X_t}. \end{aligned}$$

Considerando o modelo proposto, os casos simulados são para valores fixos de κ ($\kappa = 0, 3; 0, 8$), de T ($T = 30, 48, 60$), em meses, de ν ($\nu = 0, 1$) e do valor inicial do processo $X_0 = 0, 1$. Os valores de T foram adotados pensando em situações realistas e o valor de σ fixado em 0,06. O valor de Δ é de $1/21 = 0.04761905$ pensando-se em observações diárias, com juros mensais. O método de discretização utilizado foi o de Euler descrito no Capítulo 2.

Para os seis casos simulados apresentamos os gráficos referentes às estimativas do coeficiente de tendência e referentes às estimativas do coeficiente de difusão quando pegamos uma realização do processo. O primeiro gráfico de cada figura nos mostra a estimativa do coeficiente de difusão quando utilizamos o método de máxima verossimilhança proposto por Cleur e Manfredi (1999) e quando utilizamos o método não-paramétrico de acordo com Fan e Gijbels (1995). Neste último método consideramos o caso somente com a janela piloto (Estágio 1) e também o caso em que a janela é estimada no Estágio 2 do método. Os comentários se encontram no final do último caso simulado.

Na subseção 6.3.8, após a simulação de 100 realizações para cada caso, se encontra a tabela com os Erros definidos na seção 6.2 para cada um dos métodos utilizados para a estimação de ambos os coeficientes.

6.3.1 $\kappa = 0,3$ e $T = 30$

As estimativas das funções $\mu(X_t)$ e $b(X_t)$ para uma realização do processo podem ser vistas na Figura 6.1.

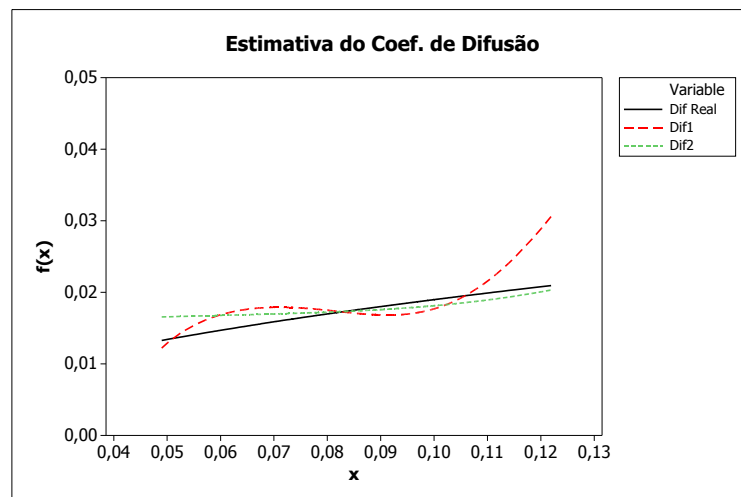
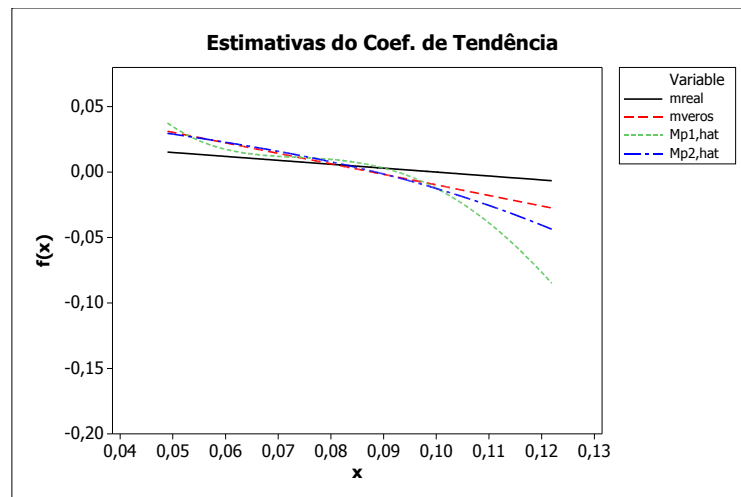


Figura 6.1: Exemplo CIR $\kappa = 0,3$ e $T = 30$.

6.3.2 $\kappa = 0,8$ e $T = 30$

As estimativas das funções $\mu(X_t)$ e $b(X_t)$ para uma realização do processo podem ser vistas na Figura 6.2.

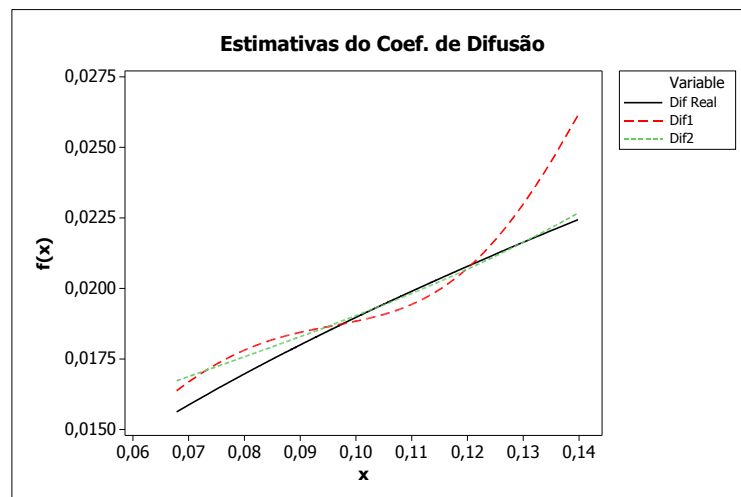
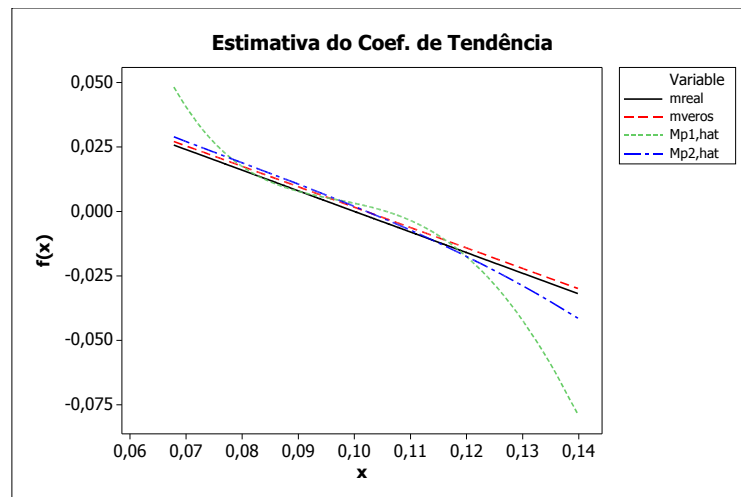


Figura 6.2: Exemplo CIR $\kappa = 0,8$ e $T = 30$.

6.3.3 $\kappa = 0,3$ e $T = 48$

As estimativas das funções $\mu(X_t)$ e $b(X_t)$ para uma realização do processo podem ser vistas na Figura 6.3.

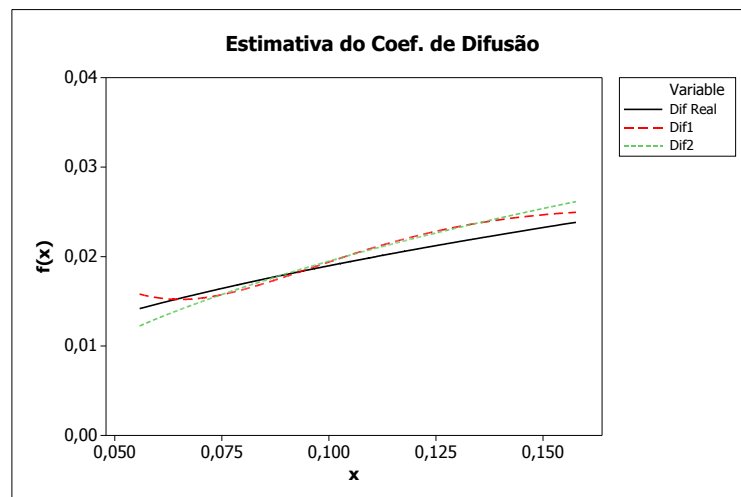
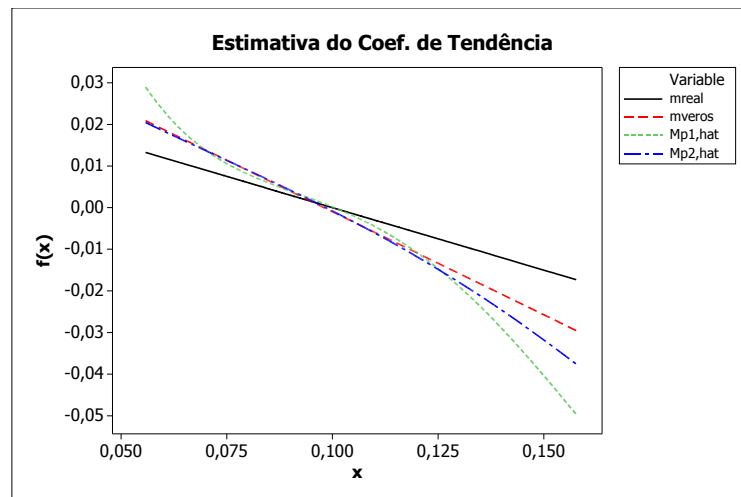


Figura 6.3: Exemplo CIR $\kappa = 0,3$ e $T = 48$.

6.3.4 $\kappa = 0,8$ e $T = 48$

As estimativas das funções $\mu(X_t)$ e $b(X_t)$ para uma realização do processo podem ser vistas na Figura 6.4.

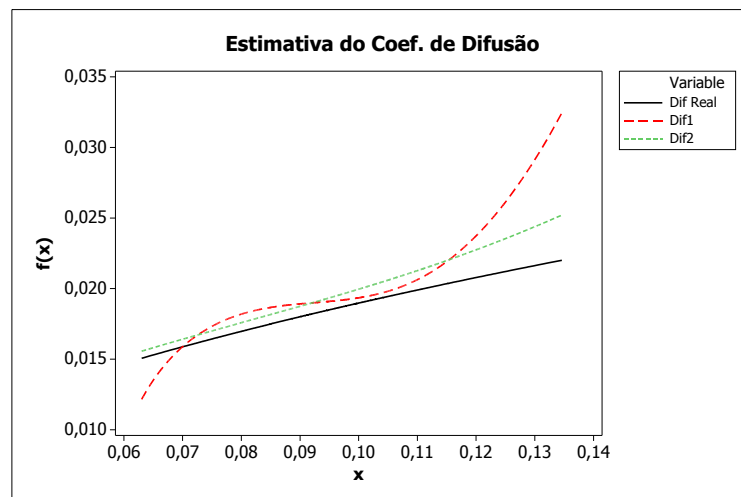
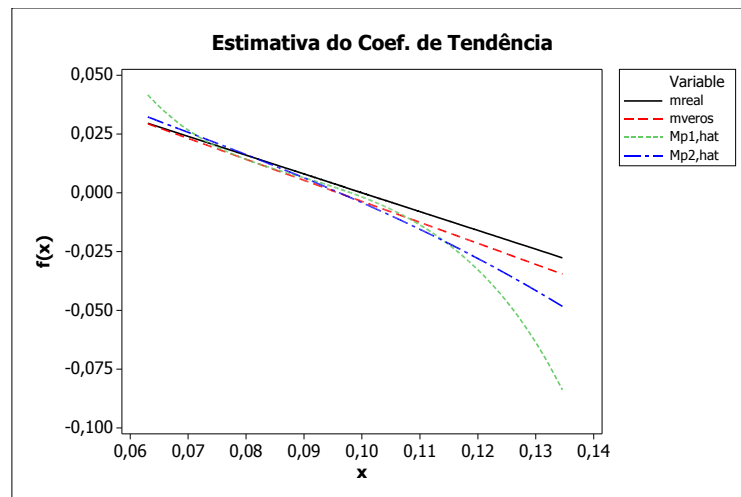


Figura 6.4: Exemplo CIR $\kappa = 0,8$ e $T = 48$.

6.3.5 $\kappa = 0,3$ e $T = 60$

As estimativas das funções $\mu(X_t)$ e $b(X_t)$ para uma realização do processo podem ser vistas na Figura 6.5.

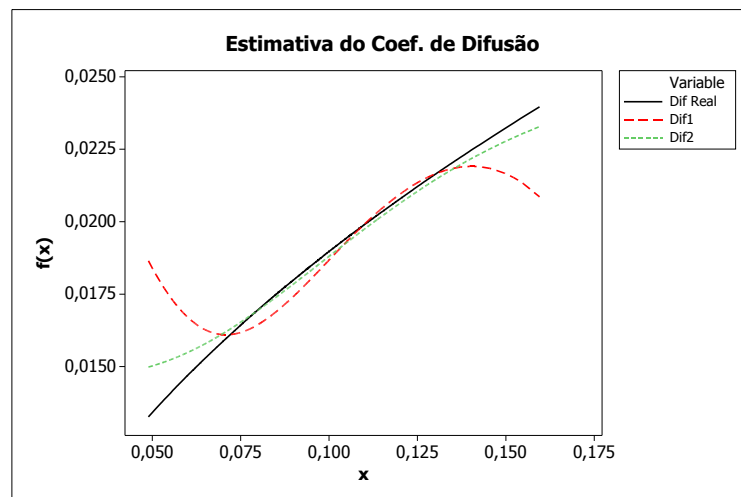
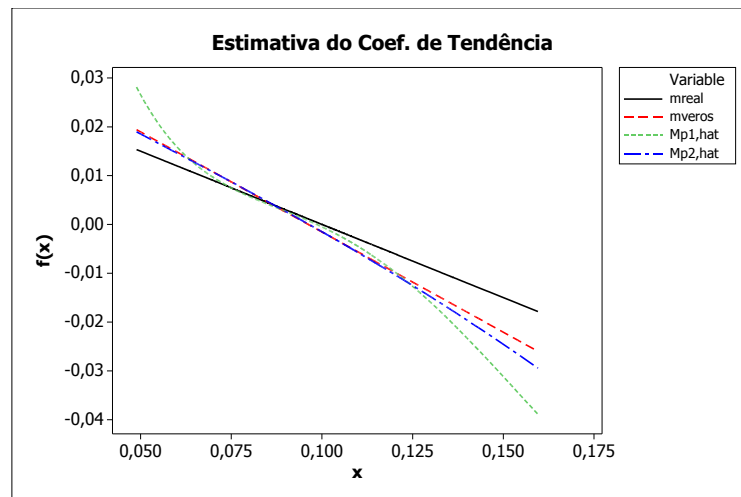


Figura 6.5: Exemplo CIR $\kappa = 0,3$ e $T = 60$.

6.3.6 $\kappa = 0,8$ e $T = 60$

As estimativas das funções $\mu(X_t)$ e $b(X_t)$ para uma realização do processo podem ser vistas na Figura 6.6.

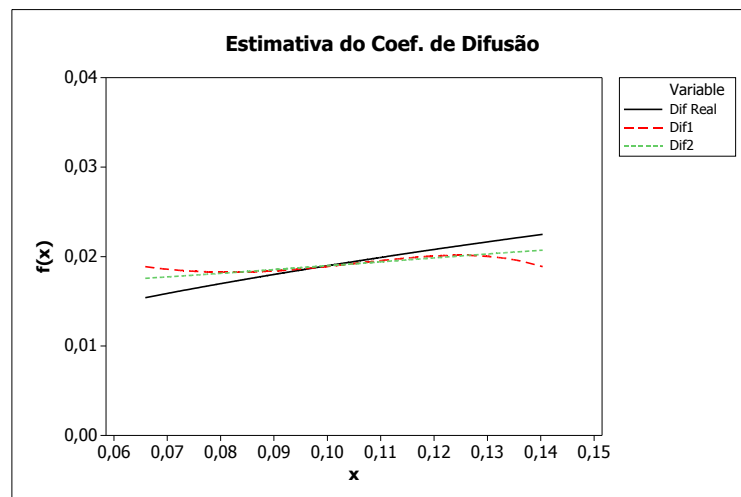
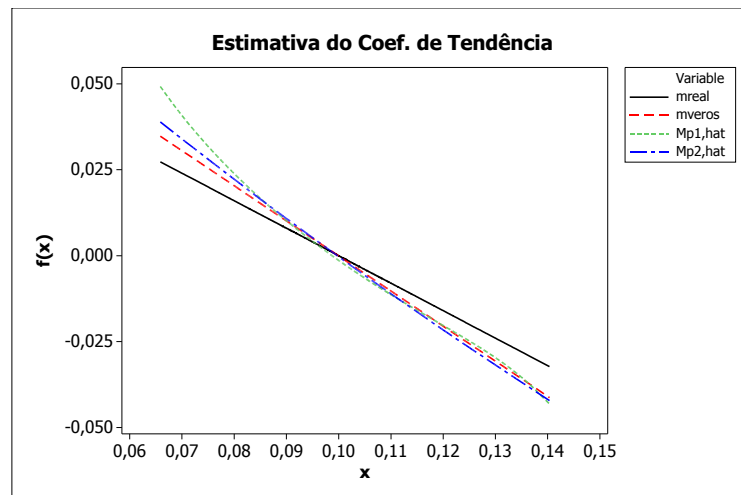


Figura 6.6: Exemplo CIR $\kappa = 0,8$ e $T = 60$.

6.3.7 Comentários

- Ao observar as figuras de 6.1 a 6.6, referentes ao modelo CIR, podemos notar que a estimativa do coeficiente de tendência quando utilizamos a janela do segundo estágio é mais próxima da curva teórica do que quando utilizamos a janela piloto.
- A estimativa por máxima verossimilhança apresentou melhores resultados nos casos considerados nas Figuras 6.1, 6.2, 6.3 e 6.4. Este resultado já era de se esperar pois estamos utilizando uma função linear no coeficiente. Apesar disso, a diferença não é tão grande entre o método de máxima verossimilhança e a estimativa produzida pelo método não-paramétrico quando utilizamos a janela do segundo estágio.
- Nos casos em que $T = 60$, observamos nas Figuras 6.5 e 6.6, que a estimativa do coeficiente de tendência utilizando o método de máxima verossimilhança e o método não-paramétrico são praticamente idênticas.
- Em relação ao coeficiente de difusão b , a curva estimada considerando-se a janela do segundo estágio, aproxima-se da curva teórica em todos os casos considerados do modelo CIR. Podemos observar este fato, considerando o segundo gráfico nas Figuras 6.1, 6.2, 6.3, 6.4, 6.5 e 6.6

Nas figuras apresentadas para o modelo CIR utilizamos somente uma realização do processo. Dessa maneira não é possível chegar a conclusões mais precisas. Após as simulações utilizando 100 realizações, apresentamos na próxima seção as tabelas referentes aos erros de estimativa de cada método e para todos os casos considerados para o modelo CIR. O objetivo é verificar se o método realmente fornece boas estimativas.

Tabela 6.9: Erros $\kappa = 0,8$ e $T = 60$ - Modelo CIR.

| | Média | DP | Mínimo | Q1 | Mediana | Q3 | Máximo |
|-------|----------|----------|----------|----------|----------|----------|----------|
| EQMV | 1,18E-05 | 1,13E-05 | 0,00E+00 | 4,15E-06 | 6,80E-06 | 1,82E-05 | 5,77E-05 |
| EQM1 | 2,51E-05 | 1,78E-05 | 8,00E-07 | 1,23E-05 | 2,04E-05 | 3,32E-05 | 8,84E-05 |
| EQM2 | 1,45E-05 | 1,32E-05 | 1,00E-07 | 5,20E-06 | 9,45E-06 | 2,15E-05 | 6,12E-05 |
| EQM1D | 7,59E-07 | 6,81E-07 | 1,00E-07 | 3,00E-07 | 5,00E-07 | 9,00E-07 | 4,10E-06 |
| EQM2D | 3,89E-07 | 4,75E-07 | 0,00E+00 | 1,00E-07 | 2,00E-07 | 5,00E-07 | 2,60E-06 |
| EQIV | 0,00E+00 | 0,00E+00 | 0,00E+00 | 0,00E+00 | 0,00E+00 | 0,00E+00 | 0,00E+00 |
| EQI1 | 0,00E+00 | 0,00E+00 | 0,00E+00 | 0,00E+00 | 0,00E+00 | 0,00E+00 | 0,00E+00 |
| EQI2 | 0,00E+00 | 0,00E+00 | 0,00E+00 | 0,00E+00 | 0,00E+00 | 0,00E+00 | 0,00E+00 |
| EQI1D | 0,00E+00 | 0,00E+00 | 0,00E+00 | 0,00E+00 | 0,00E+00 | 0,00E+00 | 0,00E+00 |
| EQI2D | 0,00E+00 | 0,00E+00 | 0,00E+00 | 0,00E+00 | 0,00E+00 | 0,00E+00 | 0,00E+00 |

A diferença entre os EQM no método de máxima verossimilhança e no método não-paramétrico só pode ser observada na quinta casa decimal. Isso indica que além do erro ser pequeno o método não-paramétrico não apresentou grande diferença do método de máxima verossimilhança, mesmo quando o coeficiente de tendência é linear. Além disso os Erros Quadráticos Integrados também são extremamente pequenos no método não-paramétrico. Sendo assim, apesar do ajuste por máxima verossimilhança ser melhor quando temos o modelo CIR, o método não-paramétrico fornece boas estimativas para o coeficiente de tendência.

O método de estimação utilizado por Cleur e Manfredi (1999) é baseado somente na estimação do coeficiente de tendência. Dessa forma, apresentamos para os erros relacionados ao coeficiente de difusão somente quando foi utilizado o método não-paramétrico. Como podemos observar, os valores dos EQM e do EQI foram bem pequenos, indicando uma proximidade muito grande entre a curva teórica e a curva estimada, principalmente quando utilizamos a janela do Estágio 2.

Outra consideração importante é que a variabilidade dos EQM foi menor para o método paramétrico quando comparamos com o método não-paramétrico, para as estimativas do coeficiente de tendência. Para as estimativas do coeficiente de difusão o método não-paramétrico utilizado com a janela do estágio 2, apresentou menor variabilidade em todos os casos considerados.

6.3.9 Uma Comparação entre o Esquema de Euler e o de Milstein

Consideremos o modelo CIR com o coeficiente de difusão

$$b(X_t) = \sigma\sqrt{X_t}.$$

Utilizando o esquema de Euler, temos a seguinte discretização

$$X_{t+1} = X_t + \kappa(\nu - X_t)\Delta + \sigma^2\sqrt{X_t}(B_{t+1} - B_t) \quad (6.4)$$

e utilizando o esquema de Milstein temos

$$X_{t+1} = X_t + \kappa(\nu - X_t)\Delta + \sigma\sqrt{X_t}(B_{t+1} - B_t) + \frac{\sigma}{4}((B_{t+1} - B_t)^2 - \Delta). \quad (6.5)$$

Observe que o último termo do lado direito de (6.5) é um número que na maioria das vezes é pequeno demais para que realmente exista uma diferença entre os dois esquemas de discretização, considerando os parâmetros escolhidos para simulação. Para ilustrar, voltemos ao caso com $\kappa = 0,3$ e $T = 30$. A Figura 6.7 mostra como fica a estimativa dos coeficientes considerando os dois esquemas de discretização.

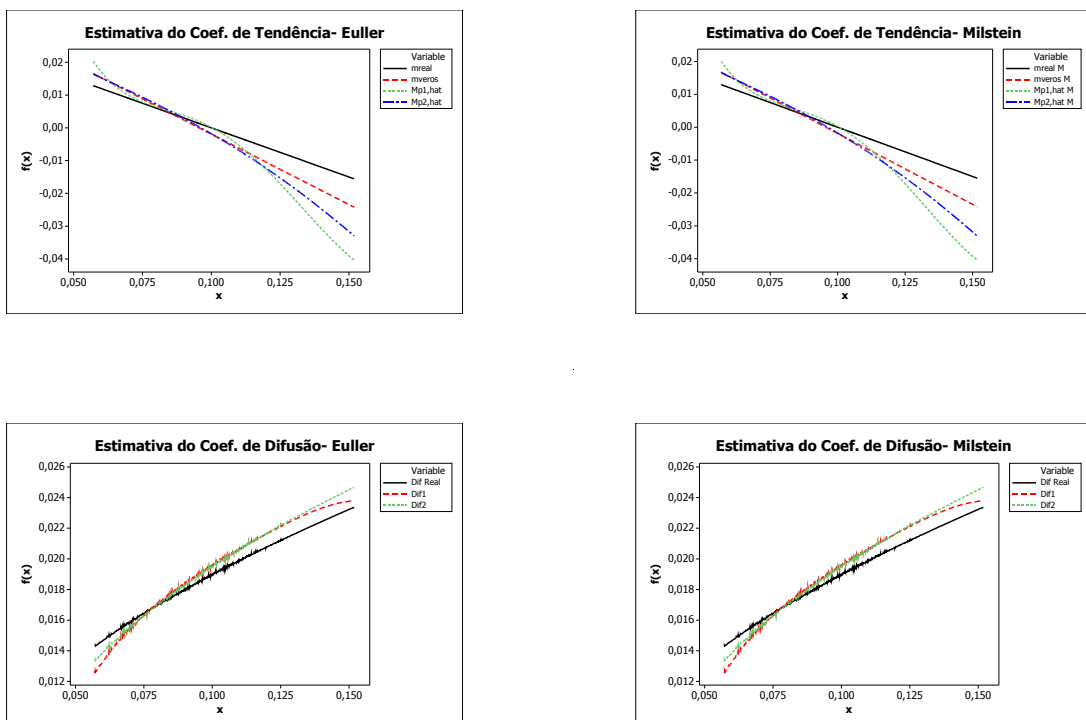


Figura 6.7: Exemplo CIR Utilizando os Dois Esquemas de Discretização.

6.4 Modelo de Brennan-Schwartz

O modelo de Brennan-Schwartz é descrito pela seguinte equação diferencial estocástica

$$dX_t = \kappa(\nu - X_t)dt + \sigma X_t dB_t, \quad \text{para } \kappa, \nu, \sigma > 0. \quad (6.6)$$

O modelo acima também é o processo de difusão mostrado na equação (1.1) com

$$\mu(X_t) = \kappa(\nu - X_t)$$

$$b(X_t) = \sigma X_t.$$

Analogamente, considerando o modelo proposto, os casos simulados são para valores fixos de κ ($\kappa = 0, 3; 0, 8$), de T ($T = 30, 48, 60$), em meses, de ν ($\nu = 0, 1$) e do valor inicial do processo $X_0 = 0, 1$. Os valores de T foram adotados pensando em situações realistas e o valor de σ fixado em 0,06. O valor de delta é de $1/21 = 0.04761905$ pensando-se em observações diárias, com juros mensais. O método de discretização utilizado foi o de Euler descrito no capítulo 2.

Para os seis casos simulados apresentamos os gráficos referentes às estimativas do coeficiente de tendência e referentes às estimativas do coeficiente de difusão quando pegamos uma realização do processo. O primeiro gráfico de cada figura nos mostra a estimativa do coeficiente de difusão quando utilizamos o método de máxima verossimulhança proposto pelos autores Cleur e Manfredi (1999), quando utilizamos o método não-paramétrico de acordo com Fan e Gijbels (1995). Neste último método consideramos o caso somente com a janela piloto (Estágio 1) e também o caso em que a janela é estimada no Estágio 2 do método. Os comentários se encontram no final do último caso simulado.

Na subseção 6.4.8 após a simulação de 100 realizações do processo para cada caso, se encontra a tabela com os Erros definidos na seção 6.2 para cada um dos métodos utilizados para a estimação de ambos os coeficientes.

6.4.1 $\kappa = 0,3$ e $T = 30$

As estimativas das funções $\mu(X_t)$ e $b(X_t)$ para uma realização do processo podem ser vistas na Figura 6.8.

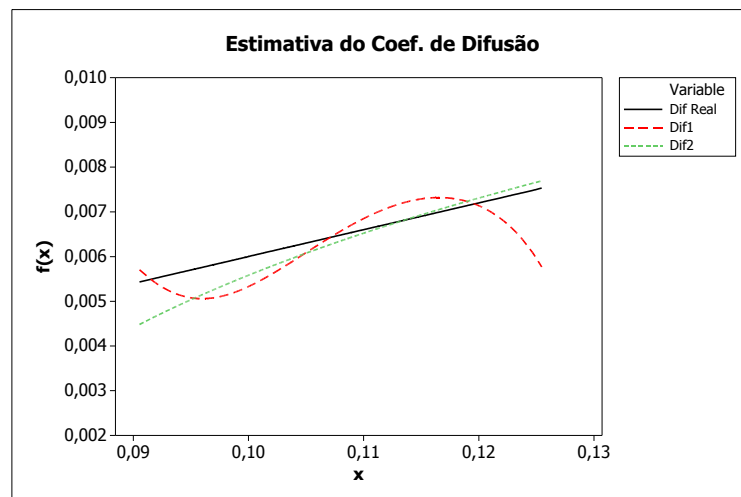
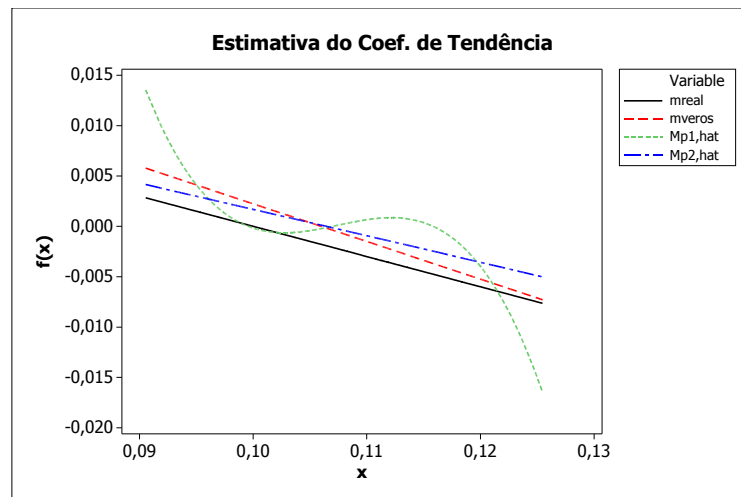


Figura 6.8: Exemplo Brennan-Schwartz $\kappa = 0,3$ e $T = 30$.

6.4.2 $\kappa = 0,8$ e $T = 30$

As estimativas das funções $\mu(X_t)$ e $b(X_t)$ para uma realização do processo podem ser vistas na Figura 6.9

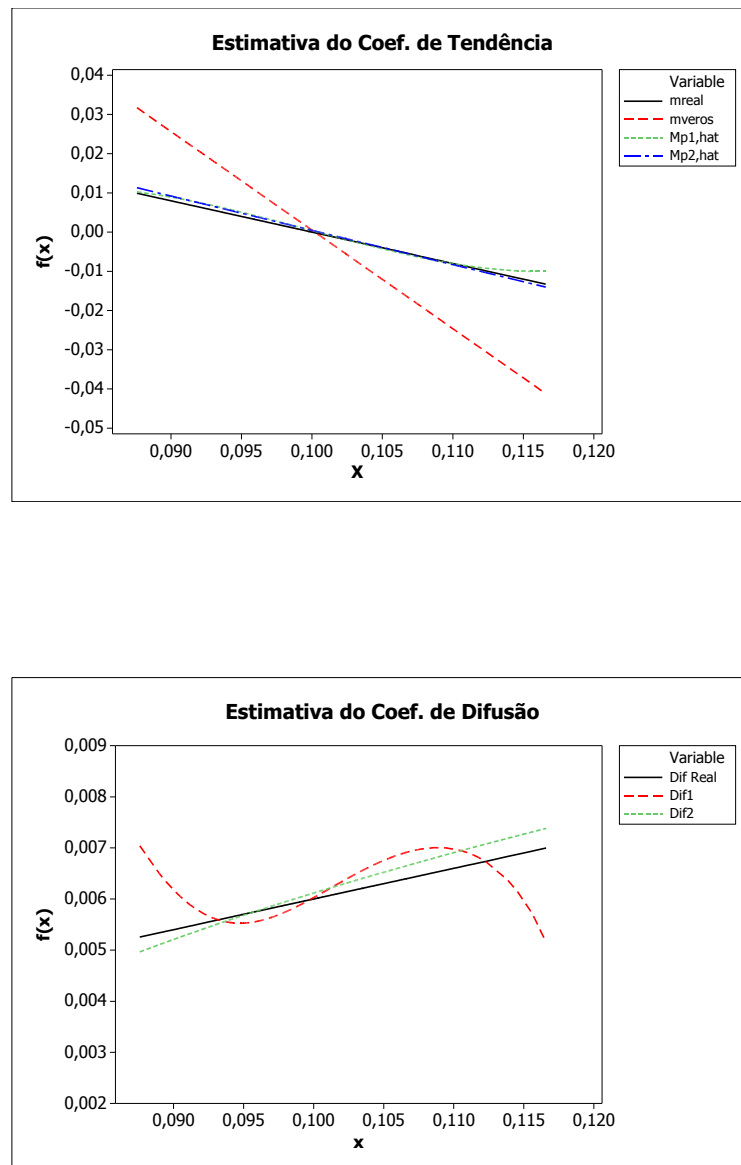


Figura 6.9: Exemplo Brennan-Schwartz $\kappa = 0,8$ e $T = 30$.

6.4.3 $\kappa = 0,3$ e $T = 48$

As estimativas das funções $\mu(X_t)$ e $b(X_t)$ para uma realização do processo podem ser vistas na Figura 6.10

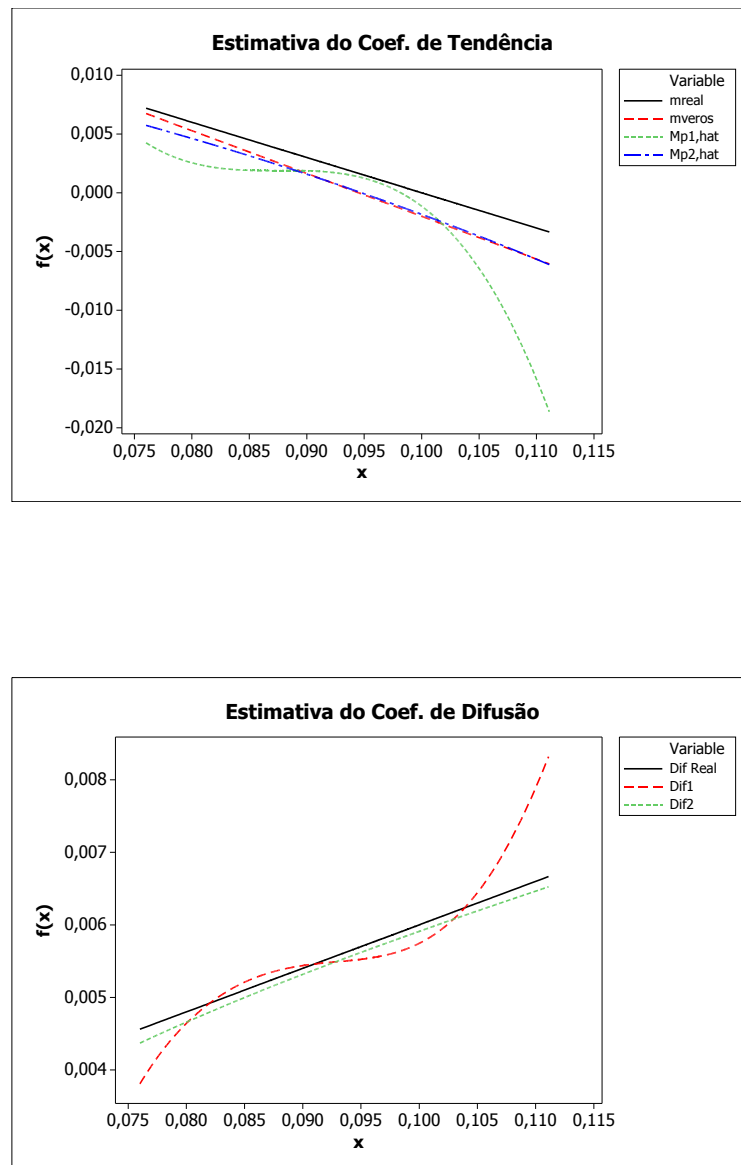


Figura 6.10: Exemplo Brennan-Schwartz $\kappa = 0,3$ e $T = 48$.

6.4.4 $\kappa = 0,8$ e $T = 48$

As estimativas das funções $\mu(X_t)$ e $b(X_t)$ para uma realização do processo podem ser vistas na Figura 6.11

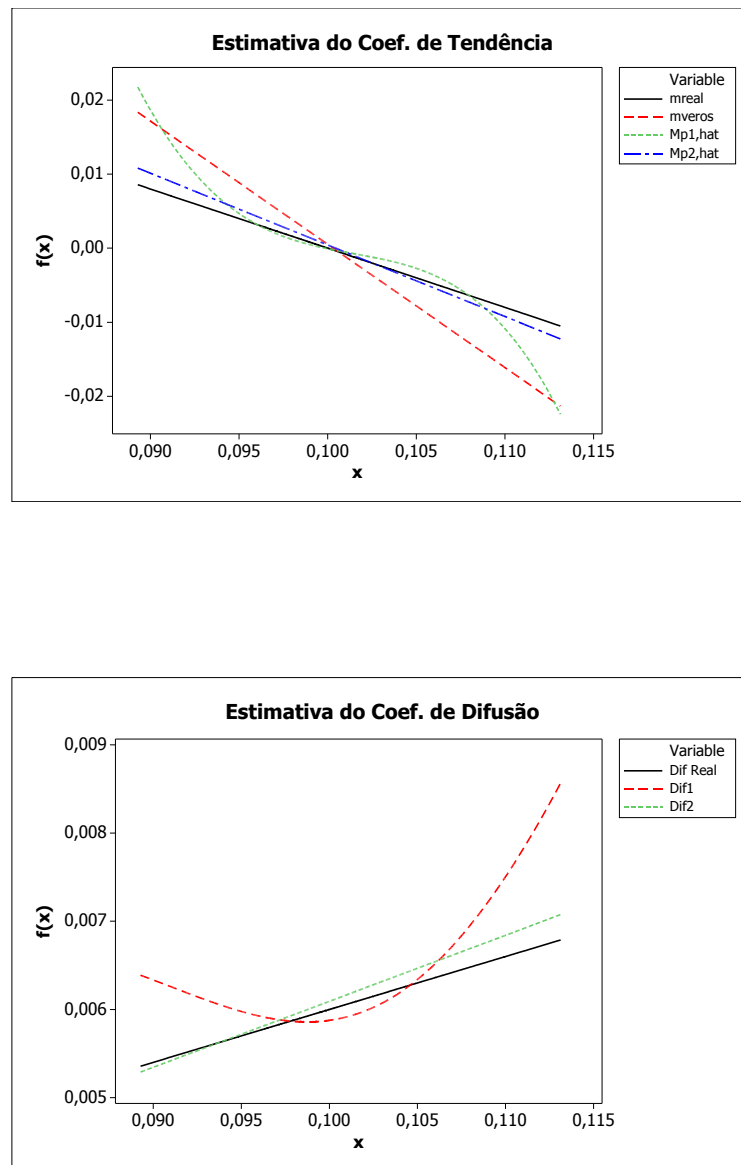


Figura 6.11: Exemplo Brennan-Schwartz $\kappa = 0,8$ e $T = 48$.

6.4.5 $\kappa = 0,3$ e $T = 60$

As estimativas das funções $\mu(X_t)$ e $b(X_t)$ para uma realização do processo podem ser vistas na Figura 6.12

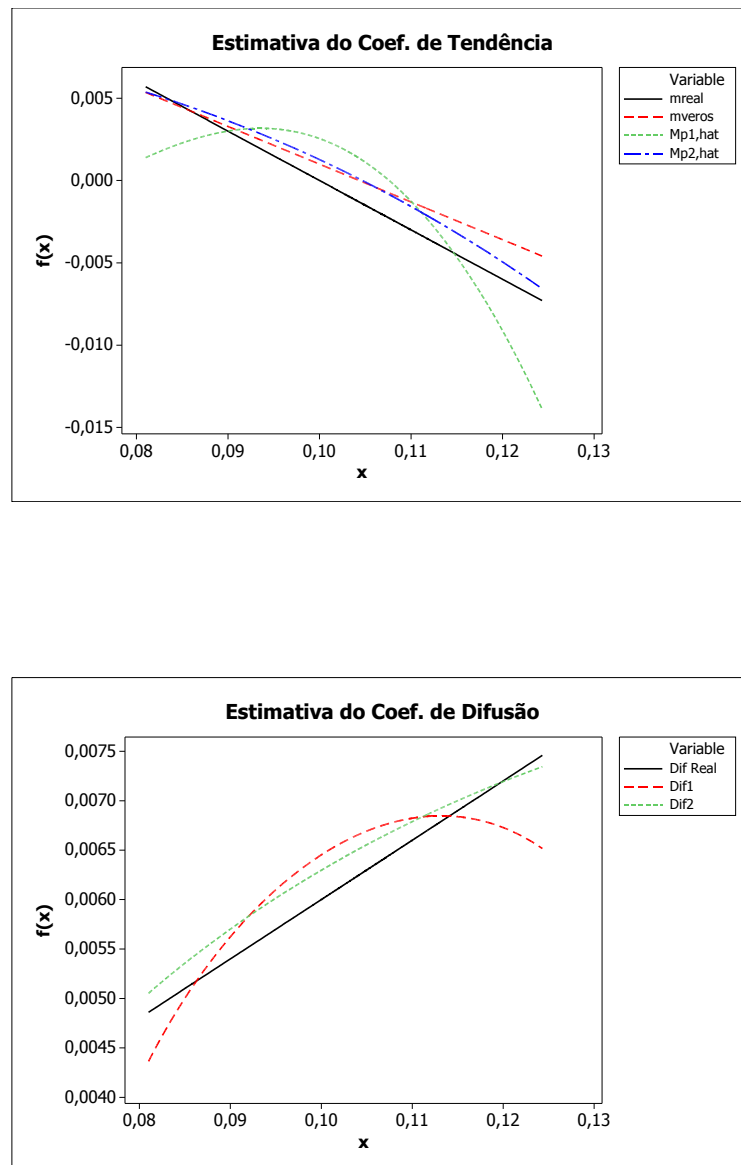


Figura 6.12: Exemplo Brennan-Schwartz $\kappa = 0,3$ e $T = 60$.

6.4.6 $\kappa = 0,8$ e $T = 60$

As estimativas das funções $\mu(X_t)$ e $b(X_t)$ para uma realização do processo podem ser vistas na Figura 6.13.

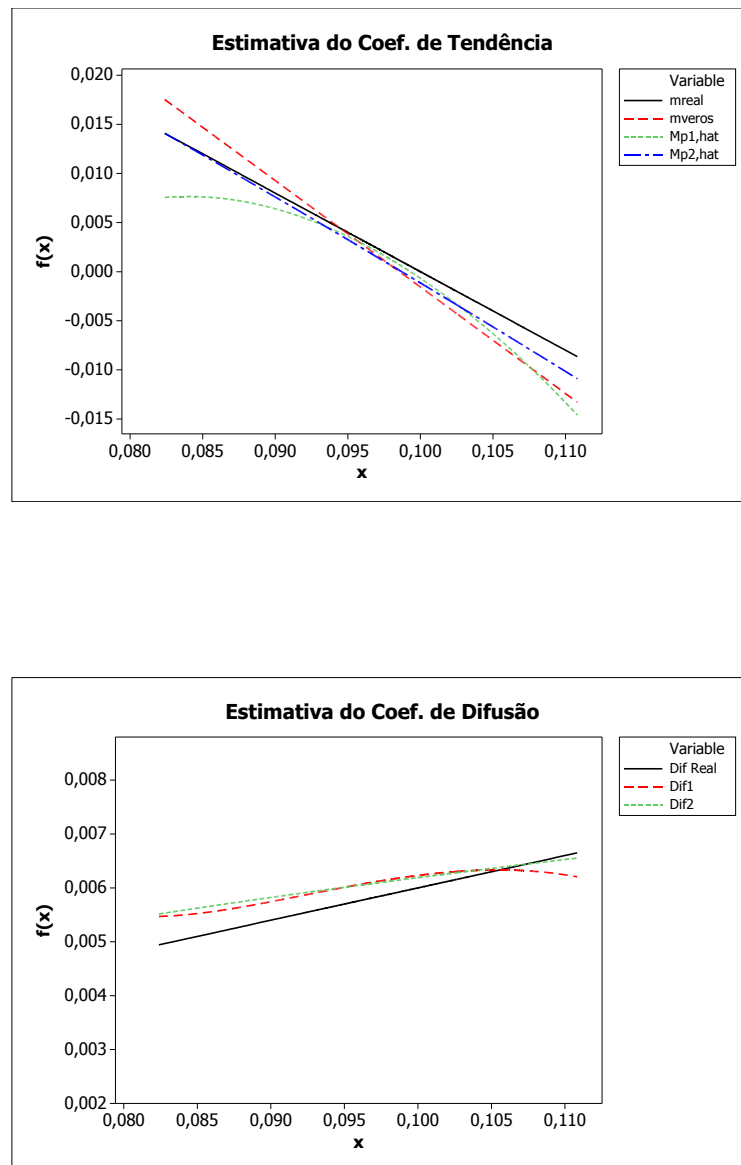


Figura 6.13: Exemplo Brennan-Schwartz $\kappa = 0,8$ e $T = 60$.

6.4.7 Comentários

- No primeiro gráfico da Figura 6.8 podemos observar que a estimativa do coeficiente de tendência utilizando o método de máxima verossimilhança apresentou pouca diferença em relação ao método não-paramétrico, apesar de sua curva estimada estar mais próxima da curva teórica. Porém, observamos nas figuras 6.9, 6.10, 6.11, 6.12 e 6.13, que o método não-paramétrico quando utilizamos a janela do Estágio 2 funciona melhor para estimar o coeficiente $\mu(\cdot)$ para os exemplos mostrados.
- Um fato interessante de ser observado é que mesmo se tratando de um coeficiente de tendência linear, pois estamos considerando o modelo de Brennan-Schwartz, a estimativa pelo método não-paramétrico apresentou melhores resultados na maioria dos casos considerados.
- Em relação ao coeficiente de difusão b , que pode ser observado no segundo gráfico das figuras, a curva estimada considerando-se a janela do segundo estágio, aproxima-se da curva teórica em todos os casos considerados do modelo Brennan-Schwartz.
- Considerando-se os casos simulados, observamos que a utilização da janela do segundo estágio, melhora consideravelmente as estimativas das funções $\mu(\cdot)$ e $b(\cdot)$.

Nas figuras apresentadas para o modelo Brennan-Schwartz também utilizamos somente uma realização do processo. Como não é possível chegar a conclusões mais precisas, apresentamos na próxima seção as tabelas referentes aos erros de estimativa de cada método e para todos os casos considerados para o modelo Brennan-Schwartz, após a simulações com 100 realizações.

Tabela 6.15: Erros $\kappa = 0,8$ e $T = 60$ - Modelo Brennan-Schwartz.

| | Média | DP | Mínimo | Q1 | Mediana | Q3 | Máximo |
|-------|----------|----------|----------|----------|----------|----------|----------|
| EQMV | 6,28E-06 | 6,13E-06 | 1,00E-07 | 1,68E-06 | 4,40E-06 | 8,80E-06 | 2,87E-05 |
| EQM1 | 2,35E-06 | 1,63E-06 | 0,00E+00 | 1,30E-06 | 2,00E-06 | 3,00E-06 | 1,04E-05 |
| EQM2 | 1,34E-06 | 1,18E-06 | 0,00E+00 | 4,00E-07 | 1,00E-06 | 1,90E-06 | 5,80E-06 |
| EQM1D | 7,00E-08 | 6,89E-08 | 0,00E+00 | 0,00E+00 | 1,00E-07 | 1,00E-07 | 3,00E-07 |
| EQM2D | 3,00E-08 | 5,41E-08 | 0,00E+00 | 0,00E+00 | 0,00E+00 | 1,00E-07 | 3,00E-07 |
| EQIV | 0,00E+00 | 0,00E+00 | 0,00E+00 | 0,00E+00 | 0,00E+00 | 0,00E+00 | 0,00E+00 |
| EQI1 | 0,00E+00 | 0,00E+00 | 0,00E+00 | 0,00E+00 | 0,00E+00 | 0,00E+00 | 0,00E+00 |
| EQI2 | 0,00E+00 | 0,00E+00 | 0,00E+00 | 0,00E+00 | 0,00E+00 | 0,00E+00 | 0,00E+00 |
| EQI1D | 0,00E+00 | 0,00E+00 | 0,00E+00 | 0,00E+00 | 0,00E+00 | 0,00E+00 | 0,00E+00 |
| EQI2D | 0,00E+00 | 0,00E+00 | 0,00E+00 | 0,00E+00 | 0,00E+00 | 0,00E+00 | 0,00E+00 |

O Erro Quadrático Médio para a estimativa do coeficiente de tendência foi, em média, maior quando utilizamos o método de máxima verossimilhança do que quando utilizamos o método não-paramétrico, independente da janela utilizada, exceto para o caso na Tabela 6.14. Neste caso, somente com a janela mais refinada obtivemos um ajuste melhor do método não-paramétrico do que do método de máxima verossimilhança utilizado.

Os Erros Quadráticos Integrados para o coeficiente de tendência apresentaram valores extremamente pequenos, o que ajuda a reforçar a idéia de que o método não-paramétrico fornece boas estimativas para o coeficiente.

Quanto ao coeficiente de difusão, os valores dos EQM foram bem pequenos indicando uma proximidade muito grande entre a curva teórica e a curva estimada, principalmente quando utilizamos a janela do Estágio 2.

Outra consideração importante é que a variabilidade dos EQM foi menor para o método não-paramétrico com a janela do estágio 2 quando comparamos com o método paramétrico, para as estimativas do coeficiente de tendência. Para as estimativas do coeficiente de difusão o método não-paramétrico utilizado com a janela do estágio 2, apresentou menor variabilidade em todos os casos considerados.

Os Erros Quadráticos Integrados para as estimativas tanto do coeficiente de tendência como o de difusão foram extremamente pequenos, considerando um arredondamento acima da décima casa decimal.

6.5 Modelo Heuristically Derived Logistic

O modelo de Heuristically Derived Logistic é descrito pela seguinte equação diferencial estocástica

$$dX_t = \kappa(\nu - X_t)X_t dt + \sigma X_t dB_t, \quad \text{para } \kappa, \nu, \sigma > 0. \quad (6.7)$$

O modelo acima também é o processo de difusão mostrado na equação (1.1) com

$$\mu(X_t) = \kappa(\nu - X_t)X_t$$

$$b(X_t) = \sigma X_t.$$

Analogamente, $\kappa(\kappa = 0, 3; 0, 8)$, de $T(T = 30, 48, 60)$, em meses, $\nu(\nu = 0, 1)$ o valor inicial do processo $X_0 = 0, 1$. O valor de σ fixado em $0, 06$. O valor de Δ é de $1/21 = 0.04761905$ pensando-se em observações diárias, com juros mensais.

Para os seis casos simulados apresentamos os gráficos referentes às estimativas do coeficiente de tendência e referentes às estimativas do coeficiente de difusão quando observamos uma realização do processo. O primeiro gráfico de cada figura nos mostra a estimativa do coeficiente de difusão quando utilizamos o método de máxima verossimulhança proposto por Cleur e Manfredi (1999) e quando utilizamos o método não-paramétrico de acordo com Fan e Gijbels (1995). Neste último método consideramos o caso com a janela piloto (Estágio 1) e também o caso em que a janela é estimada no Estágio 2 do método. Os comentários se encontram no final do último caso simulado.

Na subseção 6.5.8 após a simulação de 100 realizações para cada caso, se encontra a tabela com os Erros definidos na seção 6.2 para cada um dos métodos utilizados para a estimação de ambos os coeficientes.

6.5.1 $\kappa = 0,3$ e $T = 30$

As estimativas das funções $\mu(X_t)$ e $b(X_t)$ para uma realização do processo podem ser vistas na Figura 6.14.

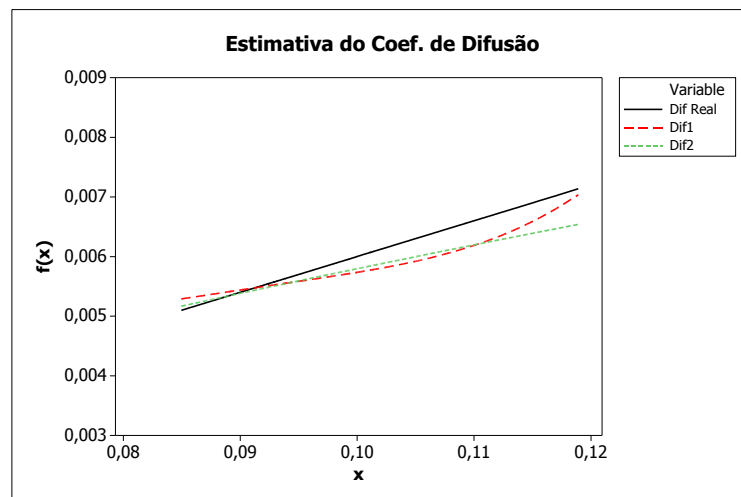
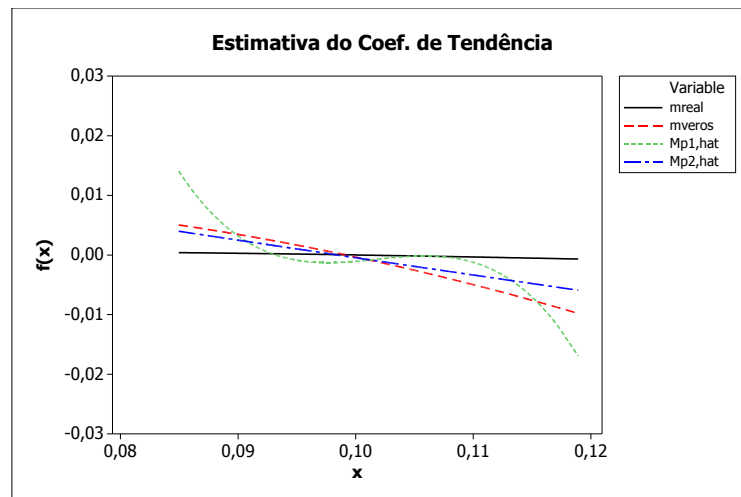


Figura 6.14: Exemplo HD Logistic $\kappa = 0,3$ e $T = 30$.

6.5.2 $\kappa = 0,8$ e $T = 30$

As estimativas das funções $\mu(X_t)$ e $b(X_t)$ para uma realização do processo podem ser vistas na Figura 6.15

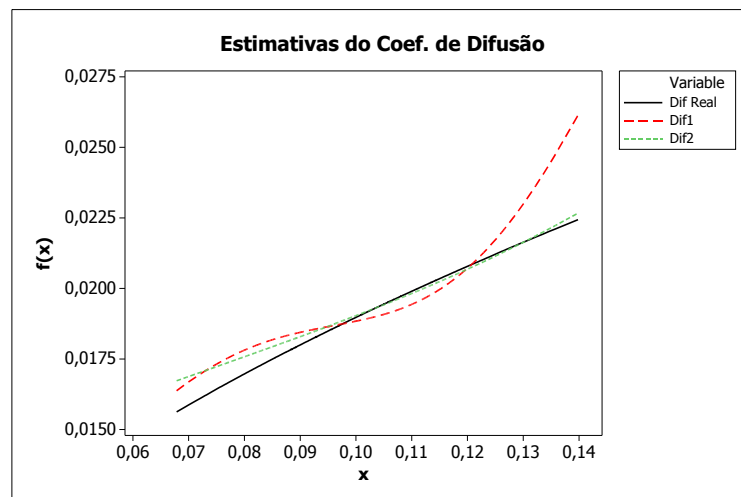
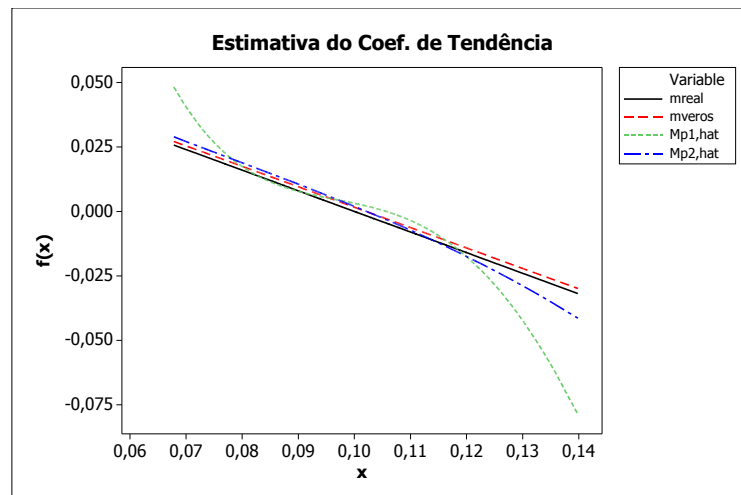


Figura 6.15: Exemplo HD Logistic $\kappa = 0,8$ e $T = 30$.

6.5.3 $\kappa = 0,3$ e $T = 48$

As estimativas das funções $\mu(X_t)$ e $b(X_t)$ para uma realização do processo podem ser vistas na Figura 6.16

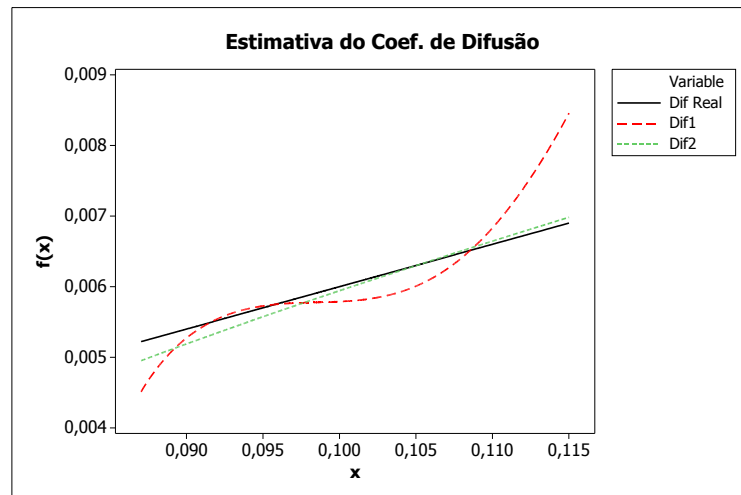
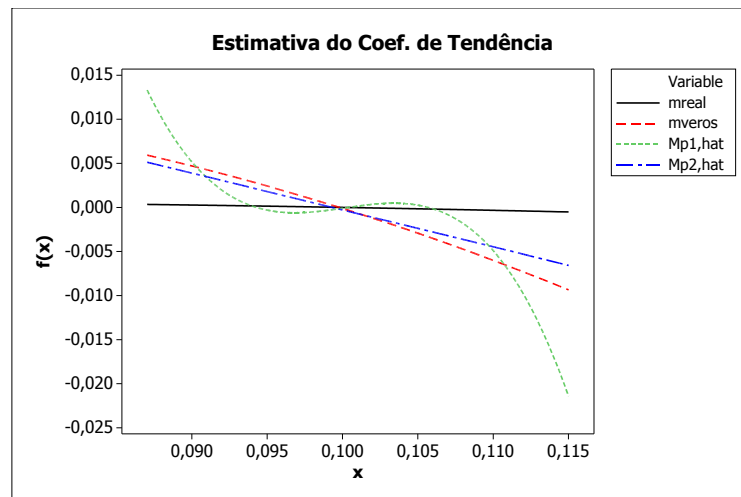


Figura 6.16: Exemplo HD Logistic $\kappa = 0,3$ e $T = 48$.

6.5.4 $\kappa = 0,8$ e $T = 48$

As estimativas das funções $\mu(X_t)$ e $b(X_t)$ para uma realização do processo podem ser vistas na Figura 6.17

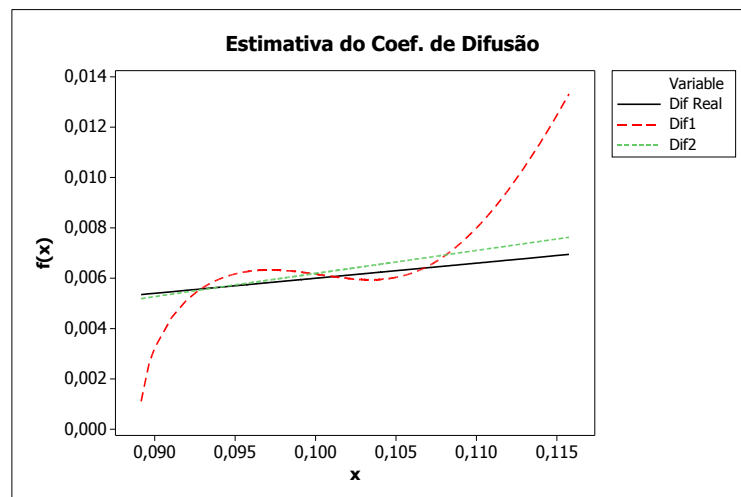
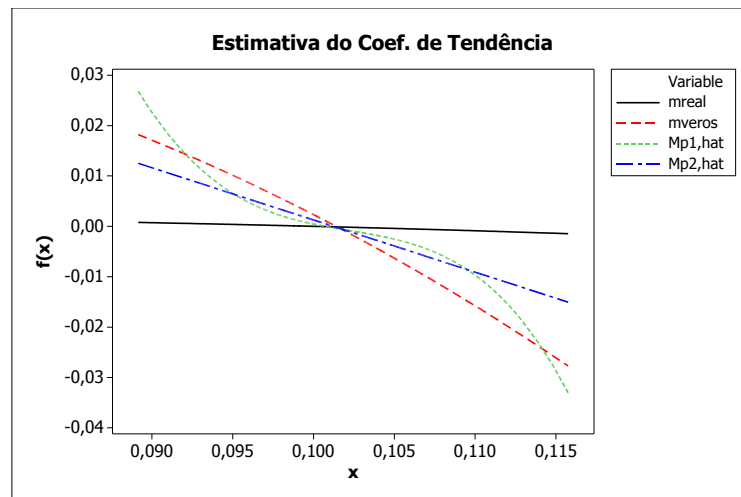


Figura 6.17: Exemplo HD Logistic $\kappa = 0,8$ e $T = 48$.

6.5.5 $\kappa = 0,3$ e $T = 60$

As estimativas das funções $\mu(X_t)$ e $b(X_t)$ para uma realização do processo podem ser vistas na Figura 6.18

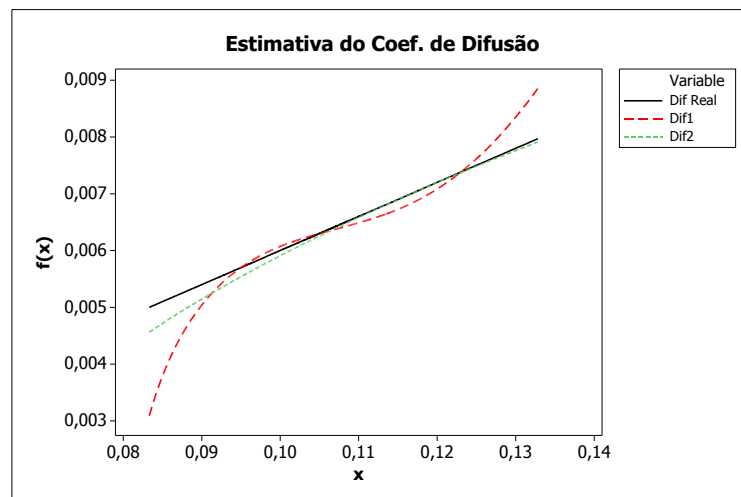
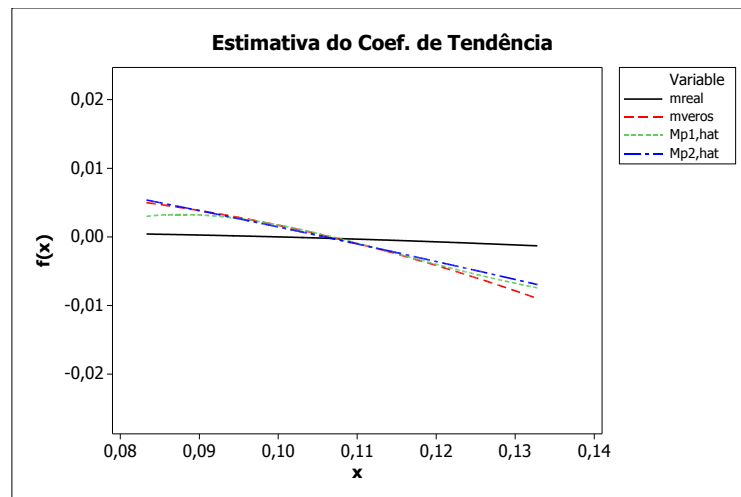


Figura 6.18: Exemplo HD Logistic $\kappa = 0,3$ e $T = 60$.

6.5.6 $\kappa = 0,8$ e $T = 60$

As estimativas das funções $\mu(X_t)$ e $b(X_t)$ para uma realização do processo podem ser vistas na Figura 6.19

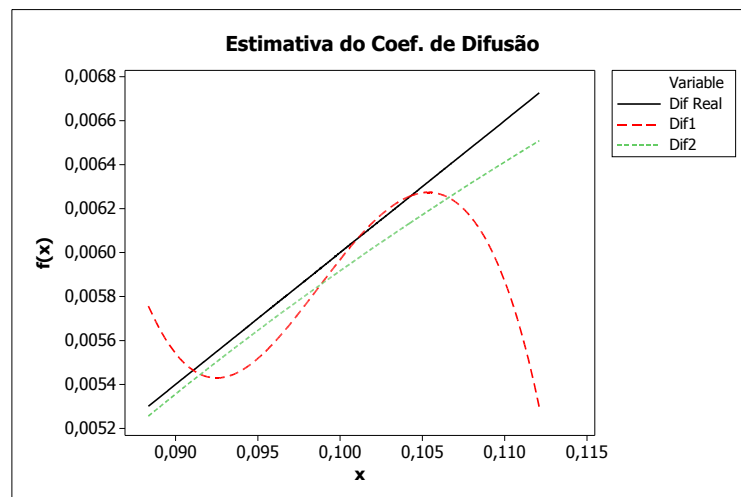
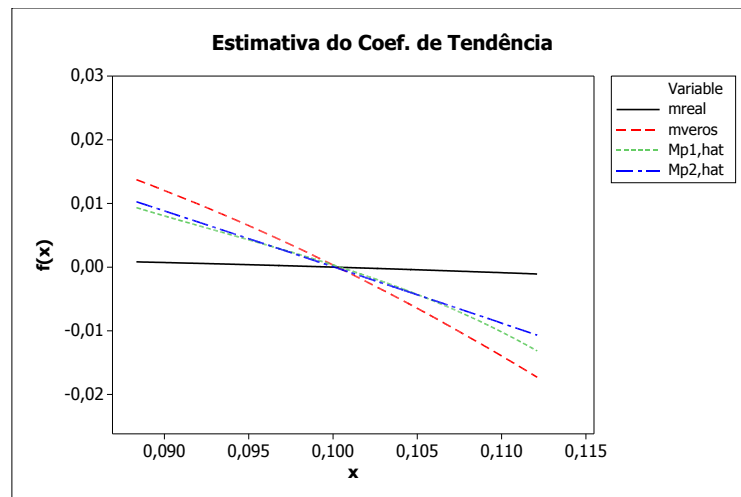


Figura 6.19: Exemplo HD Logistic $\kappa = 0,8$ e $T = 60$.

6.5.7 Comentários

- Exceto na Figura 6.15, podemos observar que a estimativa do coeficiente de tendência utilizando o método não-paramétrico está mais próxima da curva teórica do que a estimativa obtida no método de máxima verossimilhança, quando consideramos o modelo Heuristically Derived Logistic.
- Em relação ao coeficiente de difusão b , que pode ser observado no segundo gráfico das figuras, a curva estimada considerando-se a janela do segundo estágio, aproxima-se da curva teórica em todos os casos considerados do modelo Heuristically Derived Logistic.
- Considerando-se os casos simulados, observamos que a utilização da janela do segundo estágio, melhora consideravelmente as estimativas das funções $\mu(\cdot)$ e $b(\cdot)$.

Nas figuras apresentadas para o modelo Heuristically Derived Logistic também utilizamos somente uma realização do processo. Como não é possível chegar a conclusões mais precisas, apresentamos na próxima seção as tabelas referentes aos erros de estimativa de cada método e para todos os casos considerados para o modelo, após a simulações com 100 realizações.

Tabela 6.21: Erros $\kappa = 0,8$ e $T = 60$ - HD Logistic.

| | Média | DP | Mínimo | Q1 | Mediana | Q3 | Máximo |
|-------|----------|----------|----------|----------|----------|----------|----------|
| EQMV | 2,83E-05 | 1,06E-05 | 1,03E-05 | 2,08E-05 | 2,74E-05 | 3,39E-05 | 7,54E-05 |
| EQM1 | 1,46E-05 | 3,32E-06 | 7,20E-06 | 1,21E-05 | 1,43E-05 | 1,70E-05 | 2,54E-05 |
| EQM2 | 1,33E-05 | 3,17E-06 | 6,60E-06 | 1,09E-05 | 1,32E-05 | 1,52E-05 | 2,54E-05 |
| EQM1D | 6,67E-08 | 6,39E-08 | 0,00E+00 | 0,00E+00 | 1,00E-07 | 1,00E-07 | 3,00E-07 |
| EQM2D | 3,00E-08 | 5,22E-08 | 0,00E+00 | 0,00E+00 | 0,00E+00 | 1,00E-07 | 2,00E-07 |
| EQIV | 0,00E+00 | 0,00E+00 | 0,00E+00 | 0,00E+00 | 0,00E+00 | 0,00E+00 | 0,00E+00 |
| EQI1 | 0,00E+00 | 0,00E+00 | 0,00E+00 | 0,00E+00 | 0,00E+00 | 0,00E+00 | 0,00E+00 |
| EQI2 | 0,00E+00 | 0,00E+00 | 0,00E+00 | 0,00E+00 | 0,00E+00 | 0,00E+00 | 0,00E+00 |
| EQI1D | 0,00E+00 | 0,00E+00 | 0,00E+00 | 0,00E+00 | 0,00E+00 | 0,00E+00 | 0,00E+00 |
| EQI2D | 0,00E+00 | 0,00E+00 | 0,00E+00 | 0,00E+00 | 0,00E+00 | 0,00E+00 | 0,00E+00 |

O Erro Quadrático Médio para a estimativa do coeficiente de tendência foi, em média, maior quando utilizamos o método de máxima verossimilhança do que quando utilizamos o método não-paramétrico, independente da janela utilizada, exceto para o caso na Tabela 6.17.

Os Erros Quadráticos Integrados para o coeficiente de tendência apresentaram valores extremamente pequenos, o que ajuda a reforçar a idéia de que o método não-paramétrico fornece boas estimativas para o coeficiente.

Quanto ao coeficiente de difusão, os valores dos EQM foram bem pequenos indicando uma proximidade muito grande entre a curva teórica e a curva estimada, principalmente quando utilizamos a janela do Estágio 2.

Outra consideração importante é que a variabilidade dos EQM foi menor para o método não-paramétrico com a janela do Estágio 2 quando comparamos com o método paramétrico, para as estimativas do coeficiente de tendência, exceto para o caso na Tabela 6.17. Para as estimativas do coeficiente de difusão o método não-paramétrico utilizado com a janela do Estágio 2, apresentou menor variabilidade em todos os casos considerados.

Os Erros Quadráticos Integrados para as estimativas tanto do coeficiente de tendência como o de difusão foram extremamente pequenos, considerando um arredondamento acima da décima casa decimal.

7 Aplicações

O objetivo deste capítulo é aplicar o método não paramétrico e obter uma estimativa para o coeficiente de tendência e para o coeficiente de difusão quando temos um conjunto de dados reais. Foram escolhidos três bancos de dados, diretamente do site do IPEA no endereço <http://www.ipeadata.gov.br>. O primeiro banco de dados de depósito interbancário foi fornecido e utilizado por Gonçalves (2006) em sua dissertação de mestrado.

7.1 Depósito Interbancário

As observações deste conjunto de dados são de uma série diária brasileira de depósito interbancário (DI) com taxa de juros pré-fixada anual. A série contém 497 observações no período de 25/03/2003 a 22/03/2005. Os valores são em porcentagem ao ano (%a.a). O gráfico da série pode ser observado na Figura 7.1.

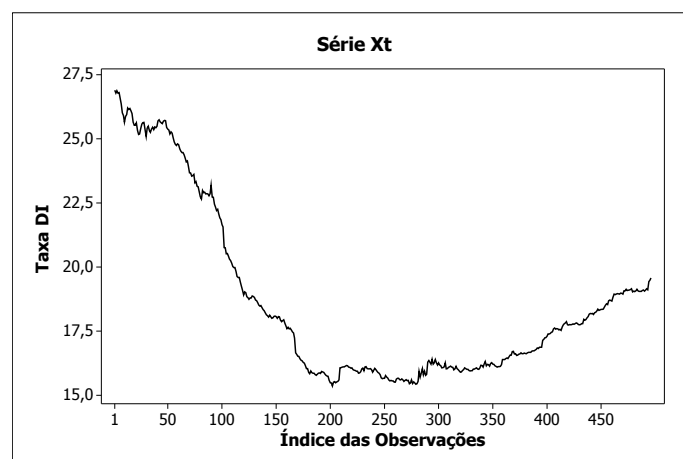


Figura 7.1: Série de Depósito Interbancário-DI.

Além das estimativas de primeira ordem, construímos para este exemplo aproximações de segunda e terceira ordem de acordo com Stanton (1997) e utilizando a generalização mostrada no Capítulo 3, de acordo com Fan e Zhang (2003).

É importante ressaltar que devemos ter muito cuidado ao utilizar as aproximações de ordem superior. Stanton (1997) afirma que quanto maior a ordem da aproximação, mais rápido as estimativas das funções μ e b convergem para seus verdadeiros valores. Porém, Fan e Zhang (2003) provam que a variância condicional $\text{var}(Y_{i\Delta}^*|X_{i\Delta}^*)$ e $\text{var}(Z_{i\Delta}^*|X_{i\Delta}^*)$ crescem exponencialmente da seguinte forma:

$$\text{Var}(Y_{i\Delta}^*|X_{i\Delta}^*) = \sigma^2(x_0)V_1(k)\Delta^{-1}\{1 + O(\Delta)\} \quad (7.1)$$

$$\text{Var}(Z_{i\Delta}^*|X_{i\Delta}^*) = 2\sigma^4(x_0)V_2(k)\Delta^{-1}\{1 + O(\Delta)\}. \quad (7.2)$$

Os valores de $V_1(k)$ e $V_2(k)$ foram calculados por Fan e Zhang (2003) e seus valores se encontram na Tabela 7.1

Tabela 7.1: Valores de $V_1(k)$ e $V_2(k)$.

| | Ordem K | | | | | | | | | |
|----------|---------|------|------|-------|-------|--------|--------|---------|---------|----------|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| $V_1(k)$ | 1,00 | 2,50 | 4,83 | 9,25 | 18,95 | 42,68 | 105,49 | 281,65 | 798,01 | 2364,63 |
| $V_2(k)$ | 1,00 | 3,00 | 8,00 | 21,66 | 61,50 | 183,40 | 570,66 | 1837,28 | 6076,25 | 20527,22 |

7.1.1 Aproximação de Primeira Ordem

A estimativa do coeficiente de tendência pode ser observada na Figura 7.3. Como $\mu(X_t)$ tenta capturar a relação existente entre os pares (X_t, Y_t) , onde $Y_t = (X_{t+1} - X_t)\Delta^{-1}$. Usando o esquema de discretização de Euler e tomando-se a esperança no modelo (1.1), temos:

$$E[X_{t+1} - X_t/X_t] = \mu(X_t)\Delta. \quad (7.3)$$

Então nosso interesse na Figura 7.2 é mostrar como $Y_t = (X_{t+1} - X_t)\Delta^{-1}$ e a estimativa $\hat{\mu}(X_t)$ se relacionam com X_t .

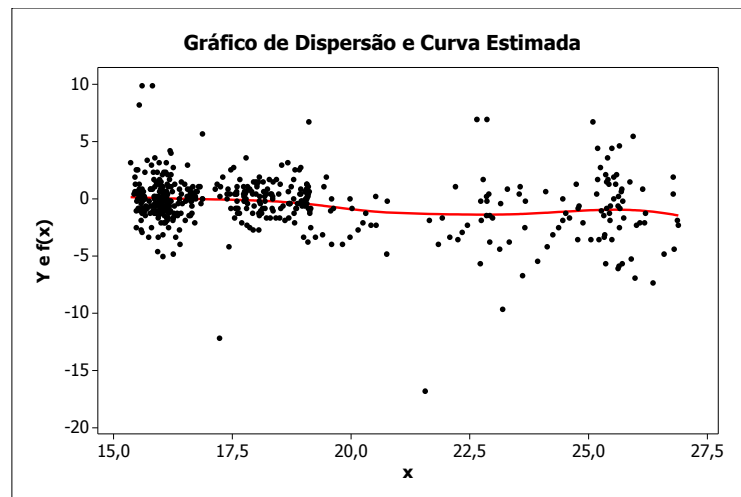


Figura 7.2: Gráfico de Dispersão e Estimativa do Coef. de Tendência-DI.

Em uma escala diferente podemos observar o comportamento do coeficiente de tendência na Figura 7.3.

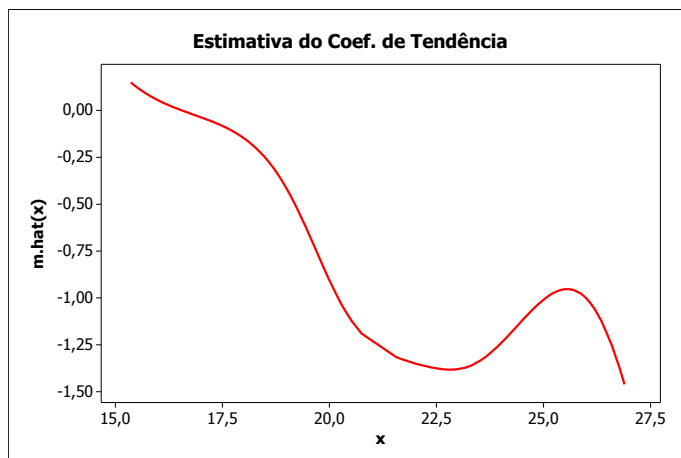


Figura 7.3: Estimativa do Coef. de Tendência-DI.

A estimativa do coeficiente de difusão pode ser observada na Figura 7.4.

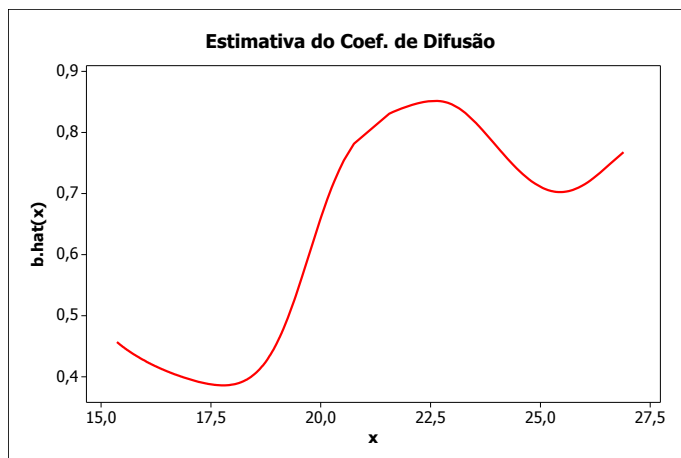


Figura 7.4: Estimativa do Coef. de Difusão-DI.

Como pode ser observado na Figura 7.3, a estimativa do coeficiente de tendência apresenta forte não-linearidade. Dessa forma, os modelos que assumem linearidade não devem ser utilizados.

Para reconstruir a série utilizando as estimativas do coeficiente de tendência, consideramos:

$$\hat{X}_{t+1} = \hat{\mu}(X_t)\Delta + X_t. \quad (7.4)$$

Na figura 7.5 temos as séries reconstruídas utilizando as estimativas do coeficiente de tendência.

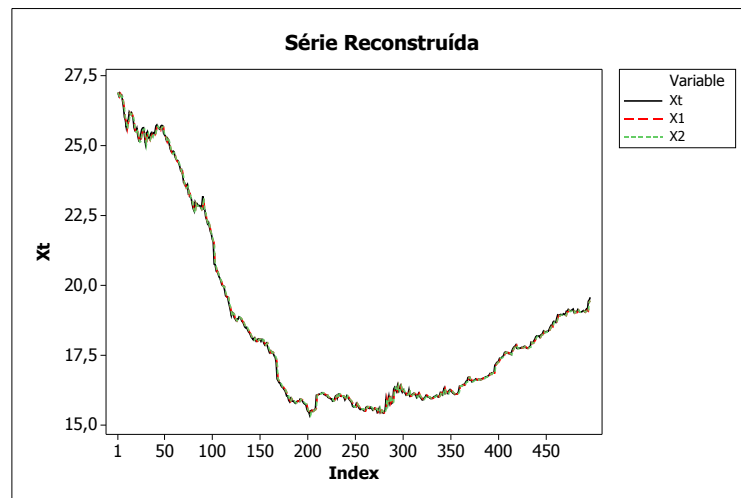


Figura 7.5: Séries Reconstruídas Usando as Estimativa do Coef. de Tendência.

Observamos que as séries estimadas são bem próximas da verdadeira série de taxas de depósito interbancário.

7.1.2 Aproximação de Segunda Ordem

A estimativa do coeficiente de tendência pode ser observada na Figura 7.3. Como $\mu(X_t)$ tenta capturar a relação existente entre os pares (X_t, Y_t^*) , onde

$$Y_t^* = \sum_{j=1}^2 a_{2j}(X_{t+j} - X_t)$$

e os a_{21} e a_{22} são obtidos pelo teorema 3.2.1.

Usando o esquema de discretização de Euler e tomando-se a esperança no modelo (1.1), temos:

$$\sum_{j=1}^2 a_{2j} E[X_{t+j} - X_t / X_t] = \mu(X_t) \Delta. \quad (7.5)$$

Então, nosso interesse na Figura 7.6 é mostrar como Y_t^* e a estimativa $\hat{\mu}(X_t)$ se relacionam com X_t .

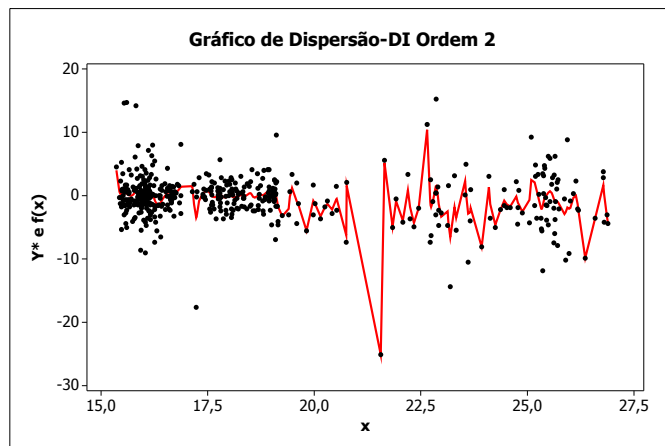


Figura 7.6: Gráfico de Dispersão e Estimativa do Coef. de Tendência-DI- Ordem 2.

Em uma escala diferente podemos observar o comportamento do coeficiente de tendência na Figura 7.7.

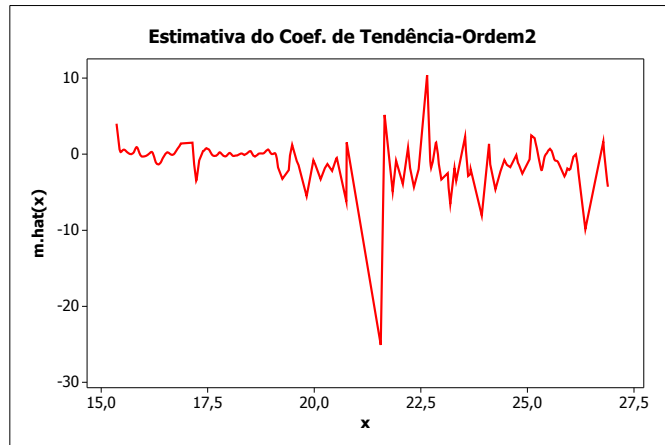


Figura 7.7: Estimativa do Coef. de Tendência-DI Ordem2.

E podemos observar a estimativa do coeficiente de difusão na Figura 7.8.

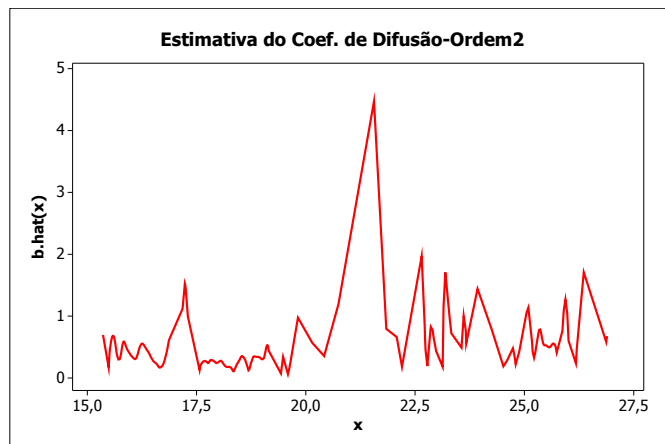


Figura 7.8: Estimativa do Coef. de Difusão-DI Ordem2.

A reconstrução da série quando utilizamos aproximação de segunda ordem é feita da seguinte forma:

$$X_{(t+2)} = \frac{1}{a_{22}} [\hat{\mu}(X_t)\Delta - a_{21}(X_{t+1} - X_t)] + X_t. \quad (7.6)$$

Na figura 7.9 temos as séries reconstruídas utilizando as estimativas do coeficiente de tendência.

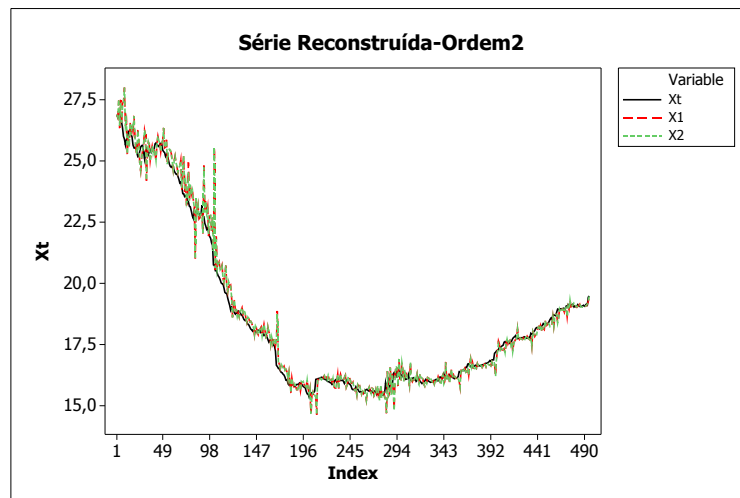


Figura 7.9: Séries Reconstruídas Usando as Estimativa do Coef. de Tendência - Ordem2.

Observamos que as séries estimadas já não são tão próximas da verdadeira série de taxas de depósito interbancário. Este fato ocorre devido ao aumento da variância condicional.

7.1.3 Aproximação de Terceira Ordem

A estimativa do coeficiente de tendência pode ser observada na Figura 7.3. Como $\mu(X_t)$ tenta capturar a relação existente entre os pares (X_t, Y_t^*) , onde

$$Y_t^* = \sum_{j=1}^3 a_{2j}(X_{t+j} - X_t)$$

e os a_{31} , a_{32} e a_{33} são obtidos pelo teorema 3.2.1.

Usando o esquema de discretização de Euler e tomando-se a esperança no modelo (1.1), temos:

$$\sum_{j=1}^3 a_{2j} E[X_{t+j} - X_t / X_t] = \mu(X_t) \Delta. \quad (7.7)$$

Então nosso interesse na Figura 7.10 é mostrar como Y_t^* e a estimativa $\hat{\mu}(X_t)$ se relacionam com X_t .

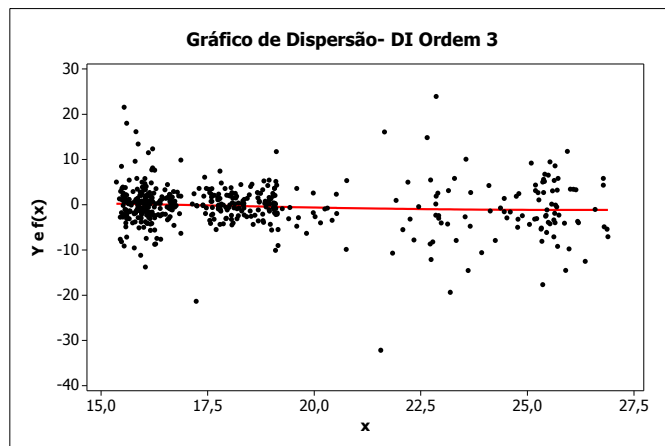


Figura 7.10: Gráfico de Dispersão e Estimativa do Coef. de Tendência-DI- Ordem 3.

Em uma escala diferente podemos observar o comportamento do coeficiente de tendência na Figura 7.11.

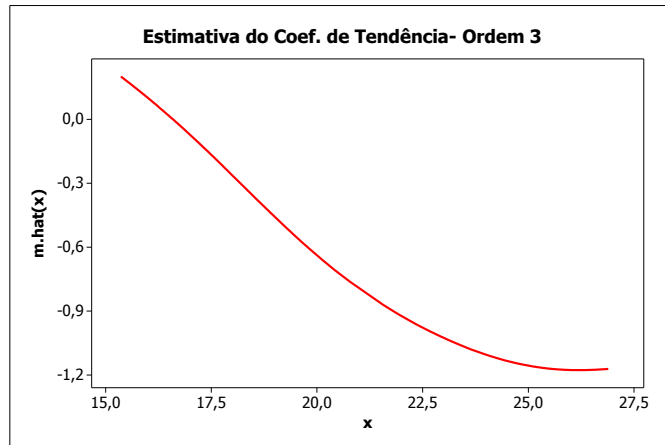


Figura 7.11: Estimativa do Coef. de Tendência-DI Ordem3.

E podemos observar a estimativa do coeficiente de difusão na Figura 7.12.

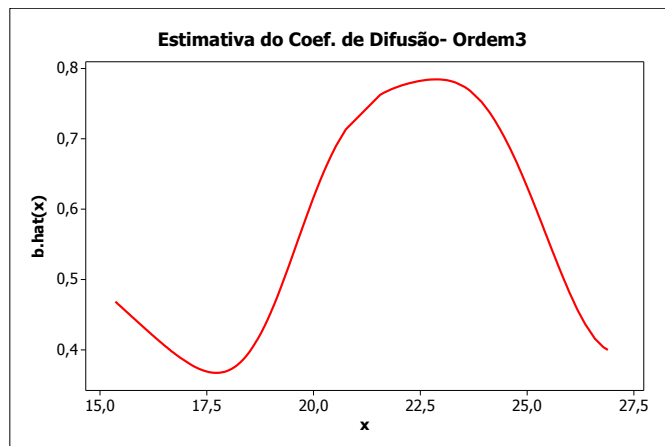


Figura 7.12: Estimativa do Coef. de Difusão-DI Ordem3.

A reconstrução da série quando utilizamos aproximação de terceira ordem é feita da seguinte forma:

$$X_{(t+3)} = \frac{1}{a_{33}}[\hat{\mu}(X_t)\Delta - a_{31}(X_{t+1} - X_t) - a_{32}(X_{t+2} - X_t)] + X_t. \quad (7.8)$$

Na figura 7.13 temos as séries reconstruídas utilizando as estimativas do coeficiente de tendência.

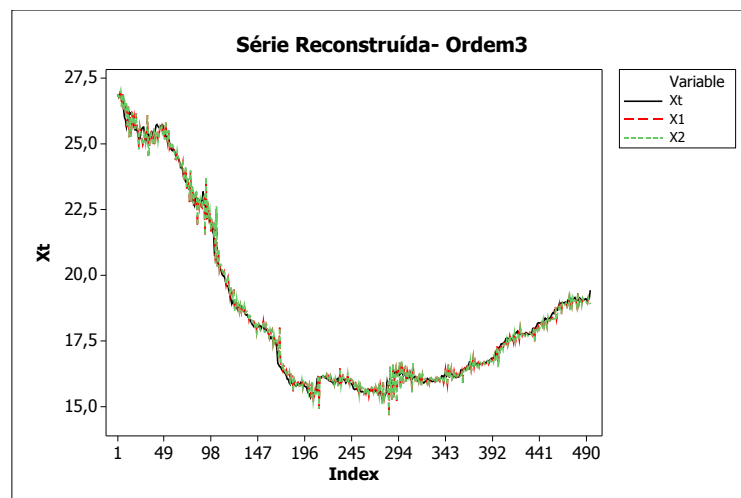


Figura 7.13: Séries Reconstruídas Usando as Estimativa do Coef. de Tendência - Ordem3.

Observamos que as séries estimadas já não são tão próximas da verdadeira série de taxas de depósito interbancário. Este fato ocorre devido ao aumento da variância condicional.

7.2 C-Bond spread

As observações deste banco de dados são do C-Bond, o principal título da dívida externa brasileira negociado no mercado internacional. Os valores são de uma série diária, em pontos-base (p.p) sobre o título do Tesouro dos EUA. Quanto maior a procura pelo papel, maior é o seu valor de mercado e maior o sinal de confiança dos investidores na economia do país. A série contém 485 observações no período de 20/10/2003 a 20/10/2005. O gráfico da série pode ser observado na Figura 7.14.

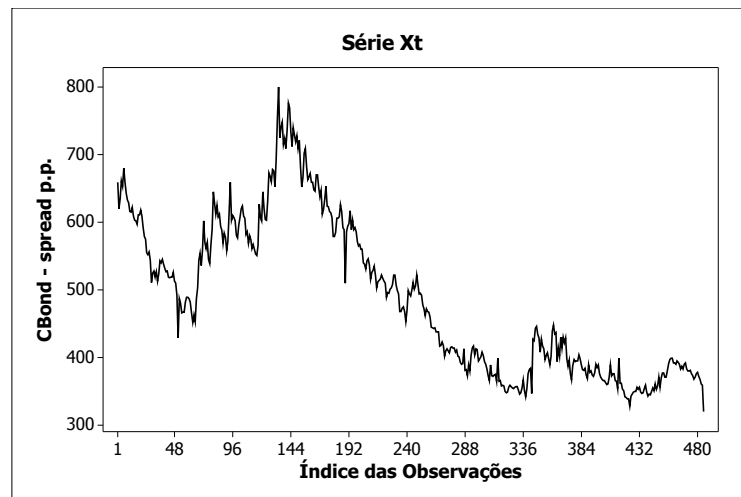


Figura 7.14: Série C-Bond spread.

Analogamente ao exemplo anterior, a estimativa do coeficiente de tendência pode ser observada na Figura 7.15. Como $\mu(X_t)$ tenta capturar a relação existente entre os pares (X_t, Y_t) , onde $Y_t = (X_{t+1} - X_t)\Delta^{-1}$. Usando o esquema de discretização de Euler e tomando-se a esperança no modelo (1.1), temos:

$$E[X_{t+1} - X_t/X_t] = \mu(X_t)\Delta. \quad (7.9)$$

Então nosso interesse na Figura 7.15 é mostrar como $Y_t = (X_{t+1} - X_t)\Delta^{-1}$ e a estimativa $\hat{\mu}(X_t)$ se relacionam com X_t .

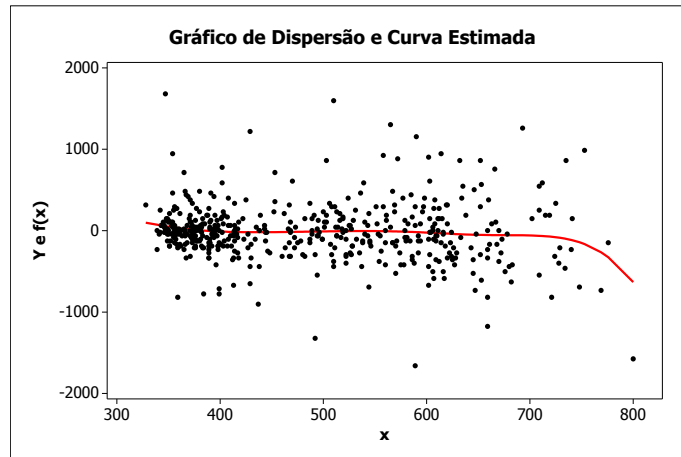


Figura 7.15: Gráfico de Dispersão e Estimativa do Coef. de Tendência-CBond.

Em uma escala diferente podemos observar o comportamento do coeficiente de tendência na Figura 7.16.

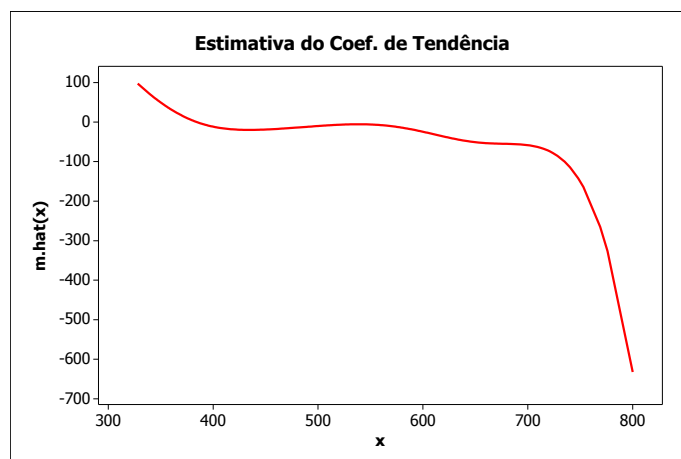


Figura 7.16: Estimativa do Coef. de Tendência-CBond.

Como pode ser observado na Figura 7.16, a estimativa do coeficiente de tendência apresenta não-linearidade. Dessa forma, os modelos que assumem linearidade não devem ser utilizados.

A estimativa do coeficiente de difusão pode ser observada na Figura 7.17.

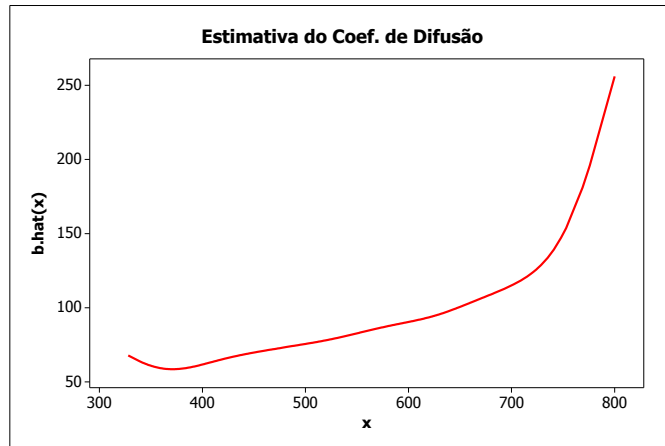


Figura 7.17: Estimativa do Coef. de Difusão-CBond.

Na figura 7.18 temos as séries reconstruídas utilizando as estimativas do coeficiente de tendência, analogamente ao exemplo anterior.

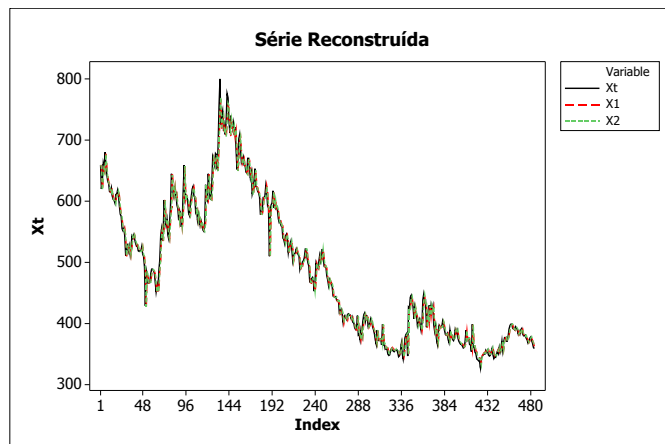


Figura 7.18: Séries Reconstruídas Usando as Estimativa do Coef. de Tendência.

Observamos que as séries estimadas são bem próximas da verdadeira série dos valores do CBond.

7.3 Overnight/Selic

A série deste conjunto de dados é a taxa Overnight/Selic, média dos juros que o governo paga aos bancos que lhe emprestaram dinheiro. Esta taxa serve de referência para outras taxas de juros do país e é a taxa básica de juros da economia. Os valores são em porcentagem ao ano (%a.a) e compreendem o período de 04/01/2003 a 22/02/2007 totalizando 1040 observações. O gráfico da série pode ser observado na Figura 7.19.

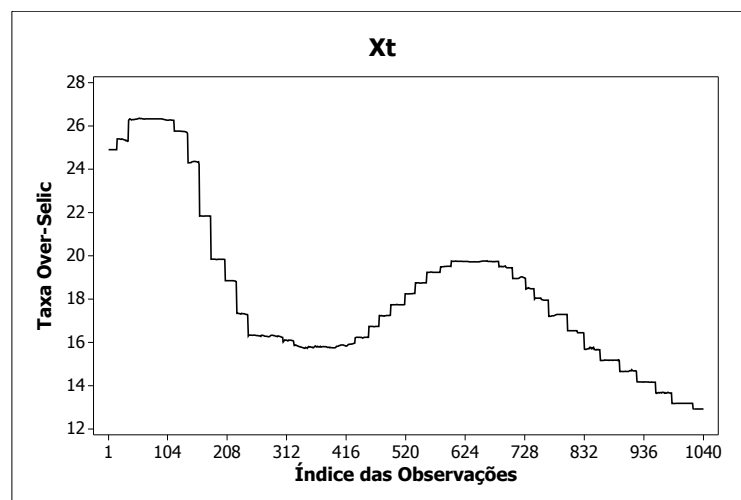


Figura 7.19: Série Over Selic.

Analogamente ao exemplo anterior, a estimativa do coeficiente de tendência pode ser observada na Figura 7.20. Como $\mu(X_t)$ tenta capturar a relação existente entre os pares (X_t, Y_t) , onde $Y_t = (X_{t+1} - X_t)\Delta^{-1}$. Usando o esquema de discretização de Euler e tomando-se a esperança no modelo (1.1), temos:

$$E[X_{t+1} - X_t/X_t] = \mu(X_t)\Delta. \quad (7.10)$$

Então nosso interesse na Figura 7.20 é mostrar como $Y_t = (X_{t+1} - X_t)\Delta^{-1}$ e a estimativa $\hat{\mu}(X_t)$ se relacionam com X_t .

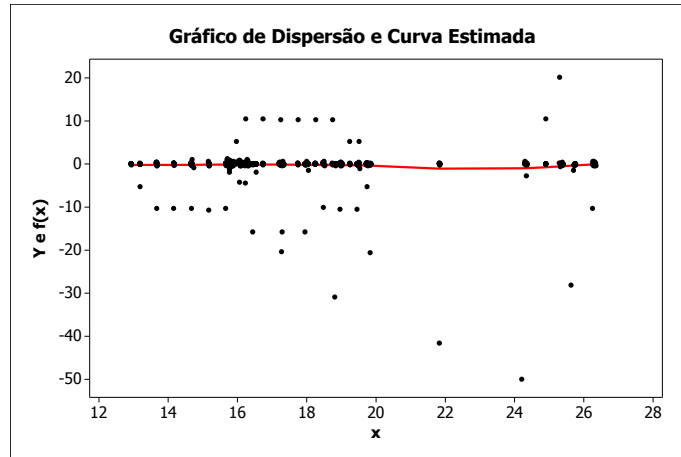


Figura 7.20: Gráfico de Dispersão e Estimativa do Coef. de Tendência-Over/Selic.

Em uma escala diferente podemos observar o comportamento do coeficiente de tendência na Figura 7.21.

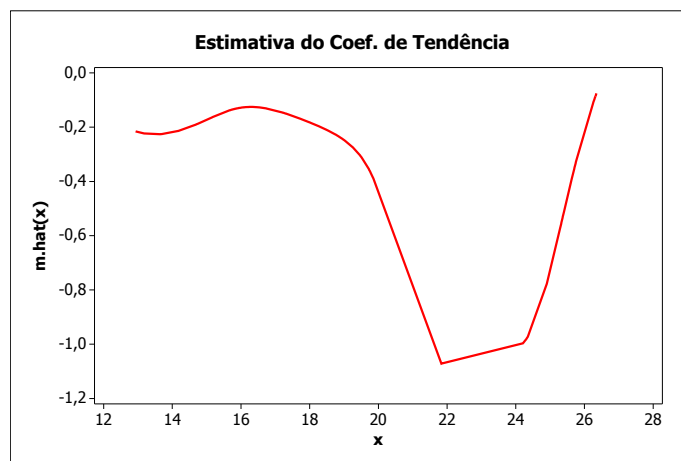


Figura 7.21: Estimativa do Coef. de Tendência-OverSelic.

Analogamente aos exemplos anteriores, os modelos que assumem linearidade para o coeficiente de tendência não devem ser utilizados. Como podemos observar na Figura 7.21, o coeficiente apresenta não-linearidade.

A estimativa do coeficiente de difusão pode ser observada na Figura 7.22.

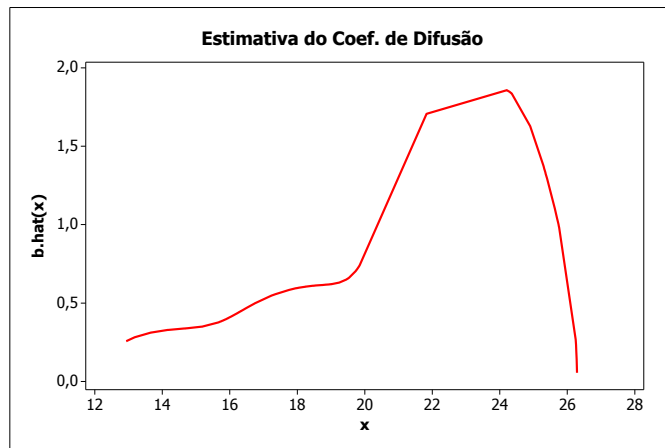


Figura 7.22: Estimativa do Coef. de Difusão-OverSelic.

Na figura 7.23 temos as séries reconstruídas utilizando as estimativas do coeficiente de tendência, analogamente ao exemplo anterior

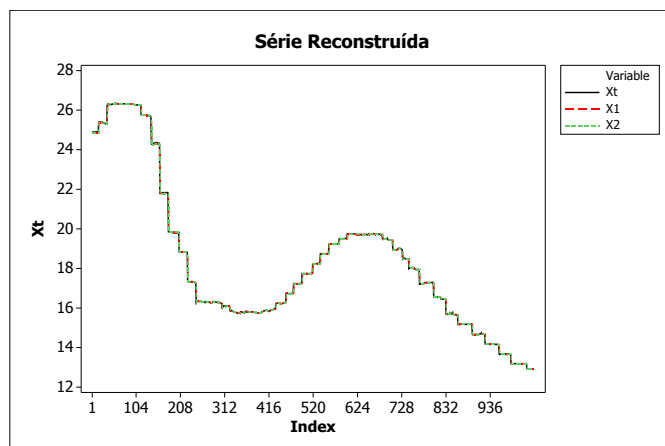


Figura 7.23: Séries Reconstruídas Usando as Estimativa do Coef. de Tendência.

Observamos que as séries estimadas são bem próximas da verdadeira série dos valores da Taxa Overnight/Selic.

8 *Considerações Finais*

Ao considerar o interesse em estimar os coeficientes μ e b evitando a especificação de formas funcionais para os coeficientes, o estudo do método não-paramétrico é de grande relevância. Neste trabalho, estudamos as aproximações para as funções μ e b de acordo com a abordagem de Stanton (1997) porém usando a regressão polinomial local de acordo com Fan e Zhang (2003), cuja estimação da janela está descrita em Fan e Gijbels (1995).

Fan e Zhang (2003) também propõem um método para testar se os dados se ajustam a um modelo de difusão conhecido *versus* a alternativa dada pelo modelo geral na expressão (1.1). Se as observações se adequarem ao modelo escolhido para teste, a estimação dos coeficientes pode ser facilmente obtida usando o método de máxima verossimilhança de acordo com Cleur e Manfredi (1999). Caso contrário, quando não conseguimos determinar a princípio formas funcionais para os coeficientes, o método não-paramétrico deve ser utilizado. Como podemos observar após o estudo de simulação no Capítulo 6, ambos os métodos apresentaram erros extremamente pequenos e mostraram ser eficientes para estimação de μ e b .

Como mostrado por Fan e Zhang (2003), as aproximações de ordem superior embora reduzam o vício das estimativas, aumentam as variâncias condicionais. Este fato pode ser observado no exemplo da seção 7.1 do Capítulo 7. Sendo assim, devemos ter cautela em utilizá-las.

Para trabalhos futuros, após a estimação dos coeficientes, é de grande interesse estudarmos métodos de previsão para as observações futuras e utilizarmos a estimativa do coeficiente de difusão para construirmos intervalos de confiança para essas estimativas.

Além disso, é de interesse estudar métodos para testar se as observações seguem processos de difusão não-homogêneos no tempo e para os casos onde podemos verificar essa não-homogeneidade, estudar formas de estimar os coeficientes.

Outro interesse é o estudo de outros processos, por exemplo os processos de Lèvy e estudar formas de estimação para as funções desconhecidas do processo.

Referências Bibliográficas

- AKAIKE, H. Statistical predictor information. *Annals of the Institute of Statistical Mathematics*, v. 22, p. 203–217, 1970.
- AKAIKE, H. A new look at the statistical model identification. *IEEE Transactions of Automatic Control AC*, v. 19, p. 716–723, 1974.
- ARFI, M. Non-parametric drift estimation from ergodic samples. *Journal of Nonparametric Statistics*, v. 5, p. 381–289, 1995.
- ARFI, M. Non-parametric variance estimation from ergodic samples. *Scandinavia Journal of Statistics*, v. 25, p. 225–234, 1998.
- AIT-SAHALIA, Y. Nonparametric pricing of interest rate derivatives securities. *Econometrica*, v. 64, p. 527–560, 1996a.
- BIANCHI, C.; CESARI, R.; PANATTONI, L. Alternative estimators of a diffusion model of the term structure of interest rates: A monte carlo comparison. In: *Temi di Discussione della Banca d'Italia*. [S.l.: s.n.], 1994.
- BIANCHI, C.; CLEUR, E. Indirect estimation of stochastic differential equation models: some computational experiments. *Computational Economics*, v. 9, p. 257–274, 1996.
- CHIU, S. T. Some stabilized bandwidth selectors for nonparametric regression. *Annals of Statistics*, v. 19, p. 1528–1546, 1991.
- CLARK, R. A calibration curve for radiocarbon dates. *Antiquity*, v. 49, p. 251–266, 1975.
- CLEUR, E. M.; MANFREDI, P. One dimensional sde models, low order numerical methods and simulation based estimation: a comparison of alternative estimators. *Computational Economics*, v. 13 (2), p. 177–197, 1999.
- COX, J.; INGERSOLL, J.; ROSS, S. A theory of the term structure of interest rates. *Econometrica*, v. 53, p. 385–467, 1985.
- CRAVEN, P.; WAHBA, G. Smoothing noisy data with spline functions. *Numerische Mathematik*, v. 31, p. 377–403, 1979.
- FAN, J.; GIJBELS, I. Data-driven bandwidth selection in local polynomial fitting: Variable bandwidth and spatial adaptation. *Journal of the Royal Statistical Society B*, v. 57, p. 371–394, 1995.
- FAN, J.; YAO, Q. Efficient estimation of conditional variance functions in stochastic regression. *Biometrika*, v. 85, p. 645–660, 1998.

- FAN, Z.; ZHANG, C. A reexamination of diffusion estimators with applications to financial model validation. *Journal of the American Statistical Association*, v. 98, p. 118–134, 2003.
- GASSER, T.; KNEIP, A.; KÖHLER, W. A flexible and fast method for automatic smoothing. *Journal of the American Statistical Association*, v. 86, p. 643–652, 1991.
- GONÇALVES, R. Q. *Estimação do Coeficiente de Tendência de uma Equação Diferencial Estocástica: Uma Aplicação à Estrutura a Termo das Taxas de Juros*. 2006. Dissertação (Mestrado) — Universidade Federal de Minas Gerais.
- HALL, P.; MARRON, J.; PARK, B. Smoothed cross-validation. *Probability Theory and Related Fields*, v. 92, p. 1–20, 1992.
- HANSEN, L. Large samples properties of generalized method of moments estimators. *Econometrica*, v. 50, p. 1029–1054, 1982.
- HÄRDLE, W.; HALL, P.; MARRON, J. S. How far are automatically chosen regression smoothing parameter from their optimum? *Journal of the American Statistical Association*, v. 83, p. 86–99, 1988.
- HÄRDLE, W.; HALL, P.; MARRON, J. S. Regression smoothing parameters that are not far from their optimum. *Journal of the American Statistical Association*, v. 87, p. 227–233, 1992.
- HÄRDLE, W.; MARRON, J. S. Random approximations to an error criterion of nonparametric statistics. *Journal of Multivariate Analysis*, v. 20, p. 91–113, 1986.
- JIANG, G.; KNIGHT, J. L. A nonparametric approach to the estimation of diffusion processes, with an application to a short-term interest rate model. *Econometric Theory*, v. 13, p. 615–645, 1997.
- KLOEDEN, P. E.; PLATEN, E. *Numerical Solution of Stochastic Differential Equations*. [S.l.]: Springer-Verlag, 1992.
- MATHEWS, J. H. *Numerical Methods*. [S.l.: s.n.], 1987.
- OKSENDAL, B. *Stochastic Differential Equations: An Introduction With Applications*. Oslo, Norway: Springer-Verlag., 2003.
- PEARSON, N.; SUM, T. Exploiting conditional density in estimating the term structure: An application to the cox, ingersoll, and ross model. *Journal of Finance*, v. 49, p. 1279–1304, 1994.
- PHAM, D. Nonparametric estimation of the drift coefficient in the diffusion equation. *Mathematische Operationsforschung und Statistik Series Statistics*, v. 12, p. 61–73, 1981.
- RAO, B. P. Estimation of the drift for diffusion process. *Statistics*, v. 16, p. 263–275, 1985.
- RAO, B. P. *Statistical inference for Diffusion Type Processes*. [S.l.]: Arnold, 1999.
- RICE, J. Bandwidth choice for nonparametric regression. *Annals of Statistics*, v. 12, p. 1215–1230, 1984.

- SHIBATA, R. An optimal selection of regression variables. *Biometrika*, v. 68, p. 45–54, 1981.
- SIMONOFF, J. *Smoothing Methods in Statistics*. New York:: Springer-Verlag, 1996.
- STANTON, R. A nonparametric model of term structure dynamics and the market price of interest rate risk. *The Journal of Finance*, v. 52, p. 1973–2002, 1997.
- STONE, M. Cross-validatory choice and assessment of statistical predictions (with discussion). *Journal of the Royal Statistical Society B*, v. 36, p. 111–147, 1974.