

Regis Queiroz Gonçalves

**ESTIMAÇÃO DO COEFICIENTE DE TENDÊNCIA DE
UMA EQUAÇÃO DIFERENCIAL ESTOCÁSTICA:
UMA APLICAÇÃO À ESTRUTURA A TERMO
DAS TAXAS DE JUROS**

Dissertação apresentada ao Programa de Pós-Graduação do Departamento de Estatística como requisito parcial à obtenção do título de Mestre em Estatística.

Área de concentração: Estatística.

Orientador: Prof. Gregorio Saravia Atúncar, PhD.

Belo Horizonte

Instituto de Ciências Exatas da UFMG

2006

Dedico este trabalho à Cleonice, Adeir, Regina e Aurita.

gregorio-regis-icex-ufmg

Agradecimentos

Em primeiro lugar agradeço a Deus, por Sua justiça e bondade infinitas, começando pela linda família no seio da qual me permitiu nascer.

À minha amada família: aos meus pais, Adeir e Cleonice, pela luta incessante para me dar condições de prosseguir em minha vida, por todo o amor e carinho que me apoiam e pela educação e valores que me ensinaram; à minha irmã Regina, pelo exemplo de autoconfiança e perseverança e à minha avó Aurita, minha segunda mãe, por me apoiar tanto desde o início de minha caminhada.

Ao meu tio Isidoro, cuja intervenção, demonstrando seu carinho e preocupação para comigo, foi decisiva para que eu pudesse continuar meus estudos.

Ao Gregorio, meu orientador e amigo, pela sabedoria, disponibilidade, compreensão, paciência, atenção, incentivo, e, principalmente, por acreditar em mim até quando eu mesmo duvidei. Obrigado por toda a ajuda e pela condução rigorosa deste trabalho.

Aos professores do mestrado, pelo conhecimento e experiência transmitidos nas aulas e seminários, que me possibilitaram avançar em meus estudos e conhecer um mundo fantástico. Em especial ao Professor Luiz Henrique Duczmal.

Aos membros da banca examinadora, pela importância fundamental para a relevância e a qualidade deste trabalho.

Aos meus ex-mestres e amigos da graduação, professores Antônio José Manzato, Luiz

Carlos Baida, Fernando Ferrari e Trajano Nóbrega Neto, pelo incentivo e pelas colaborações que me levaram ao mestrado.

Ao amigo Flávio Ferreira, que me recebeu quando cheguei à Belo Horizonte. Seu exemplo de luta e de superação na vida, sempre com alegria, não serão jamais esquecidos por mim.

Ao Frank, a quem muito considero, pela honra de sua amizade e parceria e por me servir de exemplo de determinação e objetividade. À minha grande amiga Jú Mambrini, todo o reconhecimento, carinho, respeito e amizade. Aos amigos do mestrado, em especial ao Augusto, à Cléia, à Josenete e ao Ricardo Tavares.

À Dona Darcinha e à Juliana, pela hospitalidade e amizade, em cujo lar senti-me novamente em uma família e onde encontrei paz e inspiração para iniciar o esboço do presente trabalho.

Aos amigos Ricardo Tolentino, Débson, Daniel, Nancy, Savana, Elio, Serginho, Márcio, João Batista, Henrique Geanezi, Henrique Leão, Reginaldo, Estênio e Janaína Fagundes. Por todos os momentos bons e ruins, pelos risos e as lágrimas, pelo apoio que sempre me deram e por não me deixarem só.

De forma alguma posso deixar de mencionar as valorosas contribuições do Breno, pela ajuda com a programação, do Helvécio e do Rodrigo, pela agilidade e a quantidade de artigos que conseguiram para a minha pesquisa e da Rogéria e da Marcinha, que nunca faltaram com a atenção e a disponibilidade que precisei quando passei pela secretaria. À todos os demais funcionários do Departamento de Estatística.

À todas as pessoas que participaram direta ou indiretamente do processo que se encerrou com a produção deste trabalho e que não tenham sido mencionadas.

“Deus não joga dados.”

Albert Einstein

gregorio-regis-icex-ufmg

Sumário

Resumo	7
Introdução	8
1 Revisão de Literatura	12
1.1 Abordagens Paramétricas	12
1.2 Abordagens Não-paramétricas	20
2 Cálculo Estocástico	30
2.1 Conceitos Preliminares	30
2.2 O Cálculo de Itô	32
3 Taxas de Juros	37
3.1 Formação das Taxas de Juros	37
3.1.1 Taxa Livre de Risco	38
3.2 Títulos Públicos	39
3.2.1 Mercado de Bonds	40
3.2.2 Rendimento na Maturidade	42
3.2.3 Volatilidade dos Bonds	43
3.3 A Estrutura a Termo das Taxas de Juros	44
3.4 Modelos de Um Fator e com Reversão à Média	45
3.4.1 Modelo de Vasicek	46
3.4.2 Modelo de Cox-Ingersoll-Ross (CIR)	46

4	Simulação	48
4.1	Geração de Números Pseudo-aleatórios	48
4.2	Simulação de Processos de Difusão	49
4.2.1	Casos Simulados	52
5	Aplicações	69
5.1	Títulos dos EUA	69
5.2	Depósito Interbancário (DI) Prefixado do Brasil	69
5.3	Títulos do Canadá	71
5.4	Resumo das Estimativas	73
6	Considerações Finais	78
6.1	Conclusões	78
6.2	Propostas de Trabalhos Futuros	78
	Abstract	80
	Referências Bibliográficas	81

Lista de Figuras

4.1	Trajetoária simulada de um modelo de CIR $\kappa = 0,3$, $\nu = 0,1$, $\sigma = 0,06$, $T = 60$ e $\Delta = 0,04761$	56
4.2	Trajetoária simulada de um modelo de Vasicek $\kappa = 0,3$, $\nu = 0,1$, $\sigma = 0,06$, $T = 60$ e $\Delta = 0,04761$	56
4.3	Trajetoária simulada de um modelo de CIR $\kappa = 0,8$, $\nu = 0,1$, $\sigma = 0,06$, $T = 60$ e $\Delta = 0,04761$	57
4.4	Trajetoária simulada de um modelo de Vasicek $\kappa = 0,8$, $\nu = 0,1$, $\sigma = 0,06$, $T = 60$ e $\Delta = 0,04761$	57
4.5	Trajetoárias simuladas dos modelos CIR e Vasicek $\kappa = 0,3$, $\nu = 0,1$, $\sigma = 0,06$, $T = 60$ e $\Delta = 0,04761$	58
4.6	Trajetoárias simuladas dos modelos CIR e Vasicek $\kappa = 0,8$, $\nu = 0,1$, $\sigma = 0,06$, $T = 60$ e $\Delta = 0,04761$	58
4.7	Varição do vício relativo em função de T para o modelo CIR, com $\Delta = 0,1$ e $\kappa = 0,3$	59
4.8	Varição do vício relativo em função de T para o modelo CIR, com $\Delta = 0,01$ e $\kappa = 0,3$	59
4.9	Varição do vício relativo em função de T para o modelo de Vasicek, com $\Delta = 0,1$ e $\kappa = 0,3$	60
4.10	Varição do vício relativo em função de T para o modelo de Vasicek, com $\Delta = 0,01$ e $\kappa = 0,3$	60
4.11	Varição do vício relativo em função do Δ para o modelo CIR, com $T = 60$ e $\kappa = 0,3$	61

4.12	Variação do vício relativo em função do Δ para o modelo CIR, com $T = 120$ e $\kappa = 0, 3$	61
4.13	Variação do vício relativo em função do Δ para o modelo de Vasicek, com $T = 60$ e $\kappa = 0, 3$	62
4.14	Variação do vício relativo em função do Δ para o modelo de Vasicek, com $T = 120$ e $\kappa = 0, 3$	62
4.15	Variação do vício relativo de $\hat{\kappa}_{mvd}$ para o modelo de Vasicek e $\kappa = 0, 3$. . .	63
4.16	Variação do vício relativo de $\hat{\kappa}_{mvd}$ para o modelo CIR e $\kappa = 0, 3$	63
4.17	Variação do vício relativo de $\hat{\kappa}_{mqo}$ para o modelo de Vasicek e $\kappa = 0, 3$. . .	64
4.18	Variação do vício relativo de $\hat{\kappa}_{mqo}$ para o modelo de CIR e $\kappa = 0, 3$	64
4.19	Variação do desvio padrão de $\hat{\kappa}_{mvd}$ para o modelo de Vasicek e $\kappa = 0, 3$. . .	65
4.20	Variação do desvio padrão de $\hat{\kappa}_{mvd}$ para o modelo de CIR e $\kappa = 0, 3$	65
4.21	Variação do desvio padrão de $\hat{\kappa}_{mqo}$ para o modelo de Vasicek e $\kappa = 0, 3$. . .	66
4.22	Variação do desvio padrão de $\hat{\kappa}_{mqo}$ para o modelo de CIR e $\kappa = 0, 3$	66
4.23	Distribuição de $\hat{\kappa}_{mvd}$ para o modelo de CIR com $\kappa = 0, 8$, $\nu = 0, 1$, $\sigma =$ $0, 06$, $T = 60$ meses e $\Delta = 0, 04761$	67
4.24	Distribuição de $\hat{\kappa}_{mqo}$ para o modelo de CIR com $\kappa = 0, 8$, $\nu = 0, 1$, $\sigma =$ $0, 06$, $T = 60$ meses e $\Delta = 0, 04761$	68
4.25	Distribuição de $\hat{\kappa}_{mvd}$ para o modelo de Vasicek com $\kappa = 0, 3$, $\nu = 0, 1$, $\sigma = 0, 06$, $T = 60$ meses e $\Delta = 0, 04761$	68
5.1	Taxas semanais dos Títulos de 3 meses dos EUA, de 05/01/1962 até 30/08/1996.	70
5.2	Dispersão das diferenças $Y_i = X_{i+1} - X_i$ das taxas contra as taxas X_i do tesouro dos EUA.	70
5.3	Taxas dos títulos dos EUA e tendências ajustadas pelos modelos de CIR e Vasicek.	71
5.4	Curvas de rendimento das taxas DI diárias em três datas distintas.	72
5.5	Dispersão das diferenças $Y_i = X_i - X_{i-1}$ das taxas DI de 3 meses pelas taxas X_{i-1}	72
5.6	Taxas DI diárias de 3 meses brasileiras, de 23/03/2003 até 22/03/2005. . .	73

5.7	Dispersão das diferenças $Y_i = X_i - X_{i-1}$ das taxas DI de 60 meses pelas taxas X_{i-1}	73
5.8	Taxas DI diárias de 60 meses e tendências ajustadas pelos modelos CIR e Vasicek.	74
5.9	Dispersão das diferenças $Y_i = X_i - X_{i-1}$ das taxas DI de 84 meses pelas taxas X_{i-1}	74
5.10	Taxas DI diárias de 84 meses e tendências ajustadas pelos modelos de CIR e Vasicek.	75
5.11	Dispersão das diferenças $Y_i = X_{i+1} - X_i$ das taxas contra as taxas X_i canadenses.	75
5.12	Taxas diárias dos títulos canadenses com maturidade de 3 meses, de 02/01/1996 até 30/03/2006.	76

gregorio-regis-icex-ufmg

Lista de Tabelas

3.1	<i>Títulos da dívida pública do Governo Federal.</i>	40
3.2	<i>Resultados de um bond para diferentes maturidades e vencimentos.</i>	44
4.1	<i>Integrais definindo os EMVTC.</i>	50
4.2	<i>EMVDs de κ do modelo de Vasicek; $T = 100$ e $\Delta = 0,1$.</i>	51
4.3	<i>Valores de Δ e seus respectivos números de observações, considerando o ano financeiro, de 252 dias.</i>	52
4.4	<i>EMVD e EMQO com desvios-padrão para $\kappa = 0,3$, $\Delta = 0,1$ e T em meses.</i>	53
4.5	<i>EMVD e EMQO com desvios-padrão para $\kappa = 0,3$, $\Delta = 0,04761$ e T em meses.</i>	54
4.6	<i>EMVD e EMQO com desvios-padrão para $\kappa = 0,3$, $\Delta = 0,01$ e T em meses.</i>	54
4.7	<i>EMVD e EMQO com desvios-padrão para $\kappa = 0,8$, $\Delta = 0,1$ e T em meses.</i>	54
4.8	<i>EMVD e EMQO com desvios-padrão para $\kappa = 0,8$, $\Delta = 0,04761$ e T em meses.</i>	55
4.9	<i>EMVD e EMQO com desvios-padrão para $\kappa = 0,8$, $\Delta = 0,01$ e T em meses.</i>	55
5.1	<i>Estimativas por EMVD para κ e ν.</i>	76
5.2	<i>Estimativas por EMQO para κ e ν.</i>	76

Resumo

Devido à sofisticação dos produtos financeiros nos anos recentes, os modelos probabilísticos de tempo contínuo dados por equações diferenciais estocásticas têm merecido a atenção crescente por parte de analistas do mercado financeiro e de acadêmicos. Para superar a limitação de alguns modelos, é necessário considerar a forma como as taxas de juros se desenvolvem ao longo do tempo. Neste trabalho a estimação do coeficiente de tendência dos modelos de Vasicek e de Cox-Ingersoll-Ross é realizada utilizando-se estimadores de máxima verossimilhança discretizados e de mínimos quadrados ordinários. Contorna-se a necessidade de se conhecer o valor do coeficiente de difusão e confirma-se a presença do vício dos estimadores, por se considerar observações em tempo discreto para modelos de tempo contínuo. São apresentadas comparações dos desempenhos dos estimadores variando-se o intervalo de discretização e o tempo de observação do processo. Os coeficientes de tendência de ambos os modelos são estimados pelos dois métodos, obtendo-se a trajetória esperada das taxas de juros de cinco conjuntos de dados reais.

Palavras-chave: Equações diferenciais estocásticas, Estrutura a termo das taxas de juros, modelo de Vasicek, modelo de Cox-Ingersoll-Ross, Esquema de discretização de Euler.

Introdução

Nas áreas quantitativas da ciência, o estudo dos fenômenos naturais e a busca de soluções para problemas reais realizam-se com base em modelos para estes fenômenos e problemas. Em muitos destes casos a modelagem proposta se traduz numa equação diferencial. Por exemplo, considere o modelo

$$\frac{dN(t)}{dt} = a(t)N(t), \quad N(0) = A$$

Este modelo representa o crescimento populacional. $N(t)$ é o tamanho da população no instante de tempo t , e $a(t)$ é a taxa de crescimento relativo no instante de tempo t . Este modelo matemático se torna mais realístico se alguma aleatoriedade for incluída em sua formulação. Assim, podemos pensar que $a(t)$ não é completamente conhecido, mas que está sujeito a algum efeito aleatório do meio ambiente, de modo a termos, na verdade,

$$a(t) = r(t) + \text{“ruído”}$$

onde não conhecemos o valor exato do termo “ruído”, somente sua distribuição de probabilidade, e $r(t)$ é assumido como determinístico. Deste modo, obtem-se

$$\frac{dN(t)}{dt} = [r(t) + \text{“ruído”}]N(t), \quad N(0) = A$$

que é o mesmo que

$$\frac{dN(t)}{dt} = r(t)N(t) + \text{“ruído”} N(t), \quad N(0) = A. \quad (1)$$

De acordo com Øksendal [32], a equação (1) é um dos exemplos mais simples de equação diferencial estocástica (EDE). As EDEs abordadas neste trabalho têm a seguinte forma geral

$$dX_t = f(t, X_t)dt + g(t, X_t)dW_t, \quad X_0 = x_0, \quad 0 \leq t \leq T \quad (2)$$

onde $\{W_t, t \geq 0\}$ é um Movimento Browniano padrão e T é o instante de tempo até o qual o processo é observado. As funções f e g são, em geral, chamadas de coeficientes, ou funções de tendência e de difusão, respectivamente.

A fim de não deixar dúvidas sobre a equação (2), é interessante observar a definição de Movimento Browniano padrão, segundo Billingsley [6]:

“O Movimento Browniano ou Processo de Wiener é um processo estocástico $[W_t : t \geq 0]$ sobre algum (Ω, \mathcal{F}, P) , com as seguintes propriedades:

- (i) O processo começa em 0: $P[W_0 = 0] = 1$.
- (ii) Os incrementos são independentes: se $0 \leq t_1 < t_2 < \dots < t_{n-1} < t_k$, então

$$P[W_{t_i} - W_{t_{i-1}} \in H_i, i \leq k] = \prod_{i \leq k} P[W_{t_i} - W_{t_{i-1}} \in H_i].$$

- (iii) Para $0 \leq s < t$ o incremento $W_t - W_s$ é normalmente distribuído com média 0 e variância $t - s$:

$$P[W_t - W_s \in H_i] = \frac{1}{\sqrt{2\pi(t-s)}} \int_H e^{-x^2/2(t-s)} dx."$$

Mais precisamente, a equação dada em (2) vem da equação

$$\frac{dX_t}{dt} = f(t, X_t) + g(t, X_t) \text{“ruído”}$$

que passa a ser escrita na forma

$$dX_t = f(t, X_t)dt + g(t, X_t) \text{“ruído” } dt$$

e então

$$dX_t = f(t, X_t)dt + g(t, X_t)dW_t.$$

Finalmente, a interpretação que usamos para (2) é a interpretação segundo Itô (ver Øksendal, [32]), significando que o processo X_t é uma solução da equação integral

$$X_T = X_0 + \int_0^T f(t, X_t)dt + \int_0^T g(t, X_t)dW_t, \quad (3)$$

na qual a última integral,

$$\int_0^T g(t, X_t)dW_t, \quad (4)$$

que é uma integral estocástica, está tratada neste trabalho segundo a interpretação de Itô (ver capítulo 2). Esta é uma entre várias escolhas razoáveis para se interpretar (4) e um detalhamento sobre integrais de Itô está fora do escopo do trabalho. A solução de uma equação do tipo (2) pode ser imaginada como a descrição matemática do movimento de uma pequena partícula num fluido em movimento e continuamente sujeita a um grande número de pequenos choques moleculares que produzem um comportamento aleatório da partícula, daí tais processos estocásticos serem chamados de processos de difusão. Veja Øksendal [32].

No contexto da aplicação deste trabalho, a pequena partícula é considerada como a taxa de juros de curto prazo, o meio no qual ela está imersa é o mercado financeiro e os choques aleatórios são a infinidade de fatores que influenciam a taxa.

O objetivo deste trabalho é estimar o Coeficiente de Tendência de duas EDEs: uma é conhecida na literatura de processos estocásticos como o Processo de Difusão de Ornstein-Uhlenbeck e a outra como o Processo de Difusão de Feller (Cleur e Manfredi, [11]). Na literatura financeira esses mesmos processos são conhecidos como Modelo de Vasicek e Modelo de Cox, Ingersoll e Ross (CIR), respectivamente. Suas expressões são dadas por: Vasicek:

$$dX_t = \kappa(\nu - X_t)dt + \sigma dW_t, \quad \alpha > 0, \kappa > 0 \quad (5)$$

CIR:

$$dX_t = \kappa(\nu - X_t)dt + \sigma\sqrt{X_t}dW_t, \quad \alpha > 0, \kappa > 0 \quad (6)$$

Embora o problema de se estimar o Coeficiente de Difusão também seja de interesse e esteja relacionado com o análogo do Coeficiente de Tendência, em nosso trabalho o procedimento para se obter as estimativas de κ e ν não leva em consideração o conhecimento a priori de σ , conforme adotado em alguns outros trabalhos, como em Cleur e Manfredi, [11]. Como procuramos tratar de um problema relevante do ponto de vista prático, em nossas simulações as estimativas são obtidas como se não conheçêssemos o verdadeiro valor de σ , usado apenas na geração dos processos.

O restante deste trabalho foi dividido em quatro capítulos, da forma como segue. No capítulo 1 fazemos uma revisão da literatura existente, relacionando algumas pesquisas com abordagens variadas, subdivididas nas abordagens paramétrica e não-paramétrica. O capítulo 2 apresenta conceitos básicos, mas importantes do Cálculo Estocástico, relevantes para o desenvolvimento deste trabalho. O resultado chave utilizado não só em nosso trabalho como também frequentemente nessa área de pesquisa é a Fórmula de Itô. O capítulo 3 apresenta a descrição da Estrutura a Termo das Taxas de Juro¹, além de outros conceitos de finanças relacionados à essa área de aplicação. O capítulo 4 trata das simulações, em que os experimentos comparam dois métodos de estimação para um dos parâmetros que fazem parte do Coeficiente de Tendência dos modelos de EDE aqui abordados. O capítulo 5 ilustra a abordagem utilizada neste trabalho com aplicações à alguns conjuntos de dados reais de taxas de juros. Por fim, no capítulo 6, temos as considerações finais e idéias de trabalhos futuros.

¹Ver seção 3.3, página 44

Capítulo 1

Revisão de Literatura

O objetivo deste capítulo é relacionar o maior número possível de trabalhos existentes na literatura dessa área, até o momento do fechamento de nosso trabalho, de modo a propiciar aos leitores uma visão geral satisfatória em termos da variedade de metodologias já empregadas. O problema da inferência para processos estocásticos de parâmetro de tempo contínuo, assim como em muitos outros casos, também admite as abordagens paramétrica e não-paramétrica.

É importante salientar que nenhum aprofundamento foi feito nessas referências, embora algumas delas mereçam certo destaque, pela atualidade ou pela contribuição que trazem aos pesquisadores teóricos e também àqueles com interesses mais voltados para a prática. São eles os trabalhos de Durham e Gallant [14], Jiang e Knight [25] e Fan e Zhang [17], os quais têm suas linhas gerais apresentadas nas duas próximas seções.

1.1 Abordagens Paramétricas

No trabalho de Chan et. al.,[10] uma variedade de modelos em tempo contínuo para a taxa de juro de curto prazo é estimada e comparada empiricamente utilizando-se o Método dos Momentos Generalizado. Encontra-se que os melhores modelos para explicar a dinâmica da taxa de juro de curto prazo são aqueles que permitem que a volatilidade (denominação empregada livremente em finanças para se referir à variabilidade) das ta-

xas mude, sendo assim modelos mais sensíveis ao nível da taxa livre de risco. Alguns modelos mais conhecidos mostram-se pobres nas comparações, devido às suas restrições implícitas sobre a volatilidade da estrutura a termo¹. Os resultados destas comparações têm implicações importantes no uso dos diferentes modelos para a estrutura a termo das taxas.

Não obstante a variedade de modelos, relativamente pouco se sabia sobre como eles se comparavam em termos de suas capacidades de capturar o comportamento real da taxa de juro de curto prazo livre de risco. A principal razão para isto foi, provavelmente, a falta de uma estrutura geral comum na qual os diferentes modelos pudessem ser aninhados como casos particulares. Desse modo os trabalhos empíricos nesta área focavam modelos específicos e não as comparações entre eles porque sem uma estrutura comum era difícil avaliar desempenhos relativos de um modo consistente.

Dois novos métodos são desenvolvidos com a finalidade de realizar inferência estatística em modelos económicos dinâmicos não lineares em Smith [38]. Ambos requerem pouca tratabilidade analítica, sendo baseados na simulação numérica do comportamento dinâmico dos modelos.

Embora um dos estimadores seja assintoticamente mais eficiente que o outro, um estudo de Monte Carlo mostra que, para uma aplicação específica, o estimador menos eficiente possui erro quadrático médio menor em amostras de tamanho tipicamente encontradas na macroeconomia. O estimador com performance superior em amostras pequenas é usado para estimar um modelo de ciclo de negócios real usando dados de séries temporais americanas. Tais métodos são chamados, respectivamente, o Método Estendido dos Momentos Simulados (EMSM, em inglês) e Quase Máxima Verossimilhança Simulada (SQML, em inglês).

O trabalho de Gourieroux et. al. [20] apresenta métodos de inferência que se baseiam em um critério “incorreto”, no sentido de que a otimização de tal critério não fornece diretamente um estimador consistente do parâmetro de interesse. O argumento desse

¹Ver seção 3.3, página 44

critério, chamado parâmetro auxiliar, pode ter uma dimensão maior do que a do parâmetro de interesse. Um segundo passo, baseado em simulações, fornece um estimador consistente e assintoticamente normal para o parâmetro de interesse. Vários testes são propostos e os métodos descritos nesse artigo exigem apenas que o modelo possa ser simulado, sendo, por esta razão, úteis para modelos cuja complexidade inviabilizam uma abordagem direta. Áreas como microeconomia, macroeconomia e finanças são sugeridas como campos de aplicação.

A teoria da estrutura a termo dos preços de títulos indexados devida a Cox, Ingersoll e Ross (CIR) é testada em Gibbons e Ramaswamy [19]. O método econométrico usado é o Método Generalizado dos Momentos e explora a distribuição de probabilidade da variável de estado do modelo CIR, que por ser um modelo de uma única variável (no caso a taxa de juros instantânea) permite testes empíricos sobre a dinâmica da estrutura a termo.

O método permite estimar um modelo em tempo contínuo baseado em dados amostrais discretos. Os testes indicam que o modelo CIR para índices de títulos desempenham razoavelmente bem quando confrontados com retornos de Notas do Tesouro. As estimativas indicam que as curvas de rendimento podem tomar várias formas. Entretanto, os modelos ajustados são pobres para explicar a correlação serial nos retornos reais das Notas do Tesouro.

Em Danielsson [13] o modelo de volatilidade estocástica é usado para estimar a dinâmica de preços diários de ativos. Devido às grandes dimensões das integrais a serem avaliadas para encontrar a densidade marginal dos preços, nenhuma técnica convencional é aplicável. Um método Monte Carlo chamado de Máxima Verossimilhança Simulada é então usado para obter a densidade marginal, onde a variável latente é simulada condicionada à informação disponível. O modelo é estimado com base em 2022 observações diárias do índice Standard and Poor 500. Para comparação, modelos ARCH são estimados com os mesmos dados.

Dois resultados importantes são demonstrados nesse artigo. Primeiro, a estrutura dinâmica dos preços de ativos é captada pelo modelo de volatilidade estocástica e a estimação do modelo, por máxima verossimilhança simulada para o conjunto de dados em

consideração, é bem sucedida. O modelo de volatilidade estocástica dinâmico consegue capturar a não normalidade dos preços de ativos. Segundo, máxima verossimilhança simulada é um método viável para estimar modelos de variáveis latentes dinâmicos, tais como os modelos de volatilidade estocástica dinâmicos. As estimativas Monte Carlo de uma verossimilhança marginal com erro de 0,01% são obtidas com a avaliação de uma integral múltipla de dimensão 2022.

Num contexto onde a literatura financeira atual já estava também muito relacionada com as taxas de juro de curto prazo, Broze et. al., [7] dá uma forma geral envolvendo os modelos mais usuais e derivam uma versão de tempo discreto bem especificada. As propriedades ergódicas são estudadas para se construir um procedimento econométrico consistente baseado numa abordagem de máxima verossimilhança e uma comparação empírica é feita usando dados de notas do tesouro americano.

Por fim uma estratégia de estimação baseada em um método de estimação indireta com a mesma idéia desenvolvida em Gouriéroux et. al., [20] é examinada. Essa estimação indireta resolve o problema do vício de discretização, que normalmente é um problema quando os dados são observações de tempo discreto de uma trajetória amostral contínua. A questão de estimar modelos de tempo contínuo baseados em observações de tempo discreto é um dos interesses deste trabalho e o modelo de Brennan-Schwartz ($dX_t = \kappa(\nu - X_t)dt + \sigma X_t dW_t$), um Movimento Browniano Geométrico com termo de longo prazo positivo, é bastante explorado.

Broze et. al. [7] ainda conclui que um vício está presente e que as estimativas usuais de máxima verossimilhança ou por momentos generalizados, baseados em modelos de tempo discreto, são inconsistentes. Conseqüentemente, valores ajustados por meio destes procedimentos de estimação deveriam ser manuseados com cuidado, especialmente quando são usados para a precificação de derivativos.

O trabalho de Bianchi e Cleur [5] considera a estimação de alguns modelos de equação diferencial estocástica utilizando o método de inferência indireta proposto em Gouriéroux et. al. [20] e usando dados discretos. A performance do método é analisada via Monte

Carlo. Em particular os modelos de Vasicek e o de Cox-Ingersoll-Ross (CIR), usados em finanças, são analisados. Os resultados mostram a capacidade da estimação indireta para remover o vício resultante da discretização de modelos de tempo contínuo. A simulação Monte Carlo é realizada de acordo com os métodos propostos em Smith [38], entre outros.

No trabalho de Overbeck e Rydén [33] são estudados dois estimadores para os parâmetros do modelo CIR baseados em quadrados mínimos condicionais e o estimador de máxima verossimilhança. As propriedades assintóticas dos vários estimadores são discutidas e suas performances são comparadas num estudo de simulação.

As propriedades estatísticas de estimadores dos parâmetros do coeficiente de tendência para processos de difusão são estudadas por métodos numéricos modernos para equações diferenciais estocásticas em Kloeden et. al. [26]. Esses métodos são particularmente úteis com dados amostrados discretamente, para os quais se pode construir estimadores por aproximações discretas das integrais estocásticas que aparecem nos estimadores de máxima verossimilhança para as difusões observadas continuamente.

Os efeitos da estimação de um processo de Ornstein-Uhlenbeck e de um processo Movimento Browniano Geométrico (ou modelos de Vasicek e de Brennan-Schwartz, respectivamente, conforme conhecidos na literatura financeira), devido a dados mal gerados, é investigado em Broze et. al. [8]. O método é chamado de Inferência Quase Indireta pelo fato de um dos passos envolvidos na inferência indireta exigir a avaliação das densidades de transição do processo de difusão relativo, que não são diretamente computáveis. Dessa forma, uma abordagem natural é usar discretizações do modelo relativo com um passo de discretização ainda menor que o passo das observações para se obter os dados simulados e, assim, o modelo simulado não é o verdadeiro. Um vício assintótico devido a este passo de discretização fixo é uma consequência desse procedimento.

Em Ait-Sahalia [1] modelos paramétricos são testados por meio da comparação de suas densidades de transição paramétricas com as respectivas densidades estimadas não-parametricamente. O modelo de tempo contínuo não é substituído por aproximações discretas, mesmo com os dados sendo em intervalos de tempo discretos. A principal

fonte de rejeição dos modelos é a falta de linearidade do coeficiente de tendência. Em torno do nível médio do processo (a taxa de juros de curto prazo esperada) a tendência é essencialmente zero e a taxa de juros instantânea se comporta como um caminho aleatório, enquanto uma forte tendência de reversão à média ocorre quando o processo está afastado desse valor esperado. A volatilidade é maior nessa situação.

Na maioria dos casos em que um processo de difusão é observado em instantes discretos, a densidade de transição e, conseqüentemente, a função de verossimilhança das observações não podem ser calculadas explicitamente. Usando polinômios de Hermite, Ait-Sahalia [3] constrói uma seqüência de funções em forma fechada e mostra que ela converge para a verdadeira (e desconhecida) função de verossimilhança. Esse trabalho documenta que tal aproximação é extremamente acurada como função das variáveis de estado, o que não necessariamente implica que as estimativas resultantes seriam, na prática, próximas das obtidas pelo verdadeiro EMV. Usando simulação de Monte Carlo obtém-se evidências de que esse método supera outros esquemas de aproximação em situações relevantes para modelos financeiros. Em Ait-Sahalia [2] a abordagem citada logo acima é aplicada a alguns modelos de taxas de juro, entre eles o modelo de Vasicek e o modelo CIR.

Outro trabalho relevante seguindo a linha da densidade de transição é o de Jensen e Poulsen, [24], que compara técnicas para aproximar as densidades de transição dos processos de difusão. Esse trabalho considera a aproximação de Euler, a aproximação binomial, métodos baseados em simulação, uma solução numérica das equações diferenciais parciais de Fokker-Plank e aproximações por polinômios de Hermite Ait-Sahalia [3].

Levando em conta os erros absoluto e relativo e os tempos computacionais gastos, Jensen e Poulsen, [24] conclui que a expansão em polinômios de Hermite é o melhor método, sendo bastante superior aos outros três. A segunda melhor abordagem é a que envolve a resolução da equação diferencial parcial usando o método de Crank-Nicolson. A terceira melhor opção é a da aproximação binomial, ficando a simulação da densidade, tanto se usados geradores pseudo aleatórios como quase aleatórios, como o pior deles.

Encontramos em Pedersen [34] uma abordagem conhecida como Máxima Verossimilhança Simulada (SML, em inglês), que tem como base uma sequência de aproximações da função de verossimilhança, com resultados de convergência provados. Segundo esse trabalho a estimação por meio dessa aproximação é fácil e de aplicação geral. A idéia principal desse trabalho é derivar as aproximações com base nas observações discretas, evitando assim os problemas encontrados quando os métodos de estimação para observações discretas são desenvolvidos com base na teoria para observações contínuas.

Essa abordagem pode, em princípio, levar a aproximações da verdadeira função de densidade de transição arbitrariamente boas. Entretanto, estudos posteriores mostraram que o método é computacionalmente pesado (ver Durham e Gallant [14]). Para uma dada realização do processo $dX_t = \mu(X_t, \theta) + \sigma(X_t, \theta)dW_t$, isto é, considerando uma amostra $\{X_i = X(t_i), i = 0, 1, \dots, n\}$, sejam $p(x_t, t; x_s, s)$, para $s < t$, a função densidade de transição de $X(t, \theta) | X(s, \theta)$ e

$$l_n(\theta) = \sum_{i=1}^n \log(p(X_i, t_i | X_{i-1}, t_{i-1}; \theta)) \quad (1.1)$$

o logaritmo da função de verossimilhança. Para aproximar (1.1) considera-se a partição do intervalo $[s, t]$: $s = \tau_1 < \tau_2 < \dots < \tau_M = t$. As variáveis $X(\tau_1), \dots, X(\tau_{M-1})$ não são observadas e devem ser integradas. Como o processo de difusão é markoviano, então a aproximação da verdadeira função de densidade é dada por

$$p^{(M)}(x_t, t; x_s, s, \theta) = \int \prod_{m=0}^{M-1} p^{(1)}(u_{m+1}, \tau_{m+1}; u_m, \tau_m) d\lambda(u_1, \dots, u_{M-1}) \quad (1.2)$$

onde λ é a medida de Lebesgue, $u_0 = x_s$, $u_M = x_t$ e

$$p^{(1)}(x_t, t; x_s, s, \theta) = \phi(x_t; x_s + \mu(x_s)(t - s), \sigma^2(x_s)(t - s)) \quad (1.3)$$

é a sub-densidade de transição, sendo $\phi(x; \mu, \sigma^2)$ a função densidade da distribuição Normal. Dessa forma a aproximação do logaritmo da função de verossimilhança é expresso como

$$l_n^M(\theta) = \sum_{i=1}^n \log[p^M(X_i, t_i | X_{i-1}, t_{i-1}; \theta)] \quad (1.4)$$

Ainda sobre Pedersen [34] é interessante notar que o esquema de discretização usado é o de Euler-Maruyama (ou apenas Euler), que é o mais simples entre a família das aproximações estocásticas de Taylor porque é o único que permite que a densidade de transição da Cadeia de Markov aproximada seja conhecida. Segue que uma convergência mais rápida é esperada quando o coeficiente de difusão da equação diferencial estocástica não depender da variável de estado, dado que, nesse caso, o esquema de Euler é equivalente ao de Milstein, que possui convergência mais rápida.

Santa-Clara [37] desenvolve um novo método econométrico para estimar processos de tempo contínuo a partir de dados amostrados discretamente. Ele estende a técnica de máxima verossimilhança para casos onde a densidade de transição não pode ser encontrada em forma fechada, mas é possível que ela seja calculada por simulação. Propriedades assintóticas do estimador são obtidas, mostrando que ele possui o mesmo comportamento do estimador de máxima verossimilhança verdadeiro em amostras grandes, ou seja, é consistente e assintoticamente normal. O método é então utilizado para estimar modelos de uma ampla família de processos para a estrutura a termo das taxas de juro e algumas restrições desses modelos são testadas.

Durham e Gallant [14] estudam técnicas numéricas desenvolvidas para melhorar o desempenho do tipo de abordagem feita por Pedersen [34]. Dados gerados para o modelo de CIR são usados com um exemplo de teste, pois desde que esse modelo possui função de verossimilhança conhecida em forma fechada, então é possível checar a qualidade dos procedimentos.

Durham e Gallant [14] argumentam que a abordagem de Pedersen [34], baseada em simulação, tem grande apelo teórico, mas que implementações prévias foram computacionalmente caras, no sentido de que o número de trajetórias amostrais a serem geradas de acordo com Pedersen [34], é muito maior do que o requerido de acordo com Durham e Gallant [14]. Em um exemplo, foram necessárias 256 amostras conforme Pedersen [34] contra apenas 8 amostras conforme Durham e Gallant [14].

O problema da eficiência computacional da abordagem proposta por Pedersen [34] é atacado em duas frentes: **primeiro**, melhorar a aproximação da função densidade de

transição dada em (1.2) para se obter uma certa acurácia com menos pontos intermediários no intervalo $[s, t]$; **segundo**, estuda-se técnicas para acelerar a convergência da integração de Monte Carlo usada para avaliar (1.2). Para isso são considerados vários amostradores por importância e uma transformação do modelo de CIR que o torna um processo com a função de difusão constante. Tal transformação representa uma melhoria tanto na acurácia da aproximação da função de verossimilhança quanto na integração numérica de Monte Carlo.

Para ilustrar as novas técnicas, são estimados um modelo simples de volatilidade estocástica (SV, em inglês),

$$\begin{aligned}dX_t &= (\theta_1 + \theta_2 X_t)dt + \theta_3 X_t^{\theta_4} e^{H_t} dW_{1t} \\dH_t &= \theta_5 H_t dt + \theta_6 dW_{2t},\end{aligned}$$

em que dW_{1t} e dW_{2t} são independentes, e um modelo de elasticidade da volatilidade constante (CEV, em inglês),

$$dX_t = (\theta_1 + \theta_2 X_t)dt + \theta_3 X_t^{\theta_4} dW_t,$$

que é a mesma forma geral do modelo já estudado com outra técnica por Chan et. al. [10] (apenas parametrizado de forma diferente). O modelo de elasticidade da volatilidade constante leva tal nome porque a especificação da função de difusão como $\sigma(X_t, \theta) = \theta_3 X_t^{\theta_4}$ é comumente assim conhecida. O conjunto de dados usado foi o das taxas de títulos de 3 meses do Tesouro dos EUA, que conseguimos obter diretamente dos autores, e também é usado na seção 5.1 deste trabalho. As idéias expostas em Durham e Gallant [14] podem ser generalizadas para os casos multivariados.

1.2 Abordagens Não-paramétricas

A teoria da estimação para processos de difusão está bem desenvolvida sob a modelagem paramétrica, quando a forma funcional das funções de tendência e de difusão

são especificadas a menos de um parâmetro de dimensão finita. Em contrapartida, a estimação não-paramétrica não é estudada em detalhe. Os resultados conhecidos se relacionam somente à inferência estatística para modelos de difusão ergódicos com um ruído pequeno ou para um grande tempo de observação T .

Sobre as referências a respeito deste assunto, tem-se Genon-Catalot et. al. [18], que aplicam ondaletas, e Fan e Gijbels [15], onde estimadores localmente polinomiais são descritos. No trabalho de Milstein e Nussbaum [31] considera-se o problema de estimar a função f de um processo de difusão $y(t)$, $t \in [0, 1]$, que satisfaz a uma equação diferencial de Itô da seguinte forma:

$$dy(t) = f(y(t))dt + \varepsilon dW(t), \quad t \in [0, 1], \quad y(0) = 0, \quad (1.5)$$

onde $dW(t)$ é um ruído branco gaussiano e ε é um parâmetro pequeno.

Mostra-se que o modelo (1.5) é assintoticamente equivalente à sua versão discretizada de acordo com o esquema de Euler para a solução de uma equação diferencial estocástica. Nesta versão discretizada, um processo solução aproximado y_i definido sobre uma malha $t_i, i = 1, \dots, n$ é obtido por meio de uma sequência de aproximações sucessivas dada por

$$y_{i+1} = y_i + n^{-1}f(y_i) + \varepsilon n^{-1/2}\xi_i, \quad i = 1, \dots, n, \quad y_1 = 0, \quad (1.6)$$

onde ξ_i é uma variável Normal padrão e n^{-1} é o passo entre t_i e t_{i+1} .

A estimação simultânea das funções de tendência e de difusão é discutida, entre outros, nos artigos de Hall e Carroll [21], Hardle e Tsybakov [22], Ruppert et. al. [36] e Fan e Yao [16].

No último trabalho citado acima (Fan e Yao [16]), sob suposições gerais, que incluem modelos de séries temporais não lineares como um caso especial, propõe-se um método eficiente para estimar a variância condicional. Aplica-se uma regressão linear local ao quadrado dos resíduos e, sem conhecer a função de regressão, estima-se a variância condicional assintoticamente tão bem quanto se a regressão fosse conhecida. Uma das suposições sob as quais se estabelece este resultado assintótico é a de que as observações vêm de um processo estacionário e isto é verificado via simulação.

A suposição de estacionaridade pode ser um tanto restritiva para propósitos práticos. Neste sentido destaca-se que esta suposição é atendida num sentido local, ou seja, tomando-se um intervalo de tempo de tal modo que o processo seja estacionário nesse intervalo. A inferência para estes casos exige o estudo de propriedades não assintóticas dos procedimentos adotados.

Mercurio e Spokoiny [30] oferecem um exemplo de estimação não-paramétrica para modelos de volatilidade estocástica sob estacionaridade local. A abordagem tem como principal problema encontrar este intervalo de homogeneidade de tempo. Um algoritmo para conseguir tal intervalo de tempo, de forma que a volatilidade possa ser estimada apenas por uma média, é estudado teoricamente e também via simulação Monte Carlo.

Uma comparação com um modelo GARCH(1,1) para a estimação e a previsão da volatilidade de séries temporais financeiras é mostrada. O método se presta bem à estimação e à previsão de curto prazo, sendo competitivo e até mesmo superando modelos paramétricos padrões para a previsão quando se trata de um horizonte pequeno.

Em estimação não-paramétrica de curvas a análise estatística pode estar centrada na inferência da estrutura qualitativa da curva analisada. Lepski et. al. [28] propõe um novo estimador do tipo Núcleo com seletor de janela variável. Este seletor leva a taxas de convergência ótimas. O procedimento é uma alternativa razoável à estimação via ondaletas. Na verdade, trata-se de uma modificação do procedimento de estimação adaptativa proposto por Lepski [27].

O modelo considerado neste artigo é o modelo ruído branco,

$$dY(t) = f(t)dt + \sigma dW(t), \quad 0 \leq t \leq 1, \quad (1.7)$$

em que $\{W(t), \quad 0 \leq t \leq 1\}$, é um Movimento Browniano e f é uma função desconhecida. Este modelo também é considerado em Lepski e Spokoiny [29], onde a estimação adaptativa ótima em um ponto de um conjunto de dados com ruído é estudada. Dois procedimentos são apresentados, ambos assintoticamente ótimos para situações diferentes.

No primeiro deles estuda-se o problema da seleção de janela na estimação não-paramétrica pontual usando núcleo estimadores com uma função núcleo dada. A otimalidade do

seletor proposto é provada no sentido assintótico. Esta otimalidade é válida não somente entre núcleo estimadores com janela variável, mas entre todos os estimadores viáveis. Um aspecto importante deste procedimento é o seu funcionamento completamente adaptativo, que serve para uma ampla classe de funções obedecendo restrições de regularidade moderadas (Lepski e Spokoiny [29]).

Ainda sobre este artigo, o segundo procedimento pode ser visto como um caso particular do primeiro sob a suposição de que a função a ser estimada pertença a alguma classe Hölder $\Sigma(\beta, L)$, com β e L parâmetros desconhecidos. Esta suposição permite escolher uma família de núcleos e o procedimento resultante é assintoticamente ótimo no sentido adaptativo em qualquer amplitude de adaptação com $\beta \leq 2$.

Em Spokoiny [39] o procedimento de seleção da janela ótima é adaptado a um suavizador localmente linear para a estimação pontual da função de tendência f em um modelo de difusão do tipo (2). Esta abordagem é não-paramétrica no sentido de que não se especifica uma forma funcional para as funções f e g , que dependem da variável de estado. Os resultados sobre a qualidade da estimativa são não assintóticos e não exigem quaisquer propriedades ergódicas ou mistas do processo observado.

No trabalho de Jiang e Knight [25] a abordagem não-paramétrica se dá por meio do uso de núcleo estimadores. São dois os objetivos desse trabalho: 1) a identificação e a estimação das funções de tendência e de difusão de um processo de difusão de Itô, baseado em amostras de observações discretas, e 2) a ilustração da técnica com uma aplicação às taxas de juro de curto prazo diárias de títulos canadenses com maturidade de 3 meses. Jiang e Knight [25] consideram o modelo

$$dX_t = \mu(X_t)dt + \sigma(X_t)dW_t, \quad X_{t_0} = X, \quad (1.8)$$

em que $\{W_t, t \geq t_0\}$ é um movimento Browniano e os termos de tendência e difusão são dados, respectivamente, por:

$$\mu(x_t) = \lim_{h \rightarrow 0} E \left\{ \frac{X_{t+h} - X_t}{h} \middle| X_t = x_t \right\}$$

$$\sigma^2(x_t) = \lim_{h \rightarrow 0} E \left\{ \frac{(X_{t+h} - X_t)^2}{h} \middle| X_t = x_t \right\}.$$

Fazendo algumas hipóteses sobre as funções de tendência e de difusão, Jiang e Knight [25] concluem que a função densidade de transição é a solução única das equações retrospectiva e prospectiva de Kolmogorov e, dessa forma, as funções de densidade de transição de (1.8) são completamente caracterizadas pelas funções $\mu(\cdot)$ e $\sigma^2(\cdot)$.

Uma maneira alternativa apresentada pelo trabalho de Jiang e Knight [25] é identificar primeiro a função de difusão, pois segundo os autores assim se evita uma série de restrições devidas a outras suposições que se deve fazer sobre os parâmetros da EDE. Dessa forma a função de tendência é estimada depois, usando as estimativas de $\sigma^2(\cdot)$ e da função de densidade marginal $p(X_t)$ do processo.

Estimador Não-paramétrico da Função de Difusão $\sigma^2(\cdot)$

Mais duas condições precisam ser impostas sobre a EDE em (1.8) a fim de se estimar a função de difusão. Como afirmado nesse trabalho de Jiang e Knight, o estimador não-paramétrico para $\sigma^2(\cdot)$ com base em observações X_t no intervalo de tempo $[0, T]$, com $0 < T_0 \leq T$ (T_0 uma constante positiva) e $\Delta_n = \frac{T}{n}$, quando a trajetória amostral do processo de difusão passa por x , é dado por

$$S_n = \frac{\sum_{i=1}^{n-1} nK \left(\frac{X_{i\Delta_n} - x}{h_n} \right) [X_{(i+1)\Delta_n} - X_{i\Delta_n}]^2}{\sum_{i=1}^n TK \left(\frac{X_{i\Delta_n} - x}{h_n} \right)} \quad (1.9)$$

sendo $\{h_n > 0, n = 1, 2, \dots\}$ uma sequência positiva tal que $nh_n \rightarrow \infty$ e $h_n \rightarrow 0$ quando $\Delta_n \rightarrow 0$ e K uma função núcleo.

Estimador Não-paramétrico da Função de Tendência $\mu(\cdot)$

A estimativa de $\mu(\cdot)$ no ponto x é obtida pela equação

$$\mu(x) = \frac{1}{2} \left[\frac{d\sigma^2(x)}{dx} + \sigma^2(x) \frac{p'(x)}{p(x)} \right], \quad (1.10)$$

que depende, por sua vez, do coeficiente de difusão, da função densidade marginal $p(X_t)$ e da derivada primeira de $p(X_t)$, ambas avaliadas no ponto $X_t = x$. O último passo para se obter o estimador não-paramétrico de Jiang e Knight para $\mu(\cdot)$ é definir uma aproximação consistente para $\frac{p'(x)}{p(x)}$ e daí utilizar a equação (1.10), pois S_n em (1.9) é diferenciável. Assim, estima-se $Q(x) = \frac{p'(x)}{p(x)}$ também a partir das observações discretas por meio do estimador

$$q_n(x) = \frac{\sum_{i=1}^n \frac{1}{h_n} K' \left(\frac{X_{i\Delta_n} - x}{h_n} \right)}{\sum_{i=1}^n K \left(\frac{X_{i\Delta_n} - x}{h_n} \right)} \quad (1.11)$$

dado que a trajetória amostral do processo passa por x e $p(x) \neq 0$.

Finalmente, o estimador consistente de $\mu(x)$ é obtido substituindo em (1.10) as quantidades $\sigma^2(x)$ por $S_n(x)$ e $Q(x)$ por $q_n(x)$. Daí:

$$\hat{\mu}(x) = \frac{1}{2} \left[\frac{dS_n(x)}{dx} + S_n(x) q_n(x) \right] \quad (1.12)$$

Passando agora ao artigo de Stanton [40], este autor considerou um processo de difusão X_t satisfazendo a EDE (1.8) e argumentou que sob restrições convenientes sobre $\mu(\cdot)$ e $\sigma(\cdot)$ uma função f arbitrária pode-se reescrever a esperança condicional

$$E_t[f(X_{t+\Delta}, t + \Delta)]$$

na forma de uma expansão em série de Taylor

$$\begin{aligned} E_t[f(X_{t+\Delta}, t + \Delta)] &= f(X_t, t) + \mathcal{L}f(X_t, t)\Delta + \frac{1}{2}\mathcal{L}^2f(X_t, t)\Delta^2 + \dots \\ &\dots + \frac{1}{n!}\mathcal{L}^n f(X_t, t)\Delta^n + O(\Delta^{n+1}), \end{aligned} \quad (1.13)$$

onde \mathcal{L} é o gerador infinitesimal do processo $\{X_t\}$, representado em Stanton [40] por

$$\begin{aligned} \mathcal{L}f(x, t) &= \lim_{\tau \downarrow t} \frac{E[f(X_\tau, \tau) | X_t = x] - f(x, t)}{\tau - t} \\ &= \frac{\partial f(x, t)}{\partial t} + \frac{\partial f(x, t)}{\partial x} \mu(x) + \frac{1}{2} \frac{\partial^2 f(x, t)}{\partial x^2} \sigma^2(x) \end{aligned} \quad (1.14)$$

(para detalhes ver Øksendal [32]). A idéia de Stanton [40] é usar a equação (1.13) para estimar o lado esquerdo, já que as funções $\mu(\cdot)$ e $\sigma(\cdot)$ são desconhecidas; exatamente o oposto do uso mais comum, que seria construir aproximações numéricas para o lado esquerdo da mesma, conhecendo-se as funções $\mu(\cdot)$ e $\sigma(\cdot)$.

Em particular, Stanton [40] trabalhou com as aproximações de primeira, segunda e terceira ordens para as funções de tendência e de difusão, cujos passos mostrados em seu trabalho são descritos resumidamente a seguir. A equação (1.13) pode ser reescrita como

$$\begin{aligned} \mathcal{L}f(X_t, t) &= \frac{1}{\Delta} E_t[f(X_{t+\Delta}, t + \Delta) - f(X_t, t)] - \frac{1}{2} \mathcal{L}^2 f(X_t, t) \Delta - \\ &\quad - \frac{1}{6} \mathcal{L}^3 f(X_t, t) \Delta^2 - \dots \end{aligned} \quad (1.15)$$

e a partir da equação (1.15) e algumas manipulações, as aproximações são obtidas. A de primeira ordem

$$\mathcal{L}f(X_t, t) = \frac{1}{\Delta} E_t[f(X_{t+\Delta}, t + \Delta) - f(X_t, t)] + O(\Delta), \quad (1.16)$$

a de segunda ordem

$$\begin{aligned} \mathcal{L}f(X_t, t) &= \frac{1}{2\Delta} \{4E_t[f(X_{t+\Delta}, t + \Delta) - f(X_t, t)] - \\ &\quad - E_t[f(X_{t+2\Delta}, t + 2\Delta) - f(X_t, t)]\} + O(\Delta^2), \end{aligned} \quad (1.17)$$

e, por fim, a de terceira ordem

$$\begin{aligned} \mathcal{L}f(X_t, t) &= \frac{1}{6\Delta} \{18E_t[f(X_{t+\Delta}, t + \Delta) - f(X_t, t)] - \\ &\quad - 9E_t[f(X_{t+2\Delta}, t + 2\Delta) - f(X_t, t)] + \\ &\quad + 2E_t[f(X_{t+3\Delta}, t + 3\Delta) - f(X_t, t)]\} + O(\Delta^3). \end{aligned} \quad (1.18)$$

Para se obter a aproximação de uma das três primeiras ordens de uma função $g(x, t)$, basta considerar $\mathcal{L}f(X_t, t) = g(x, t)$ em uma das equações (1.16) - (1.18). Assim, pela definição de \mathcal{L} , para as aproximações de $\mu(\cdot)$, faz-se $\mathcal{L}f(x, t) = \mu(x)$, o que equivale a tomar $f(x, t) = x$. Logo, as aproximações para a função de tendência $\mu(\cdot)$ são:

primeira ordem

$$\mu(X_t) = \frac{1}{\Delta} E_t[X_{t+\Delta} - X_t] + O(\Delta), \quad (1.19)$$

segunda ordem

$$\mu(X_t) = \frac{1}{2\Delta} \{4E_t[X_{t+\Delta} - X_t] - E_t[X_{t+2\Delta} - X_t]\} + O(\Delta^2), \quad (1.20)$$

e terceira ordem

$$\begin{aligned} \mu(X_t) = \frac{1}{6\Delta} \{18E_t[X_{t+\Delta} - X_t] - 9E_t[X_{t+2\Delta} - X_t] + \\ + 2E_t[X_{t+3\Delta} - X_t]\} + O(\Delta^3). \end{aligned} \quad (1.21)$$

Analogamente, para construir as aproximações para a função de difusão $\sigma(\cdot)$, toma-se $f(x, t) = (x - X_t)^2$, o que leva a $\mathcal{L}f(X_t, t) = \sigma^2(X_t)$. Daí tem-se a aproximação de primeira ordem

$$\sigma^2(X_t) = \frac{1}{\Delta} E_t[(X_{t+\Delta} - X_t)^2] + O(\Delta), \quad (1.22)$$

a aproximação de segunda ordem

$$\sigma^2(X_t) = \frac{1}{2\Delta} \{4E_t[(X_{t+\Delta} - X_t)^2] - E_t[(X_{t+2\Delta} - X_t)^2]\} + O(\Delta^2), \quad (1.23)$$

e a aproximação de terceira ordem

$$\begin{aligned} \sigma^2(X_t) = \frac{1}{6\Delta} \{18E_t[(X_{t+\Delta} - X_t)^2] - 9E_t[(X_{t+2\Delta} - X_t)^2] + \\ + 2E_t[(X_{t+3\Delta} - X_t)^2]\} + O(\Delta^3). \end{aligned} \quad (1.24)$$

Para calcular todas as esperanças a partir dos dados, Stanton [40] procedeu não-parametricamente, usando o método do núcleo estimador tendo como função núcleo a densidade Gaussiana. Investigou o desempenho dessas aproximações para alguns modelos paramétricos para as taxas de juro mais conhecidos e, segundo o autor²:

... encontramos que com dados diários até as aproximações de primeira ordem, as mais simples, são quase indistinguíveis das verdadeiras funções sobre um amplo intervalo de valores. Quando a frequência da amostragem diminui, a performance de todas as aproximações deteriora. Contudo, as aproximações de ordem mais alta permanecem indistinguíveis das funções verdadeiras mesmo com dados mensais. Isto sugere que os erros de aproximação introduzidos deveriam ser menores quanto mais a série estudada é observada com frequência mensal ou maior.

²Texto original em Inglês

Em seu trabalho, Stanton [40] ainda conclui que a parametrização da função de tendência como uma função linear não é sustentada pelas suas estimativas, enquanto parece não haver problema com a função de difusão. Indica, ainda, que a metodologia de aproximações via expansão de Taylor de ordens superiores, que nesse trabalho foram estimadas com uma abordagem não-paramétrica, garante erros de estimativas cada vez menores.

Fan e Zhang [17], em um trabalho mais recente, abordam as questões 1) as aproximações de ordens superiores funcionam melhor do que as de ordens inferiores? e 2) há procedimentos formais e razoáveis que ajudam a determinar se a não linearidade observada na função de tendência é real ou devida à variações? Fan e Zhang [17] continuam, em parte, a metodologia de Stanton [40], tratando de aproximações de ordens superiores em uma estrutura geral, com a finalidade de se avaliar o impacto das aproximações de ordens superiores na variância dos estimadores.

Com respeito à primeira questão, Fan e Zhang [17] mostram que os estimadores construídos por aproximações de ordem superior reduzem os erros de aproximação numérica nos vícios assintóticos, porém as variâncias assintóticas aumentam exponencialmente, aproximadamente. Para tentar responder a segunda questão, Fan e Zhang [17] usam um ajuste linear local baseado na aproximação de primeira ordem proposta por Fan e Yao [16]. Os autores conseguem melhorar o efeito de fronteira do método do núcleo estimador e construir testes formais para testar modelos financeiros paramétricos contra alternativas não-paramétricas.

Suas simulações mostram que o método linear local sobressai ao método do núcleo estimador. Outro resultado desse artigo é que o então chamado “teste de razão de verossimilhança generalizado” não-paramétrico é poderoso na detecção de alternativas não-paramétricas. Tal teste mostra que a evidência contra a linearidade da função de tendência da taxa de juro de curto prazo é fraca, sendo forte a evidência contra uma família popular de modelos para a função de volatilidade.

Fan e Zhang [17] concluem que Stanton [40] dá uma conclusão algo incompleta sobre as aproximações, pois lhe faltou notar o aumento da variância dos estimadores conforme

se aumenta a ordem da aproximação. Apontam também que esse fenômeno não acontece apenas em ajustes não-paramétricos, sendo devida, em parte, à aproximação de Taylor estocástica.

gregorio-regis-icex-ufmg

Capítulo 2

Cálculo Estocástico

2.1 Conceitos Preliminares

Dentre outras referências sobre as quais apoiamos este trabalho, destacamos Øksendal, [32], como referência sobre cálculo estocástico e Hull, [23], que aborda os conceitos de derivativos financeiros com um tratamento quantitativo sofisticado utilizando o cálculo estocástico. Quando se trata da existência de $\int_0^T g(X_t)dW_t$ em algum sentido, duas opções são as mais utilizadas: a interpretação de Itô e a de Stratonovich. Dessa forma, segundo a interpretação adotada, pode-se referir também ao *cálculo (estocástico) de Itô* ou ao *cálculo (estocástico) de Stratonovich*. Ainda antes de prosseguirmos com este assunto, precisamos de conceitos matemáticos básicos importantes e, logo a seguir, veremos as idéias básicas que norteiam as definições de Itô e Stratonovich.

Definição 2.1.1. (*σ -álgebra*)

Seja Ω um conjunto e seja \mathcal{F} uma família de subconjuntos de Ω . \mathcal{F} é uma σ -álgebra se, e somente se:

(i) $\emptyset \in \mathcal{F}$.

(ii) $F \in \mathcal{F} \Rightarrow F^C \in \mathcal{F}$.

(iii) $A_1, A_2, \dots \in \mathcal{F} \Rightarrow A := \bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$.

Definição 2.1.2. (*σ -álgebra gerada por uma família de conjuntos*)

Dada uma família \mathcal{U} de subconjuntos de Ω , existe uma menor σ -álgebra $\mathcal{H}_{\mathcal{U}}$ que contem \mathcal{U} . Esta σ -álgebra é dada por $\mathcal{H}_{\mathcal{U}} = \cap \{\mathcal{H}; \mathcal{H} \text{ é } \sigma\text{-álgebra de } \Omega, \mathcal{U} \subset \mathcal{H}\}$ e é chamada de σ -álgebra gerada por \mathcal{U} .

Definição 2.1.3. (*Espaço mensurável*)

Diz-se que o par (Ω, \mathcal{F}) , definido por um espaço amostral e uma σ -álgebra definida sobre Ω , é um espaço mensurável.

Definição 2.1.4. (*Medida de probabilidade*)

Seja uma função $P: \mathcal{F} \rightarrow [0,1]$ tal que

(i) $P(\emptyset) = 0$ e $P(\Omega) = 1$ e

(ii) Se $A_1, A_2, \dots \in \mathcal{F}$ com $A_i \cap A_j = \emptyset$ para quaisquer $i \neq j$ implicar $P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$, então P é uma medida de probabilidade sobre o espaço mensurável (Ω, \mathcal{F}) .

Definição 2.1.5. (*Espaço de probabilidade*)

Diz-se que a tripla (Ω, \mathcal{F}, P) , definida por um espaço amostral, uma σ -álgebra e uma medida de probabilidade, é um espaço de probabilidade.

Definição 2.1.6. (*σ -álgebra gerada por uma variável aleatória*)

Dada uma variável aleatória $X: \Omega \rightarrow \mathfrak{R}$, então X gera uma σ -álgebra \mathcal{H}_X sobre Ω de tal modo que \mathcal{H}_X é a menor σ -álgebra que contem todos os conjuntos $X^{-1}(U)$ tais que $U \in \mathcal{B}(U)$ e U é aberto. Ou seja, $\mathcal{H}_X = \{X^{-1}(B); B \text{ é conjunto de Borel}\}$.

Definição 2.1.7. (*Esperança de uma variável aleatória*)

Se $\int_{\Omega} |X(\omega)| dP(\omega) < \infty$, então o número $E[X] := \int_{\Omega} X(\omega) dP(\omega) = \int_{\mathfrak{R}} x d\mu_X(x)$ é chamado de esperança de X (com respeito à medida de probabilidade P do espaço (Ω, \mathcal{F}, P) subentendido). Se $f: \mathfrak{R} \rightarrow \mathfrak{R}$ é mensurável com respeito à σ -álgebra de Borel e $\int_{\Omega} |f(X(\omega))| dP(\omega) < \infty$, tem-se $E[f(X)] := \int_{\Omega} f(X(\omega)) dP(\omega) = \int_{\mathfrak{R}} f(x) d\mu_X(x)$.

Definição 2.1.8. (*Processo estocástico*)

Seja $\{X_t\}_{t \in T}$ uma coleção parametrizada de variáveis aleatórias definida sobre um espaço de probabilidade (Ω, \mathcal{F}, P) e assumindo valores em \mathbb{R} . Diz-se que $\{X_t\}_{t \in T}$ é um processo estocástico. T é o espaço de parâmetro e pode ser $[0, \infty)$, inteiros não negativos ou mesmo subconjuntos de \mathbb{R} .

Podemos notar que para um t fixo tem-se a variável aleatória dada por $\omega \rightarrow X_t(\omega); \omega \in \Omega$. Da mesma forma, fixando um $\omega \in \Omega$, tem-se a função $t \rightarrow X_t(\omega); t \in T$, que é chamada uma *trajetória* de X .

É interessante, para efeito de visualização, imaginar t como o tempo e cada ω como uma partícula. Dessa forma, $X_t(\omega)$ representaria a posição da partícula ω no instante de tempo t . Assim faz até mais sentido chamar X_t de trajetória, pois se estaria pensando no movimento descrito por uma dada partícula ω no tempo. Também é importante não perder de vista que um processo estocástico é uma função de duas variáveis tal que $(t, \omega) \rightarrow X_t(\omega)$, sendo, exatamente por isso, às vezes mais útil escrever $X(t, \omega)$ em vez de $X_t(\omega)$.

2.2 O Cálculo de Itô

Conforme já mencionado na introdução, o cálculo estocástico deste trabalho se faz segundo a interpretação de Itô. Decompondo o intervalo de integração da forma $0 = t_0 < t_1 < \dots < t_n = T$ é natural aproximar $\int_0^T g(X_t) dW_t$ pelo limite, em algum sentido, da soma $\sum_{i=1}^n g(X_{\tau_i})(W_{t_i} - W_{t_{i-1}})$ quando $n \rightarrow \infty$. O problema agora é decidir qual o ponto τ a se considerar no intervalo $[t_{i-1}, t_i]$, pois diferentemente da integral de Riemann-Stieltjes, faz diferença aqui a particular escolha de quais pontos τ serão tomados.

O cálculo de Itô faz $\tau = t_{i-1}$ e assim tem-se somas não-antecipativas porque em cada subintervalo o valor da função integranda é anterior ao incremento do processo. Essa escolha leva a resultados diferentes daqueles que esperaríamos obter segundo as regras do cálculo usual. Um exemplo é $\int_0^T W_t dW_t$. Pelo cálculo determinístico o resultado obtido

seria $\frac{W_T^2}{2}$, mas pelo cálculo de Itô o resultado é $\frac{W_T^2 - T}{2}$. Já o cálculo de Stratonovich faz $\tau = \frac{t_{i-1} + t_i}{2}$, ou seja, τ é o ponto médio do incremento, e assim obtem-se as mesmas regras do cálculo usual, mas este trabalho não utiliza e não aborda o cálculo de Stratonovich.

A Integral de Itô

Para apresentar a integral de Itô de maneira formal, precisaremos de duas definições:

Definição 2.2.1. *Seja $\mathcal{N} = \{\mathcal{N}_t; t \geq 0\}$ uma família crescente de σ -álgebras de subconjuntos de Ω . Um processo $g(t, \omega) : [0, \infty) \times \Omega \rightarrow \mathfrak{R}$ é chamado \mathcal{N} -adaptado se para todo $t \geq 0$ a função $\omega \rightarrow g(t, \omega)$ é \mathcal{N}_t -mensurável.*

A seguir, tem-se a classe de funções para as quais a integral de Itô será definida, de conformidade com Øksendal [32]. Considera-se as funções dessa classe no intervalo de tempo (S, T) .

Definição 2.2.2. *Seja $\mathcal{V} = \mathcal{V}(S, T)$ a classe de funções $f(t, \omega) : [0, \infty) \times \Omega \rightarrow \mathfrak{R}$ tal que:*

- (i) $(t, \omega) \rightarrow f(t, \omega)$ é $\mathcal{B} \times \mathcal{F}$ -mensurável, em que \mathcal{B} é a σ -álgebra de Borel sobre $[0, \infty)$.
- (ii) $f(t, \omega)$ é \mathcal{F} -adaptado.
- (iii) $E \left[\int_S^T (f(t, \omega))^2 dt \right] < \infty$.

Com respeito à construção da integral de Itô, pode-se encontrar em Øksendal [32] um desenvolvimento razoavelmente detalhado. Em nosso escopo apenas daremos as idéias desta construção.

O objetivo é definir o que é uma integral estocástica para as funções f da classe V da última definição, isto é, definir a *Integral de Itô*

$$I[f](\omega) = \int_S^T f(t, \omega) dW_t \quad (2.1)$$

onde W_t é um movimento browniano.

Primeiramente define-se (2.1) para uma classe de funções elementares $\phi \in \mathcal{V}$ dadas por $\phi(t, \omega) = \sum_j e_j(\omega) I_{[t_j, t_{j+1})}(t)$, onde e_j é uma função mensurável, e então mostra-se que uma função $g \in \mathcal{V}$ limitada e contínua para cada ω pode ser aproximada por uma sequência de funções ϕ_n . Desse modo define-se $\int g dW_t$ como o limite de $\int \phi_n dW_t$ quando $\phi_n \rightarrow g$.

O segundo passo é análogo ao descrito acima, porém aproximando funções $h \in \mathcal{V}$, limitadas, por uma sequência de funções $g_n \in \mathcal{V}$ com as condições do primeiro passo.

O terceiro e último passo é aproximar uma função $f \in \mathcal{V}$ qualquer por meio de uma sequência de funções $h_n \in \mathcal{V}$ com as condições do segundo passo. Com estas idéias em mente podemos dar a seguinte definição:

Definição 2.2.3. (*Integral de Itô*)

Seja $f \in \mathcal{V}(S, T)$. Então a integral de Itô de f , de S até T , é definida por

$$\int_S^T f(t, \omega) dW_t = \lim_{n \rightarrow \infty} \int_S^T \phi_n(t, \omega) dW_t \quad (2.2)$$

onde $\{\phi_n\}$ é uma sequência de funções simples tal que

$$E \left[\int_S^T (f(t, \omega) dW_t - \phi_n(t, \omega))^2 dt \right] \rightarrow 0 \quad (2.3)$$

quando $n \rightarrow \infty$.

A Isometria de Itô

Há uma consequência muito importante que decorre da definição acima, chamada Isometria de Itô, que afirma o seguinte: Para toda função $f \in \mathcal{V}(S, T)$, tem-se a identidade

$$E \left[\left(\int_S^T f(t, \omega) dW_t \right)^2 \right] = E \left[\int_S^T f^2(t, \omega) dt \right] \quad (2.4)$$

Algumas propriedades da integral de Itô

A integral de Itô mantém as mesmas propriedades de operador linear do cálculo usual. E além destas, tem-se também as seguintes:

- (i) $E \left[\int_S^T f(t, \omega) dW_t \right] = 0.$
- (ii) $\int_S^T f(t, \omega) dW_t$ é \mathcal{F}_T - mensurável.

Outra boa propriedade da integral de Itô é que ela é um martingal e isto é uma vantagem no aspecto computacional. Abaixo damos a definição de martingal contínuo.

Definição 2.2.4. (Martingal)

Um processo estocástico $\{M_t\}_{t \geq 0}$ sobre um espaço de probabilidade $(\Omega, \mathcal{F}, \mathcal{P})$ é chamado de Martingal, com respeito a um filtro $\{\mathcal{M}_t\}_{t \geq 0}$ (e com respeito à medida de probabilidade P) se:

- (i) M_t é \mathcal{M}_t -mensurável para todo t .
- (ii) $E[|M_t|] < \infty$ para todo t .
- (iii) $E[M_s | \mathcal{M}_t] = M_t$ para todo $s \geq t$.

A Fórmula de Itô

Na realidade a definição em si da integral de Itô não é muito útil para se efetuar os cálculos. Então, da mesma forma como no cálculo determinístico não se calculam as integrais diretamente da definição de integral de Riemann, também não se utiliza diretamente a definição de Itô, mas sim uma regra conhecida como *fórmula de Itô*, que enunciamos a seguir (ver Øksendal [32]):

Teorema 2.2.1. (Fórmula de Itô)

Considere a equação diferencial estocástica $dX_t = f(t, X_t)dt + g(t, X_t)dW_t$ e uma função $G(t, X_t)$ com derivadas parciais contínuas até segunda ordem em $[0, \infty) \times \mathbb{R}$. Então o processo estocástico $Y_t = G(t, X_t)$ satisfaz a seguinte equação diferencial estocástica

$$dY_t = \frac{\partial}{\partial t}G(t, X_t)dt + \frac{\partial}{\partial x}G(t, X_t)dX_t + \frac{1}{2} \frac{\partial^2}{\partial x^2}G(t, X_t) \cdot (dX_t)^2, \quad (2.5)$$

onde $(dX_t)^2 = (dX_t) \cdot (dX_t)$ é calculado de acordo com as regras

$$dt \cdot dt = dt \cdot dW_t = dW_t \cdot dt = 0 \quad e \quad dW_t \cdot dW_t = dt.$$

Capítulo 3

Taxas de Juros

Este capítulo introduz conceitos importantes para o entendimento da área de aplicação que é objeto deste trabalho. Para compreender as taxas de juros é imprescindível fazer também uma abordagem sucinta do mercado financeiro e de alguns elementos que o constituem.

3.1 Formação das Taxas de Juros

Segundo Assaf Neto, [4], “O juro exprime o preço de troca de ativos disponíveis em diferentes momentos do tempo. Em geral, receber uma quantidade de dinheiro hoje é preferível a receber amanhã, e o juro oferecido pela disponibilidade imediata do bem deve remunerar o adiamento de seu uso.”

“Quaisquer que sejam os tipos de taxas conhecidos no mercado, todos eles exprimem a remuneração pela alocação de capital. A taxa de juro reflete, portanto, o preço pago pelo sacrifício de poupar, ou seja, a remuneração exigida por um agente econômico ao decidir postergar o consumo, transferindo seus recursos a outro agente.”

“As relações entre o *poupador* e o *tomador* de recursos implicam, diante do que foi considerado, uma relação entre o poder de compra futuro e o poder de compra presente. Essa relação se mantém interessante até que os juros de mercado deixem de ser atraentes para as decisões temporais dos agentes: o genuíno valor da taxa de juro é definido pelas

operações livremente praticadas no mercado, e não imposto por decisões externas a seu âmbito de atuação.”

“Em essência, a noção básica de taxa de juros está vinculada ao conceito de *taxa preferencial temporal* dos agentes econômicos envolvidos. Se para um indivíduo for indiferente receber \$ 100 hoje ou \$ 112 ao final de um ano, demonstrando disposição em aceitar indiferentemente qualquer das propostas, pode-se afirmar que sua *taxa de preferência temporal* é de 12%.”

“Admitindo-se um mercado livre, regulado por suas próprias forças, a taxa de juro é formada com base nas taxas de preferências temporais dos agentes econômicos possuidores de recursos para empréstimos e no retorno esperado daqueles que demandam recursos.”

“De certa maneira, a taxa de juro exprime a confiabilidade dos agentes econômicos com relação ao desempenho esperado da economia. Pode ser observado que, em momentos de maior instabilidade do ambiente econômico, ocorre certa elevação nas taxas de juros de mercado, como reflexo natural da incerteza associada às decisões de seus agentes.”

3.1.1 Taxa Livre de Risco

“O Estado controla certos instrumentos de política monetária que permitem regular temporariamente os níveis das taxas de juros no mercado, sempre que for julgado recomendável para a economia. Adicionalmente, o Governo tem também o controle exclusivo (monopólio) dos meios de pagamento e emissões de títulos públicos, admitidos como *ativos livres de risco*. Com fundos praticamente ilimitados, o controle das taxas de juros, quando seus percentuais escapam para níveis economicamente indesejados, é estabelecido pelo Governo, forçando ou a procura ou a oferta de seus próprios títulos.”

“Essa taxa de juros que precifica os ativos do Governo no mercado é denominada *taxa pura* ou *taxa livre de risco*, constituindo-se na taxa de juros base do sistema econômico. Essa taxa de juro da economia, por seu lado, pode assumir diversos valores de acordo com o risco oferecido pelas diversas alternativas de aplicações. Assim, apesar de o mercado

estabelecer taxas variáveis para cada nível de risco assumido pelas várias operações financeiras, o denominador comum dessa estrutura é a taxa de juro do sistema econômico.”

“Ou seja, a taxa pura, que embute unicamente a remuneração pelo sacrifício de poupar, constitui-se no piso para a estrutura de taxas de retorno da economia, na qual teremos não apenas a remuneração pelo diferimento do consumo, mas também a remuneração de cada nível de risco envolvido pelas outras alternativas de investimento.”

“Em conclusão, a taxa de juro do sistema econômico é a taxa mínima, estando sempre abaixo dos retornos oferecidos pelos ativos que não sejam títulos governamentais. E a taxa de juro estabelecida livremente pelo mercado é a taxa referencial a ser comparada com os retornos oferecidos pelos investimentos com risco.”

3.2 Títulos Públicos

Ainda de acordo com Assaf Neto, [4], “Os Governos federal, estadual e municipal costumam captar recursos no mercado por intermédio da emissão de títulos representativos da dívida pública. Os títulos públicos federais têm mais aceitação e liquidez.”

“Essencialmente, os títulos públicos emitidos pelo Tesouro Nacional estão voltados para a execução da política fiscal do Governo, antecipando receitas orçamentárias ou financiando déficits fiscais. Os títulos de emissão do Banco Central, de menor duração, têm por objetivo principal a implementação e execução da política monetária. Os títulos da dívida pública do Governo Federal, classificados pela natureza de suas emissões, e suas principais características de negociação, são apresentados a seguir:”

“LTN - Letras do Tesouro Nacional: são negociados com deságio (desconto), pagando o investidor uma quantia inferior a seu valor de face. São emitidos com um prazo mínimo de 28 dias;

LFT - Letras Financeiras do Tesouro: têm seus rendimentos definidos pela média da taxa Selic (*overnight*), garantindo uma rentabilidade de mercado ao investidor. São papéis atraentes, e seus prazos de emissão são definidos pelo Tesouro Nacional;

Tabela 3.1: *Títulos da dívida pública do Governo Federal.*

Títulos do Tesouro Nacional	Títulos do Banco Central
LTN - Letras do Tesouro Nacional	BBC - Bônus do Banco Central
LFT - Letras Financeiras do Tesouro	LBC - Letras do Banco Central
NTN - Notas do Tesouro Nacional	NBC - Notas do Banco Central

Fonte: Assaf Neto, [4], 1999, p.83.

NTN - Notas do Tesouro Nacional: oferecem rendimentos pós-fixados e atrelados a um indexador da economia. Os juros são pagos periodicamente e o prazo mínimo de emissão é de três meses. São lançadas também séries especiais de NTN, pagando rendimentos diferenciados. Esses títulos apresentam opções de rendimentos e prazos diferentes de acordo com seu tipo de emissão;

BBC - Bônus do Banco Central: caracteristicamente de curto prazo e utilizados nos leilões de títulos públicos do Banco Central. Apresentam rendimentos prefixados na forma de desconto (deságio). Os prazos desses títulos são de 28, 35, 42 e 45 dias;

LBC - Letras do Banco Central: são títulos bastante semelhantes às Letras Financeiras do Tesouro (LFT), tendo sua remuneração definida pela taxa média diária (*overnight*) do Selic. São ativos quase-moeda de grande atratividade para os investidores;

NBC - Notas do Banco Central: oferecem rendimentos postecipados e atrelados a um indexador da economia. Os juros costumam ser pagos periodicamente, e seu prazo de emissão mínimo é de três meses. As autoridades monetárias lançaram uma série especial desse título (NTN-E), com correção atrelada à variação do dólar.”

3.2.1 Mercado de Bonds

“O mercado internacional de dívidas permite também o levantamento de recursos por meio da emissão de títulos pelos próprios tomadores. Nesse contexto, destacam-se

duas importantes formas de captações das empresas processadas pela emissão de *bonds* e *commercial papers*.”

Bonds são títulos de renda fixa com maturidade¹ superior a um ano. Esses papéis nada mais são do que notas promissórias emitidas sem garantias reais que pagam juros periódicos a seus proprietários, ou determinado montante fixo ao final do prazo de emissão. Existem no mercado inúmeros títulos de renda fixa, os quais se diferenciam essencialmente pela maturidade (prazo de resgate do título) e natureza do emitente.

“Os *bonds* apresentam atualmente grande expressão nas operações desenvolvidas pelo mercado financeiro internacional, constituindo-se em importante fonte de recursos a longo prazo. São essencialmente títulos de renda fixa, representativos de dívidas de maior *maturidade*, emitidos diretamente pelos tomadores dos recursos. Não costumam apresentar garantia real, podendo eventualmente ser negociados com garantia por fiança bancária. Os emitentes desses títulos são geralmente empresas privadas e, mesmo, Governos e instituições públicas.”

“Os *bonds* prometem geralmente pagamentos periódicos de juros e amortização do principal ao final do prazo de aplicação. O título costuma apresentar ainda uma única data de vencimento, podendo-se eventualmente encontrar *bonds* com opção de resgate antes do vencimento.”

“A remuneração dos *bonds* é basicamente definida com base nas taxas de juros pagas por um *treasury bond* (bond emitido pelo Governo dos EUA), de vencimento comparável, admitidas como livres de risco. Os *bonds* negociados no mercado, emitidos por diferentes instituições, costumam oferecer um prêmio pelo risco como forma de compensar a possibilidade de insolvência do emitente, o denominado *default premium*. Esse prêmio adicional varia de acordo com o prazo e o risco do emitente do título.”

As emissões de títulos de renda fixa com maturidade inferior a um ano são geralmente identificadas como instrumentos de curto prazo do mercado monetário, como títulos públicos, *commercial papers* e certificados de depósitos bancários. Os principais emitentes

¹Prazo para o vencimento

de bonds em uma economia são os Governos (federal, estadual e municipal) e as empresas, as quais são geralmente classificadas em diversos setores de atuação como serviços públicos, instituições financeiras, indústrias, etc.

“Os adquirentes dos bonds emitidos pelas empresas tomadoras de recursos são, em grande parte, investidores institucionais, representados por fundos de pensão, instituições financeiras, Governos, grandes corporações, fundos de investimentos, etc. Diante da expressiva negociação desses títulos no mercado europeu, atribui-se a denominação *euro-bonds* aos bonds emitidos no euromercado.”

3.2.2 Rendimento na Maturidade

Os títulos de dívida de diferentes maturidades têm, cada um, um preço e uma taxa de juro associada, chamada *rendimento na maturidade* (*yield to maturity - YTM*), ou apenas *rendimento*. O conceito de rendimento na maturidade reflete o rendimento efetivo dos títulos de renda fixa até seu vencimento.

A determinação do rendimento considera o preço de mercado do título e os fluxos de caixa associados. Na análise de um bond, se o rendimento superar a taxa de rentabilidade requerida, o investimento é considerado atraente, revelando-se desinteressante quando o retorno prometido for menor que a taxa requerida.

Genericamente, para um bond com pagamentos periódicos de juros e resgate do principal ao final, forma de remuneração usualmente adotada pelo mercado, a expressão para o cálculo do rendimento é da seguinte forma:

$$P_0 = \frac{C_1}{1 + YTM} + \frac{C_2}{(1 + YTM)^2} + \dots + \frac{C_n + P_n}{(1 + YTM)^n} \quad (3.1)$$

onde:

P_0 é o preço corrente de mercado do bond;

C_1, \dots, C_n é o fluxo de pagamentos (cupons periódicos);

P_n é o valor de face do bond;

YTM é o *rendimento*, ou seja, taxa de juro que, ao descontar o fluxo de pagamentos, apura um valor presente igual ao preço corrente de mercado do bond.

O preço de mercado de um bond é bastante sensível às oscilações verificadas nas taxas de juros, variando inversamente à maneira como os juros se movimentam no mercado. O conceito de *rendimento* assume implicitamente o reinvestimento dos fluxos intermediários de caixa à própria taxa prometida pelo investimento. Quando os pagamentos são periodicamente reinvestidos à própria taxa do YTM inerente ao fluxo de caixa, a rentabilidade periódica apurada na equação (3.1) é considerada efetiva.

Um bond pode também retornar um único pagamento no seu vencimento, incluindo os juros e o principal. Esse tipo de título é chamado bond *sem cupom* e o seu *rendimento* é obtido por:

$$P_0 = \frac{C_n + P_n}{(1 + YTM)^n} \quad (3.2)$$

3.2.3 Volatilidade dos Bonds

Ainda seguindo de acordo com Assaf Neto, [4], tem-se que “a volatilidade de um bond é entendida como as mudanças ocorridas nos preços do título em razão de modificações verificadas nas taxas de juros de mercado. É, geralmente, mensurada pelo percentual de variação dos preços dos bonds. Quanto maior a volatilidade de um título, mais elevadas apresentam-se as mudanças em seus preços diante de alterações nas taxas de juros.”

Duas propriedades importantes da relação envolvendo preços dos bonds, maturidade e taxa de juro são as seguintes: (i) Os preços dos bonds têm uma relação inversa com as taxas, conforme já comentado e (ii) A volatilidade dos títulos é menor quanto mais reduzida apresentar-se sua maturidade.

Deve ser observado também que as *variações* nos preços dos títulos, em função do aumento das taxas, também aumentam para uma certa maturidade, mas as variações desses aumentos são decrescentes se considerarmos maturidades maiores. Como exemplo,

considere um título com valor de face de \$ 1.000,00 remunerado à taxa de 10% ao ano e as maturidades de 5, 10 e 20 anos. Considerando a maturidade de 5 anos com as taxas de desconto de 8% e 12%, que produzem valores do bond de \$1.079,85 e \$927,90, respectivamente, tem-se uma variação percentual de -14%. Mantendo fixo, agora, a maturidade de 10 anos, essa variação percentual é de -21,8%, tendo, portanto, aumentado com o aumento da maturidade. No entanto passando da maturidade de 5 anos para a maturidade de 10 anos a variação das variações dos preços é de 55,7%, enquanto que, de 10 para 20 anos, a variação é menor, de 32,6%. Este exemplo pode ser melhor visualizado na tabela abaixo.

Tabela 3.2: *Resultados de um bond para diferentes maturidades e vencimentos.*

Maturidade	5 anos		10 anos		20 anos	
Taxa des- conto	8%	12%	8%	12%	8%	12%
Valor bond	\$ 1.079,85	\$ 927,90	\$ 1.134,20	\$ 887,00	\$ 1.196,36	\$ 850,61
Variação do valor	-14%		-21,8%		-28,9%	

Fonte: Assaf Neto, [4], 1999, p.174.

3.3 A Estrutura a Termo das Taxas de Juros

Observando, em determinado dia, os rendimentos de títulos de diferentes maturidades, pode-se estabelecer a relação entre rendimento e maturidade. Essa relação é chamada **Estrutura a termo das taxas de juros**, em que “a termo” refere-se ao vencimento dos títulos. A representação gráfica da Estrutura a termo é chamada **curva de rendimento**. É uma representação gráfica das taxas de juros de um certo título em diferentes vencimentos.

Para se comparar adequadamente essas taxas, refere-se ao rendimento *anualizado* até

o vencimento. Essa curva, na maioria das vezes, tem inclinação ascendente, significando que as taxas de juros de curto prazo são mais baixas que as taxas de longo prazo. Em algumas vezes ocorre de a curva apresentar inclinação descendente, mesmo que apenas em parte dela. O porquê de ativos idênticos terem rendimentos diferentes e o porquê da mudança na inclinação da curva são questões analisadas com a ajuda de algumas teorias sobre a estrutura a termo das taxas de juros. Entre elas a Teoria das Expectativas se apresenta como a mais importante.

Essa teoria baseia-se na observação de que há duas formas de se investir em um período de N anos: o investidor pode manter um título de N anos, ou adquirir um título de um ano e, no vencimento, reinvestir tudo comprando outro título de um ano e assim por diante durante os N anos. Se não existe incerteza, a **arbitragem** do mercado financeiro irá garantir que as duas escolhas resultarão em retornos idênticos.

Novamente de acordo com Assaf Neto, [4] “Essa teoria propõe que a taxa de juros de longo prazo se constitua numa média geométrica das taxas de curto prazo. As taxas esperadas de curto prazo representam uma projeção não viesada das taxas futuras de juros e os investidores estarão dispostos a tomar suas decisões em relação à maturidade de seus ativos com base nos diferenciais dos rendimentos. Se um título de longo prazo oferece ganhos acima das expectativas, ele passará a substituir outros ativos na composição do portfólio de um investidor.”

3.4 Modelos de Um Fator e com Reversão à Média

Os modelos de Vasicek e de Cox, Ingersoll e Ross são os **modelos de um fator** mais populares. São modelos de um fator porque o movimento da taxa de juro depende de apenas uma variável. A maioria dos modelos dessa formulação oferecem soluções em forma fechada. A tendência e o desvio instantâneos, de acordo com Hull [23], são funções de X_t , mas independentes do tempo. Modelos de um fator implicam que as taxas se movem na mesma direção durante qualquer intervalo pequeno de tempo, mas não na mesma proporção. Isso não implica, como às vezes se supõe, que a estrutura a termo das

taxas de juro tenha sempre o mesmo formato. Pode haver um padrão significativo de estruturas sob o modelo de um fator. A suposição de um fator não é tão restritiva quanto parece.

Ambos apresentam reversão à média: a tendência instantânea $\kappa(\nu - X_t)$ representa o efeito da atração do processo para a sua média de longo prazo ν . Essa propriedade apresentada por estes dois modelos está de acordo com o fenômeno econômico que ocorre com as taxas de juro: elas são atraídas para um valor. Quando as taxas aumentam a economia esfria e daí a demanda por dinheiro também cai, provocando uma tendência de diminuição das taxas. O contrário acontece quando as taxas diminuem: a procura por empréstimos sobe, provocando, conseqüentemente, uma tendência de aumento das taxas.

3.4.1 Modelo de Vasicek

Em 1977 Vasicek [41] propôs o processo de Ornstein-Uhlenbeck, dado por (5), como modelo para a dinâmica da taxa de juros de curto prazo. Este modelo costuma ser criticado por permitir taxas negativas e também pela volatilidade (função de difusão) ser independente do nível das taxas.

Solução Exata

Uma vantagem da equação (5) é o fato de possuir solução em forma fechada, ou seja, em termos da trajetória de X_t , o que facilita a geração do processo. A solução exata de (5) é dada logo abaixo, de acordo com Cleur e Manfredi [11]:

$$X_t = \nu + (X_0 - \nu)e^{-kt} + \sigma \int_0^t e^{ks} dW_s \quad (3.3)$$

3.4.2 Modelo de Cox-Ingersoll-Ross (CIR)

Como o modelo de Vasicek permite que as taxas de juros sejam negativas, Cox, Ingersoll e Ross [12], em 1985, propuseram como alternativa o processo de difusão dado pela equação (6)

Neste modelo, para um valor inicial não negativo da taxa, ela nunca será negativa. Isto se deve ao fato não só da tendência de retorno à média, mas também da diminuição da volatilidade quando o nível X_t se aproxima de zero. As taxas seguindo (6) somente podem alcançar zero quando $\sigma^2 > 2\kappa\nu$.

Função Densidade de Transição

O modelo CIR (6) não possui solução em forma fechada em se tratando do domínio da trajetória, mas a função densidade de transição $p(X_s|X_t)$ é dada por uma distribuição Qui-Quadrado não central com $2q$ graus de liberdade e parâmetro de não centralidade $u(t)$ tal que

$$u(t) = 2 c(t) X_t e^{-\kappa(s-t)}; \quad c(t) = \frac{2\kappa}{\sigma^2(1 - e^{-\kappa(s-t)})}; \quad q = \frac{2\kappa\nu}{\sigma^2}$$

Capítulo 4

Simulação

Este capítulo mostra a aplicação do processo de simulação de Monte Carlo à teoria de estimação de certos processos estocásticos de tempo contínuo utilizada neste trabalho. É interessante notar que, com base nas simulações realizadas até esse momento, há evidências que apontam para a não necessidade de se conhecer o valor do parâmetro de difusão σ . As comparações de alguns casos estão apresentadas.

4.1 Geração de Números Pseudo-aleatórios

O processo de geração de números pseudo-aleatórios é parte de fundamental importância de qualquer estudo que faça uso de simulação. O gerador utilizado em nossos experimentos está descrito em Press et al [35], que traz muitas implementações de algoritmos na linguagem C, para diversas finalidades.

Nosso gerador pertence à família dos *geradores congruenciais multiplicativos*, que geram uma sequência de inteiros I_1, I_2, I_3, \dots , entre 0 e $m - 1$, pela relação de recorrência

$$I_{j+1} = aI_j \pmod{m}, \quad (4.1)$$

onde m e a são chamados, respectivamente, de *módulo* e *multiplicador*. Segundo Press et al [35], essa sequência, eventualmente, irá se repetir com um período menor ou igual a m , sendo que o período máximo ocorre para m e a escolhidos apropriadamente. Este

gerador, proposto em 1969 por Lewis, Goodman e Miller - e relacionado em Press et al [35] -, baseia-se nas seguintes escolhas:

$$m = 2^{31} - 1 = 2147483647 \quad e \quad a = 7^5 = 16807 \quad (4.2)$$

Em Press et al [35] comenta-se que este gerador, apesar de simples, passou por vários testes teóricos novos e, também importante, acumulou uma grande quantidade de usos bem sucedidos. Trata-se de um “padrão mínimo”, simples e portátil, cujo período é da ordem de 10^9 .

4.2 Simulação de Processos de Difusão

Os casos simulados apresentados aqui foram escolhidos devido ao fato de serem mais próximos do que se acredita ser uma situação real, na qual espera-se que o tempo de observação T não seja de muitos anos.

Os estimadores usados são os de Mínimos Quadrados Ordinários (EMQO) e os de Máxima Verossilhança Discretizados (EMVD), de acordo com Cleur e Manfredi [11]. Usou-se a Regra dos Trapézios para calcular as integrais determinísticas numericamente e a Fórmula de Itô para avaliar as integrais estocásticas.

Os modelos usados neste trabalho são o de Vasicek, ver Vasicek [41] e o de Cox-Ingersoll-Ross, ou simplesmente CIR, ver Cox, Ingersoll e Ross [12], bastante conhecidos da literatura de processos estocásticos aplicados à finanças, mais propriamente quando se trata do problema de modelagem da estrutura a termo das taxas de juros¹. Eles aparecem, por exemplo, em Hull [23]. Esses dois modelos de difusão são escritos, neste trabalho, de acordo com (5) e (6), respectivamente.

A tabela (4.1) traz as integrais que definem os estimadores de máxima verossimilhança de tempo contínuo (EMVTC), $\hat{\kappa}$ e $\hat{\nu}$, e suas fórmulas tal como em Cleur e Manfredi [11]:

$$\hat{\kappa}_{mvd} = \frac{I_1 I_5 - I_2 I_3}{I_3 I_4 - I_5^2} \quad (4.3)$$

¹Ver seção 3.3, página 44

Tabela 4.1: *Integrais definindo os EMVTC*

MODELO	I ₁	I ₂	I ₃	I ₄	I ₅
Vasicek	$\int_0^T dX_s$	$\int_0^T X_s dX_s$	$\int_0^T ds$	$\int_0^T X_s^2 ds$	$\int_0^T X_s ds$
CIR	$\int_0^T \frac{dX_s}{X_s}$	$\int_0^T dX_s$	$\int_0^T \frac{ds}{X_s}$	$\int_0^T X_s ds$	$\int_0^T ds$

Fonte: Adaptado de Cleur e Manfredi, [11], 1999.

$$\hat{\nu}_{mvd} = \frac{I_1 I_4 - I_2 I_5}{I_1 I_5 - I_2 I_3} \quad (4.4)$$

A seguir, considerando a discretização da equação (5) tem-se

$$X_{i+1} - X_i = \kappa(\nu - X_i) \Delta t + \sigma(W_{i+1} - W_i),$$

que após tomarmos a esperança, fica

$$E[X_{i+1} - X_i] = \kappa(\nu - X_i) \Delta t,$$

dado que $E[(W_{i+1} - W_i)] = 0$. Daí, arranjando os termos de forma conveniente, obtem-se

$$E[X_{i+1} - X_i] = \kappa\nu\Delta - \kappa\Delta X_i,$$

donde, fazendo $Y_i = E[X_{i+1} - X_i]$, $\beta_0 = \kappa\nu\Delta$ e $\beta_1 = \kappa\Delta$ temos os EMQO:

$$\hat{\kappa}_{mqo} = -\frac{\hat{\beta}_1}{\Delta} \quad e \quad \hat{\nu}_{mqo} = -\frac{\hat{\beta}_0}{\beta_1}, \quad (4.5)$$

em que $\hat{\beta}_0$ e $\hat{\beta}_1$ são obtidos pelo método dos Mínimos Quadrados Ordinários do modelo $Y_i = \beta_0 + \beta_1 X_i$ e Δ é o intervalo de tempo entre observações. Para o modelo (6) os estimadores de κ e ν são exatamente os mesmos que em (4.5), pois os modelos de Vasicek e CIR possuem a mesma função de tendência, que é a única parte que resta das equações após o cálculo da esperança.

Em nossos experimentos, diferentemente de Cleur e Manfredi [11], encontramos as estimativas de κ e ν supondo que o parâmetro σ , que define a função de difusão, não é conhecido. Dessa forma as integrais que envolvem σ são calculadas simplesmente como somas de Riemann, bastando acessar a amostra $\{X_i = X(t_i), i = 0, 1, \dots, n\}$. A integral estocástica I_2 de Vasicek, que tem como solução exata

$$\int_0^T X_s dX_s = \frac{X_T^2 - X_0^2 - \sigma^2 T}{2}$$

passa a ser calculada, aproximadamente, como

$$\int_0^T X_s dX_s \approx \sum_{i=1}^{n-1} X_i (X_{i+1} - X_i).$$

Da mesma forma, a integral estocástica I_1 de CIR, que tem como solução exata

$$\int_0^T \frac{dX_s}{X_s} = \log \left(\frac{X_T}{X_0} \right) + \frac{\sigma^2}{2} I_3$$

passa a ser calculada, aproximadamente, como

$$\int_0^T \frac{dX_s}{X_s} \approx \sum_{i=1}^{n-1} \frac{1}{X_i} (X_{i+1} - X_i).$$

As duas integrais estocásticas aproximadas dessa forma (*integrais naive*) são boas aproximações. Na tabela (4.2) comparamos estimativas obtidas por Cleur e Manfredi [11] (Tabela X, p. 188) usando as soluções exatas de I_1 e I_2 , ou seja, com σ conhecido, com estimativas obtidas neste trabalho usando as integrais I_1 e I_2 naive, ou seja, σ desconhecido.

Tabela 4.2: *EMVDs de κ do modelo de Vasicek; $T = 100$ e $\Delta = 0,1$.*

κ	σ desconhecido	σ conhecido
0,3	0,3361 (0,0856)	0,3342 (0,0840)
0,8	0,8248 (0,1269)	0,8034 (0,1221)

4.2.1 Casos Simulados

Os casos tratados aqui são relacionados a valores fixos de κ ($\kappa = 0,3, 0,8$), de T ($T = 60, 72, 84, 96, 108$ e 120 meses), de ν ($\nu = 0,1$) e do valor inicial X_0 do processo ($X_0 = 0,1$), que é o valor de equilíbrio. Salientamos que tais valores de T foram adotados pensando-se em situações mais realistas. Embora a estimativa de ν seja muito menos viciada do que a estimativa de κ , nós mantivemos seu cálculo como parte do processo, para termos controle sobre os resultados das estimativas.

Nós usamos o Método de Monte Carlo para simular as trajetórias dos modelos considerados, reproduzindo o que foi feito em Cleur e Manfredi [11], porém preferimos 2000 replicações e não 1000, como no artigo citado. O vício, o desvio padrão e o erro quadrático médio das estimativas do coeficiente de tendência também foram obtidos usando o compilador Borland C++ Builder 4.0.

Para o trabalho de simulação foi usado um computador PC, com processador Pentium 4 de 2 GHz, memória RAM de 512 MB e sistema operacional Windows 98. Cada replicação da simulação de Monte Carlo varia de 1,005 segundos a 1,04 segundos, dependendo da quantidade de casos sendo executados simultaneamente. O tamanho das amostras geradas não influi no tempo gasto, apenas o número de replicações.

Tabela 4.3: Valores de Δ e seus respectivos números de observações, considerando o ano financeiro, de 252 dias.

Δ	Espaço de Tempo
0,1	a cada 2 dias
0,04761	1x ao dia
0,01	5x ao dia

As tabelas (4.4), (4.5) e (4.6) exibem os EMVD e de EMQO para o parâmetro $\kappa = 0,3$, para os modelos de Vasicek e CIR e para valores do intervalo de discretização $\Delta = 0,1$, $\Delta = 0,04761$ e $\Delta = 0,01$, respectivamente.

Observa-se que, para um Δ fixo, as estimativas por ambos os métodos e modelos tornam-se menos viesadas e se mostram consistentes conforme T aumenta, o que é ilustrado pelas figuras (4.7) a (4.10). Se observarmos para um mesmo T fixo e diminuirmos Δ o mesmo já não acontece: apenas as estimativas EMVD para o modelo CIR melhoram de acordo com Δ decrescendo, conforme ilustram as figuras (4.11) e (4.12). Para Vasicek o EMVD mostra-se numericamente pouco melhor que o EMQO e para o modelo CIR, o EMQO é pouco melhor para valores maiores de Δ , piorando conforme Δ diminui.

Pode-se notar também que para Vasicek as estimativas tornam-se mais viciadas para um T fixo e Δ diminuindo. Para CIR ocorre o contrário: as estimativas de κ melhoram conforme decresce Δ para um T fixo. Isso talvez seja devido a alguma causa numérica.

A seguir ilustramos as trajetórias dos modelos CIR e Vasicek nas figuras (4.1) e (4.2), respectivamente, com $\kappa = 0,3$ e nas figuras (4.3) e (4.4) com $\kappa = 0,8$. As figuras (4.23) e (4.24) exibem as distribuições dos dois estimadores para $\kappa = 0,8$ no modelo CIR, demonstrando diferenças entre os comportamentos dos estimadores $\hat{\kappa}_{mvd}$ e $\hat{\kappa}_{mqo}$. A figura (4.25) mostra a distribuição de $\hat{\kappa}_{mvd}$ com $\kappa = 0,3$ para Vasicek.

Tabela 4.4: EMVD e EMQO com desvios-padrão para $\kappa = 0,3$, $\Delta = 0,1$ e T em meses.

T	VASICEK		CIR	
	EMVD	EMQO	EMVD	EMQO
60	0,366036 (0,116323)	0,366332 (0,116399)	0,398002 (0,135598)	0,370973 (0,119677)
72	0,356267 (0,105444)	0,356497 (0,105503)	0,377476 (0,117823)	0,357027 (0,106947)
84	0,343792 (0,094010)	0,343995 (0,094059)	0,363290 (0,105368)	0,347787 (0,097847)
96	0,340256 (0,085816)	0,340428 (0,085855)	0,352578 (0,094972)	0,340179 (0,089700)
108	0,333775 (0,082004)	0,333926 (0,082028)	0,345852 (0,088725)	0,335794 (0,085135)
120	0,329743 (0,077625)	0,329884 (0,077663)	0,340997 (0,083189)	0,332136 (0,080573)

(Obs.: 2000 repetições.)

Tabela 4.5: *EMVD e EMQO com desvios-padrão para $\kappa = 0,3$, $\Delta = 0,04761$ e T em meses.*

T	VASICEK		CIR	
	EMVD	EMQO	EMVD	EMQO
60	0,369981 (0,118487)	0,370128 (0,118526)	0,381903 (0,129696)	0,370399 (0,123172)
72	0,356706 (0,105435)	0,356827 (0,105480)	0,367690 (0,113984)	0,359024 (0,110515)
84	0,348118 (0,098355)	0,348212 (0,098375)	0,357180 (0,102077)	0,350668 (0,100334)
96	0,342318 (0,089584)	0,342404 (0,089602)	0,349707 (0,093464)	0,343650 (0,092452)
108	0,336361 (0,081258)	0,336435 (0,081270)	0,345172 (0,088583)	0,340576 (0,088088)
120	0,333961 (0,080448)	0,334025 (0,080459)	0,339347 (0,083238)	0,335884 (0,083614)

(Obs.: 2000 repetições.)

Tabela 4.6: *EMVD e EMQO com desvios-padrão para $\kappa = 0,3$, $\Delta = 0,01$ e T em meses.*

T	VASICEK		CIR	
	EMVD	EMQO	EMVD	EMQO
60	0,371814 (0,118885)	0,371843 (0,118894)	0,379534 (0,124259)	0,378199 (0,126047)
72	0,360248 (0,107936)	0,360272 (0,107941)	0,364338 (0,110640)	0,364570 (0,112207)
84	0,350950 (0,094294)	0,350796 (0,094900)	0,353044 (0,097699)	0,353850 (0,099234)
96	0,343006 (0,085945)	0,343024 (0,085950)	0,346581 (0,091746)	0,347476 (0,092932)
108	0,339315 (0,082617)	0,339330 (0,082621)	0,343492 (0,085212)	0,344352 (0,087348)
120	0,335744 (0,077790)	0,335758 (0,077793)	0,336668 (0,079314)	0,337324 (0,082406)

(Obs.: 2000 repetições.)

Tabela 4.7: *EMVD e EMQO com desvios-padrão para $\kappa = 0,8$, $\Delta = 0,1$ e T em meses.*

T	VASICEK		CIR	
	EMVD	EMQO	EMVD	EMQO
60	0,855489 (0,168381)	0,856222 (0,168508)	1,026641 (0,244108)	0,862661 (0,173098)
72	0,847516 (0,154259)	0,848069 (0,154372)	0,981429 (0,207259)	0,853855 (0,159034)
84	0,844634 (0,139937)	0,845134 (0,140022)	0,954880 (0,178372)	0,849769 (0,143811)
96	0,836219 (0,130169)	0,836676 (0,130246)	0,928821 (0,160331)	0,840487 (0,134361)
108	0,833367 (0,122371)	0,833731 (0,122409)	0,914067 (0,146697)	0,837619 (0,125659)
120	0,831576 (0,114698)	0,831916 (0,114757)	0,902387 (0,134916)	0,834749 (0,118010)

(Obs.: 2000 repetições.)

Tabela 4.8: *EMVD e EMQO com desvios-padrão para $\kappa = 0,8$, $\Delta = 0,04761$ e T em meses.*

T	VASICEK		CIR	
	EMVD	EMQO	EMVD	EMQO
60	0,867545 (0,177975)	0,867906 (0,178065)	0,945616 (0,210895)	0,872629 (0,180271)
72	0,852039 (0,161327)	0,852330 (0,161394)	0,913513 (0,184528)	0,856684 (0,163850)
84	0,844693 (0,146192)	0,844925 (0,146239)	0,895816 (0,164329)	0,848507 (0,149136)
96	0,840394 (0,135593)	0,840610 (0,135621)	0,883606 (0,149808)	0,842798 (0,138046)
108	0,834519 (0,125745)	0,834709 (0,125776)	0,872390 (0,137433)	0,837214 (0,128225)
120	0,830462 (0,118779)	0,830627 (0,118803)	0,863598 (0,128590)	0,832492 (0,121379)

(Obs.: 2000 repetições.)

Tabela 4.9: *EMVD e EMQO com desvios-padrão para $\kappa = 0,8$, $\Delta = 0,01$ e T em meses.*

T	VASICEK		CIR	
	EMVD	EMQO	EMVD	EMQO
60	0,863860 (0,179229)	0,863933 (0,179244)	0,880323 (0,186035)	0,867013 (0,181117)
72	0,849801 (0,159559)	0,849858 (0,159572)	0,863227 (0,165765)	0,853441 (0,162926)
84	0,843115 (0,147159)	0,843165 (0,147167)	0,853192 (0,151405)	0,844845 (0,149394)
96	0,840862 (0,137354)	0,840907 (0,137363)	0,849438 (0,140820)	0,842612 (0,139864)
108	0,835190 (0,128324)	0,835229 (0,128330)	0,843178 (0,131153)	0,837333 (0,130516)
120	0,831357 (0,122352)	0,831392 (0,122359)	0,838331 (0,124421)	0,832510 (0,124086)

(Obs.: 2000 repetições.)

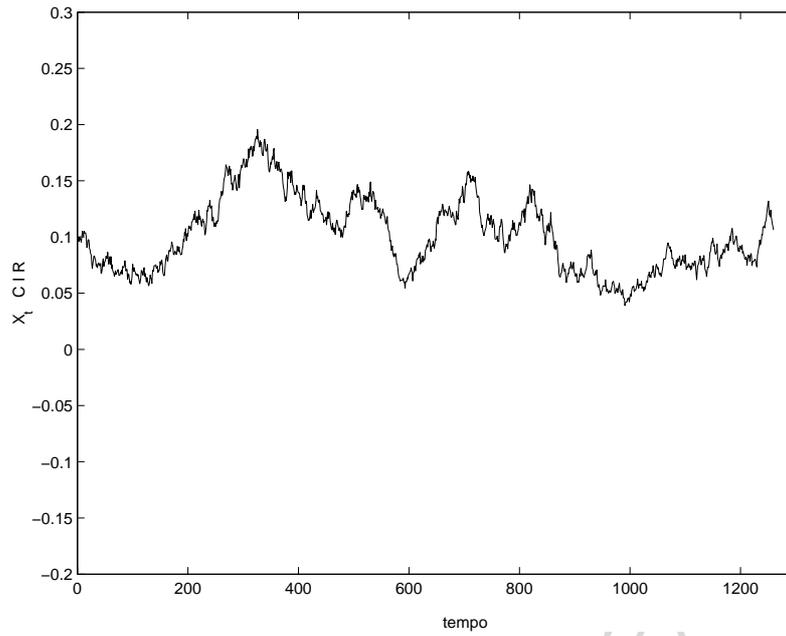


Figura 4.1: Trajetória simulada de um modelo de CIR $\kappa = 0,3$, $\nu = 0,1$, $\sigma = 0,06$, $T = 60$ e $\Delta = 0,04761$.

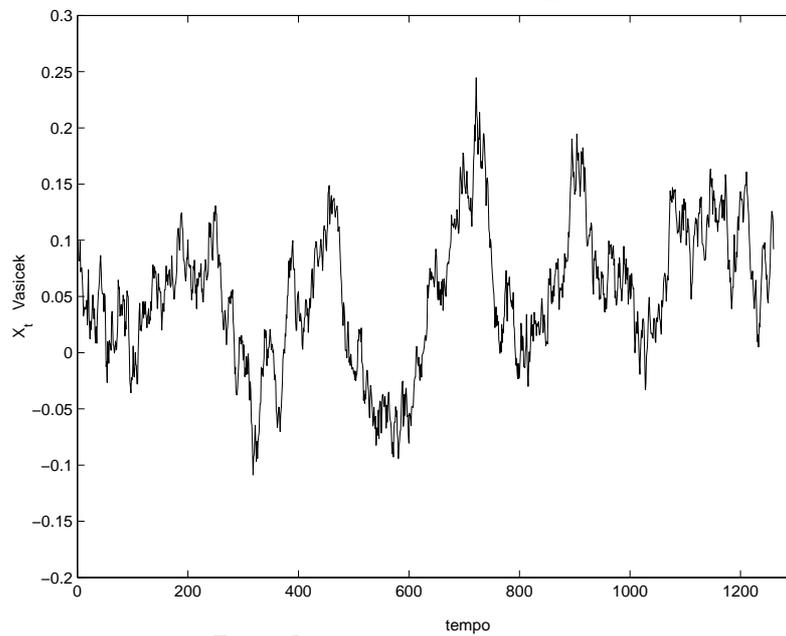


Figura 4.2: Trajetória simulada de um modelo de Vasicek $\kappa = 0,3$, $\nu = 0,1$, $\sigma = 0,06$, $T = 60$ e $\Delta = 0,04761$.

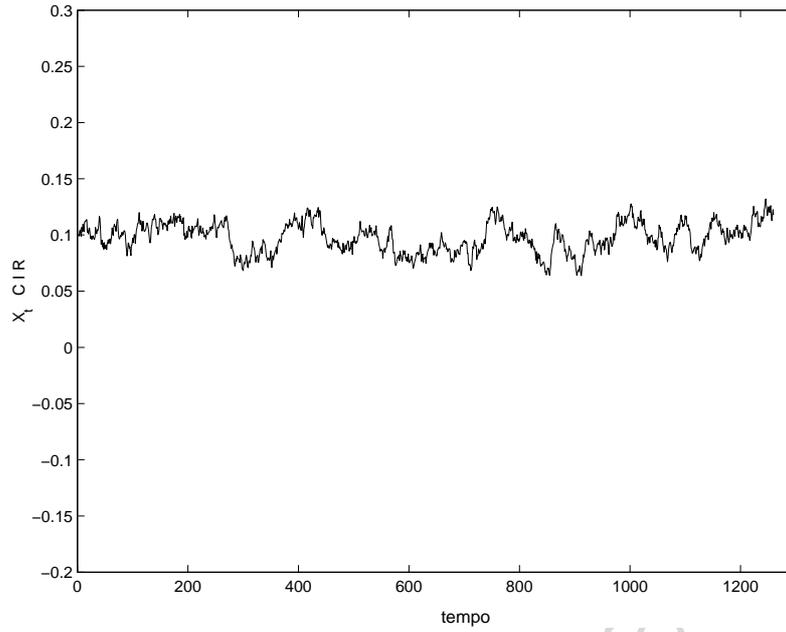


Figura 4.3: Trajetória simulada de um modelo de CIR $\kappa = 0,8$, $\nu = 0,1$, $\sigma = 0,06$, $T = 60$ e $\Delta = 0,04761$.

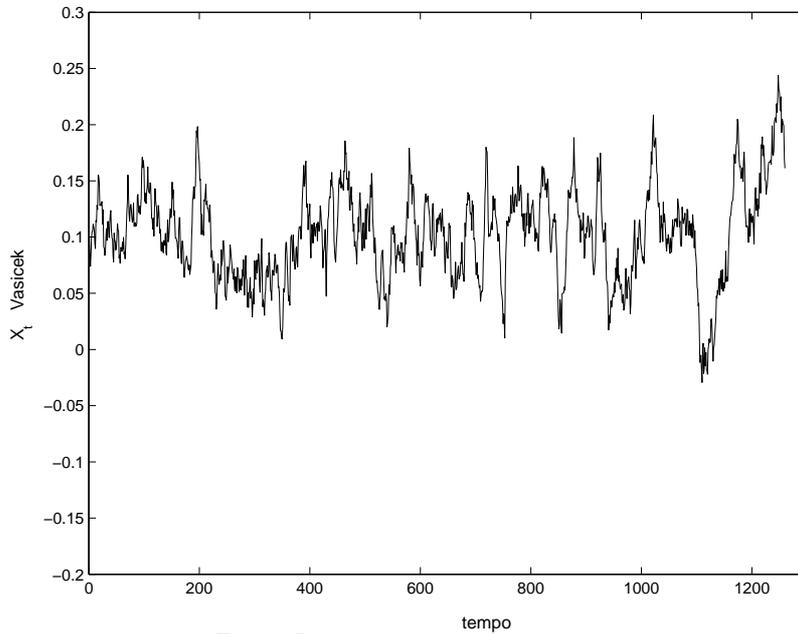


Figura 4.4: Trajetória simulada de um modelo de Vasicek $\kappa = 0,8$, $\nu = 0,1$, $\sigma = 0,06$, $T = 60$ e $\Delta = 0,04761$.

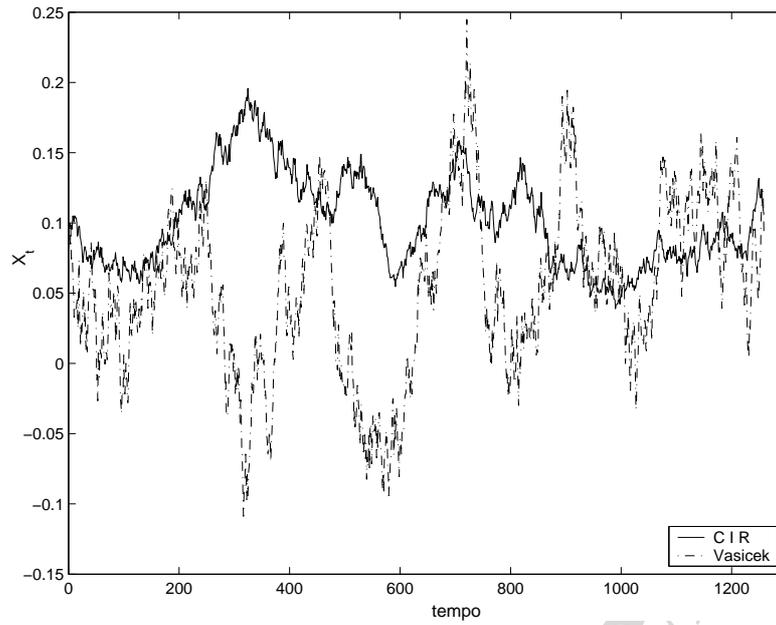


Figura 4.5: Trajetórias simuladas dos modelos CIR e Vasicek $\kappa = 0,3$, $\nu = 0,1$, $\sigma = 0,06$, $T = 60$ e $\Delta = 0,04761$.

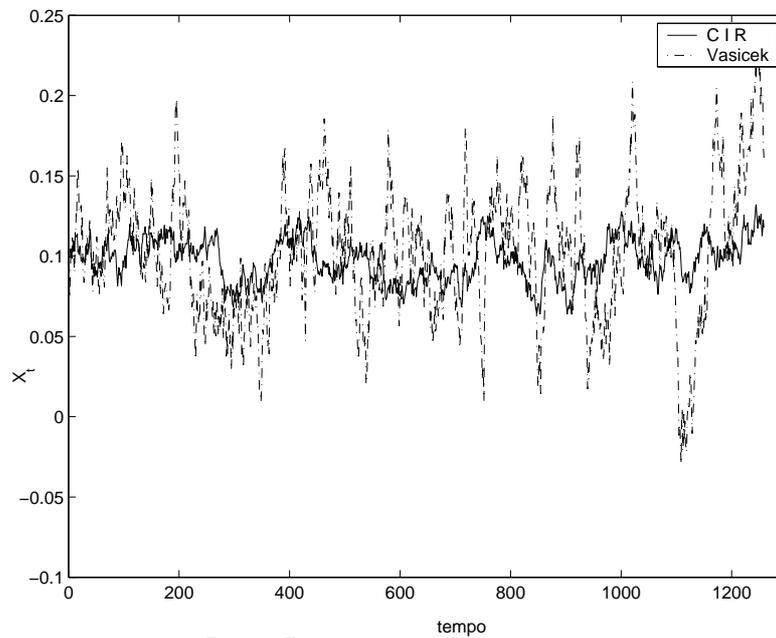


Figura 4.6: Trajetórias simuladas dos modelos CIR e Vasicek $\kappa = 0,8$, $\nu = 0,1$, $\sigma = 0,06$, $T = 60$ e $\Delta = 0,04761$.

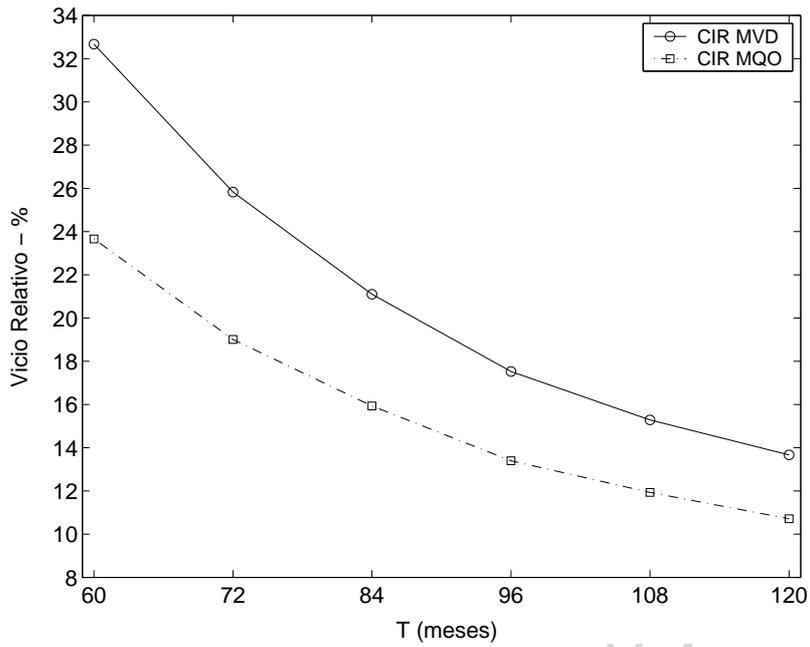


Figura 4.7: Variação do vício relativo em função de T para o modelo CIR, com $\Delta = 0, 1$ e $\kappa = 0, 3$.

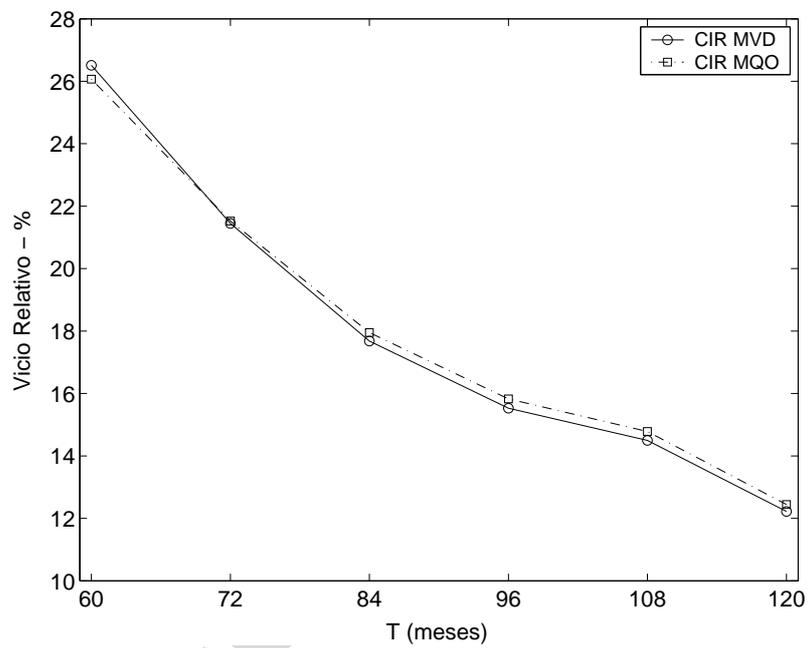


Figura 4.8: Variação do vício relativo em função de T para o modelo CIR, com $\Delta = 0, 01$ e $\kappa = 0, 3$.

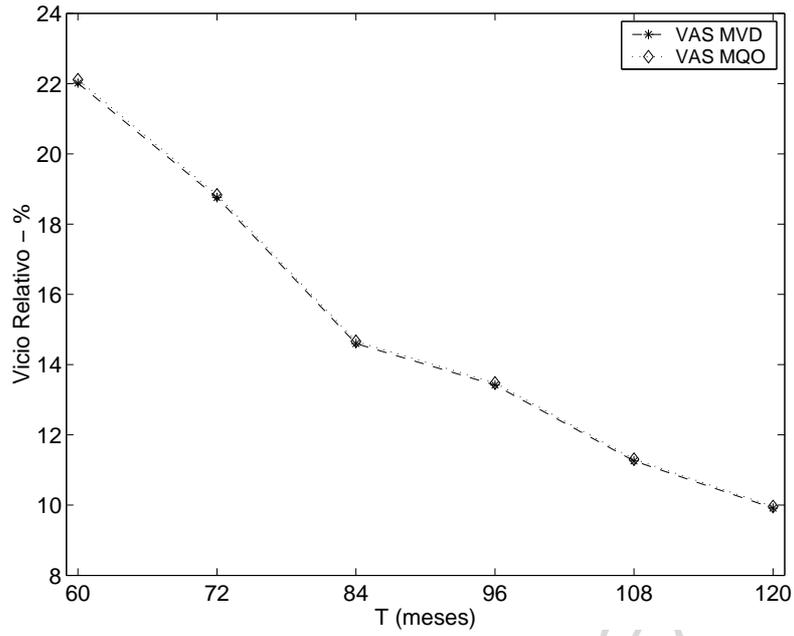


Figura 4.9: Variação do vício relativo em função de T para o modelo de Vasicek, com $\Delta = 0,1$ e $\kappa = 0,3$.

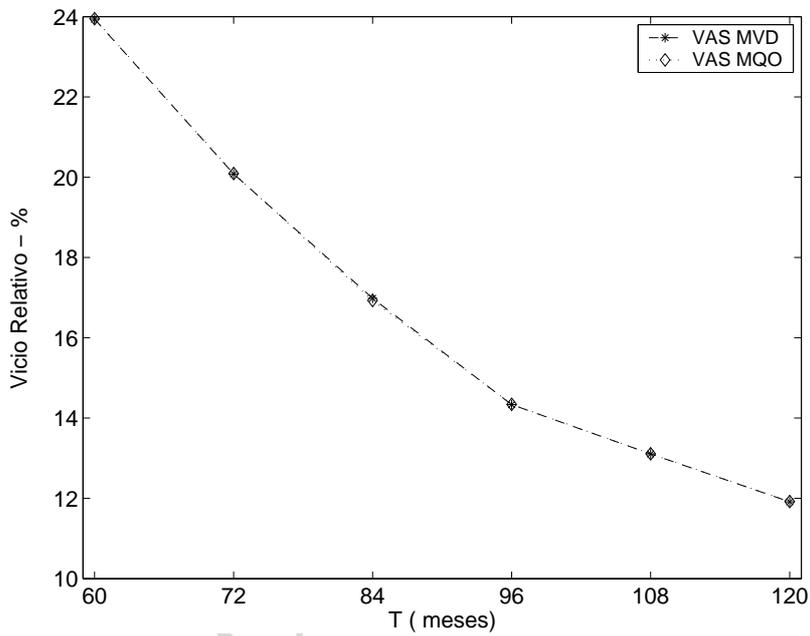


Figura 4.10: Variação do vício relativo em função de T para o modelo de Vasicek, com $\Delta = 0,01$ e $\kappa = 0,3$.

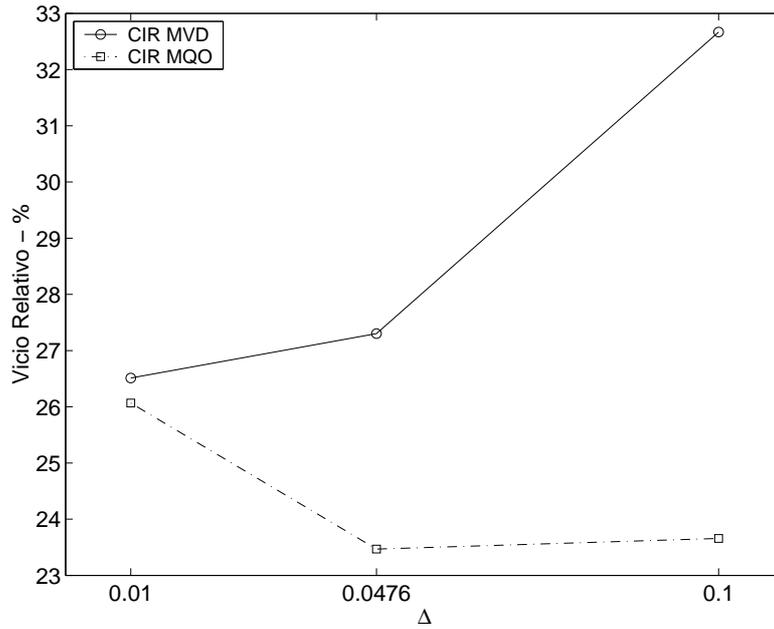


Figura 4.11: Variação do vício relativo em função do Δ para o modelo CIR, com $T = 60$ e $\kappa = 0,3$.

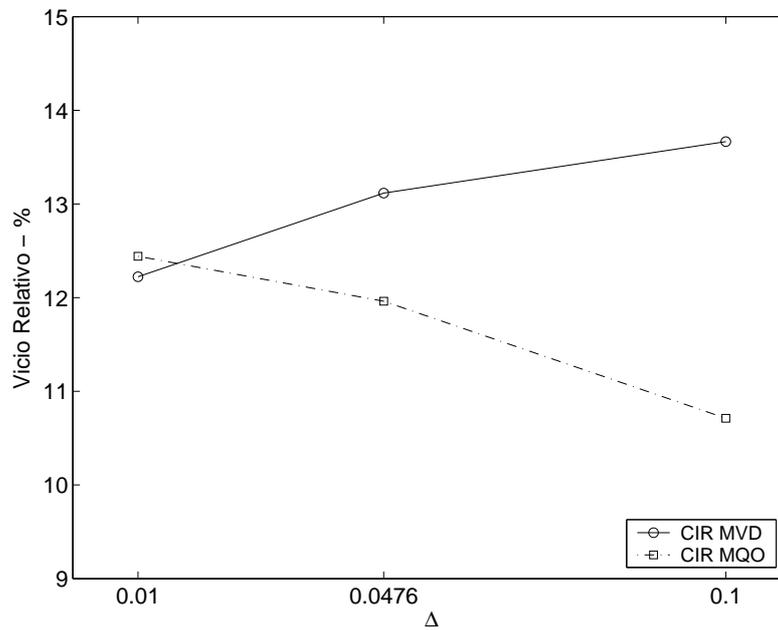


Figura 4.12: Variação do vício relativo em função do Δ para o modelo CIR, com $T = 120$ e $\kappa = 0,3$.

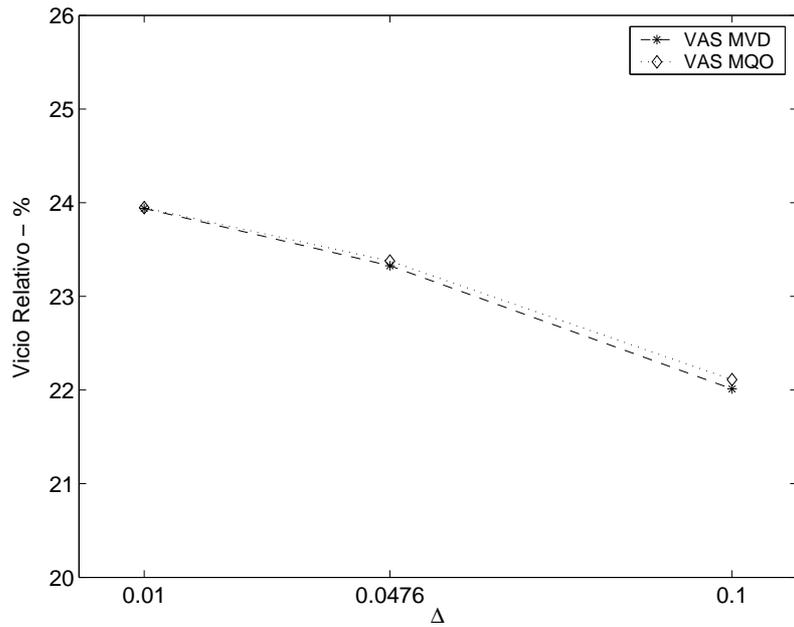


Figura 4.13: Variação do vício relativo em função do Δ para o modelo de Vasicek, com $T = 60$ e $\kappa = 0,3$.

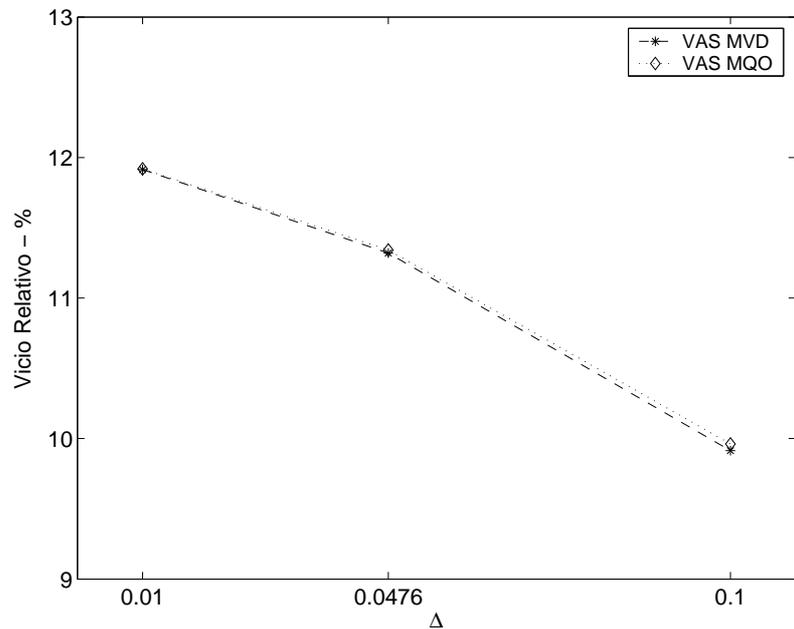


Figura 4.14: Variação do vício relativo em função do Δ para o modelo de Vasicek, com $T = 120$ e $\kappa = 0,3$.

Nas figuras (4.15) a (4.18) podemos ter uma idéia da variação do vício relativo do estimador do κ visualizando os gráficos que relacionam o mesmo ao tempo de observação T e ao intervalo de discretização Δ simultaneamente. As figuras (4.15) e (4.16) se referem ao EMVD e as figuras (4.17) e (4.18) ao EMQO.

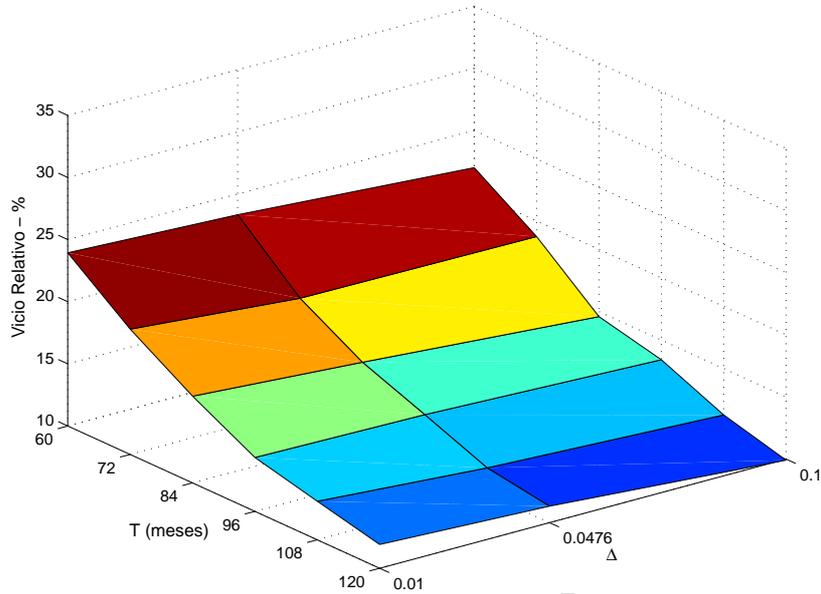


Figura 4.15: Variação do vício relativo de $\hat{\kappa}_{mvd}$ para o modelo de Vasicek e $\kappa = 0,3$.

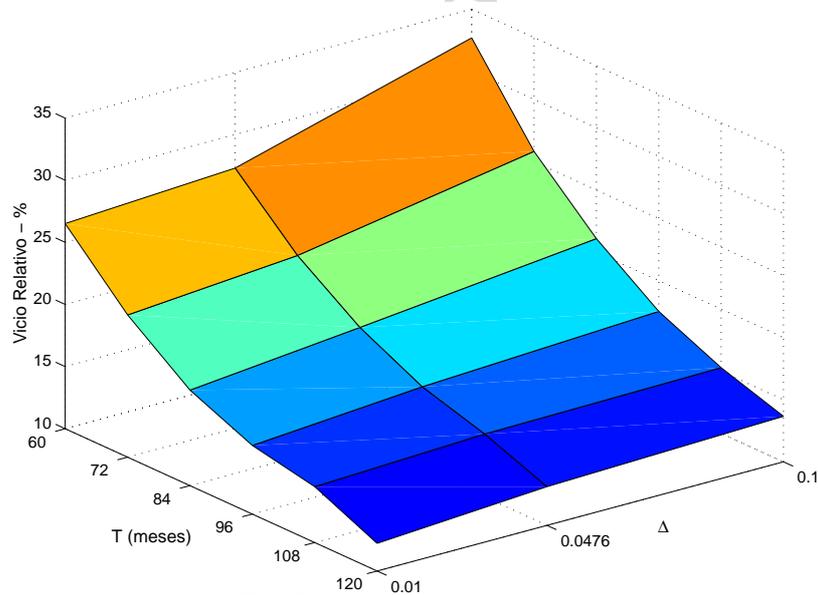


Figura 4.16: Variação do vício relativo de $\hat{\kappa}_{mvd}$ para o modelo CIR e $\kappa = 0,3$.

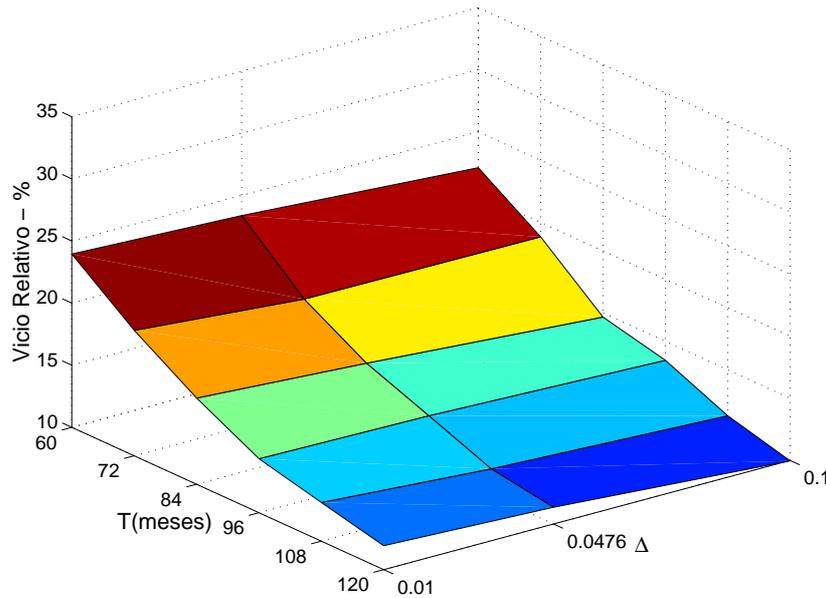


Figura 4.17: Variação do vício relativo de $\hat{\kappa}_{mqo}$ para o modelo de Vasicek e $\kappa = 0, 3$.

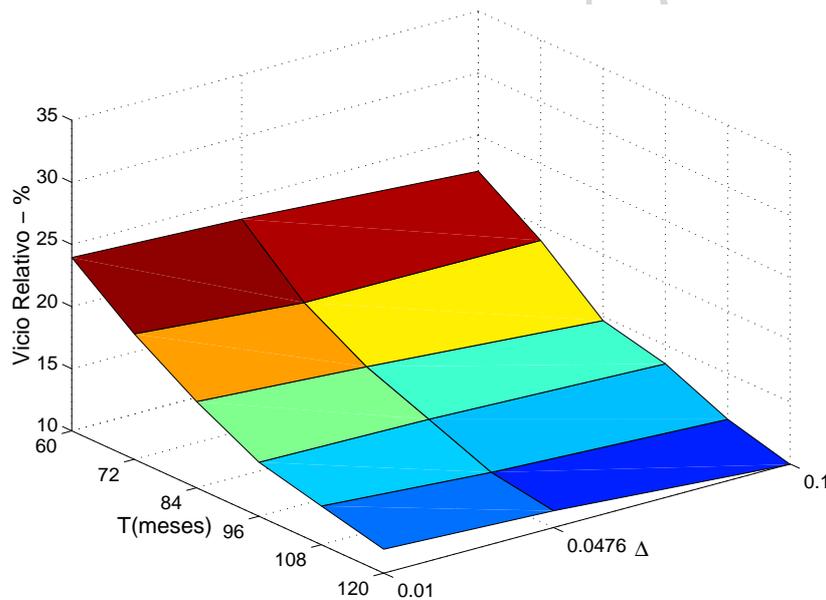


Figura 4.18: Variação do vício relativo de $\hat{\kappa}_{mqo}$ para o modelo de CIR e $\kappa = 0, 3$.

Nas figuras (4.19) a (4.22) temos a mesma análise para o caso dos vícios relativos, porém para os desvios-padrão dos estimadores. Nos gráficos das figuras (4.19) e (4.20) temos a variação dos desvios-padrão do EMVD para os modelos de Vasicek e CIR, enquanto nas figuras (4.21) e (4.22) a variação dos desvios-padrão do EMQO. Os gráficos são das estimativas de κ .

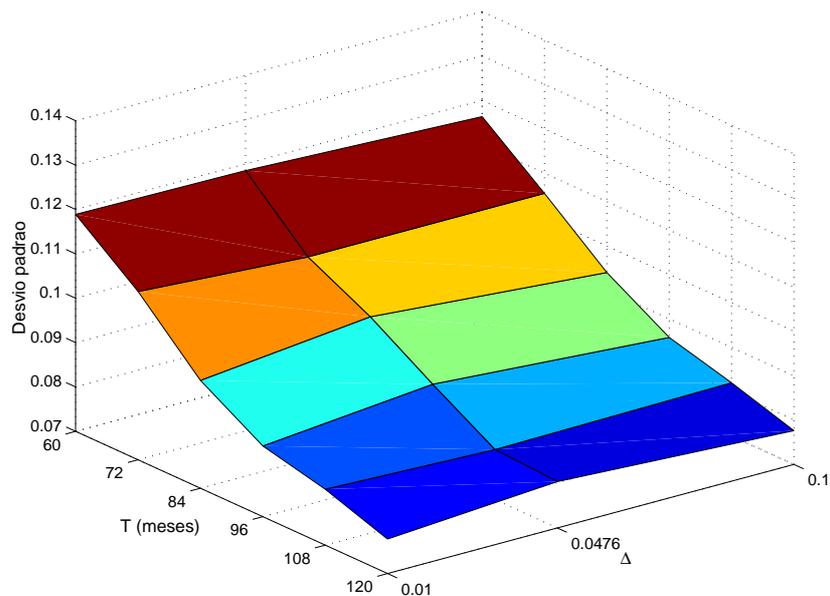


Figura 4.19: Variação do desvio padrão de $\hat{\kappa}_{mvd}$ para o modelo de Vasicek e $\kappa = 0, 3$.

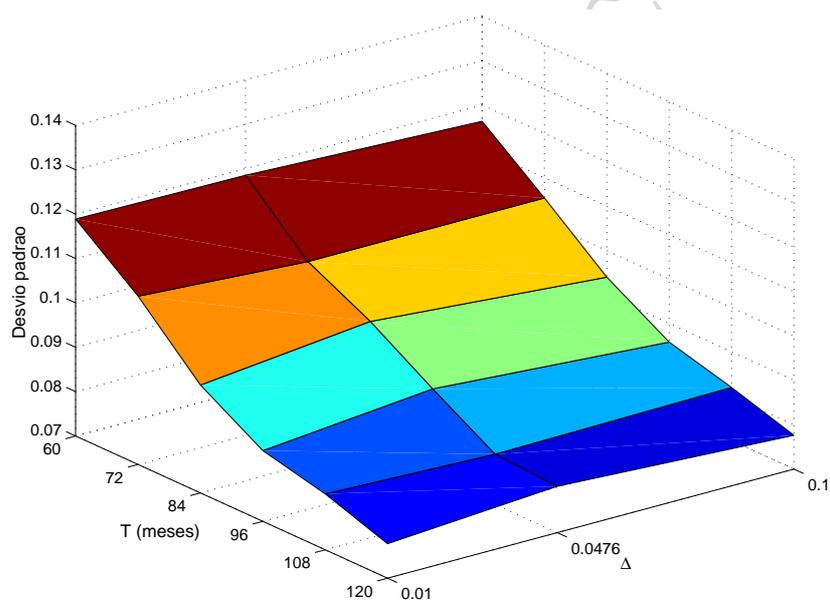


Figura 4.20: Variação do desvio padrão de $\hat{\kappa}_{mvd}$ para o modelo de CIR e $\kappa = 0, 3$.

Ao se comparar qualitativamente os comportamentos das estimativas para mesmos modelo, método de estimação e valor verdadeiro de κ , pode-se perceber que a variação do vício relativo e do desvio padrão das estimativas não apresentam o mesmo comportamento. Isso pode ser observado confrontando:

- o gráfico da figura (4.15) com o gráfico da figura (4.19): Modelo de Vasicek, EMVD

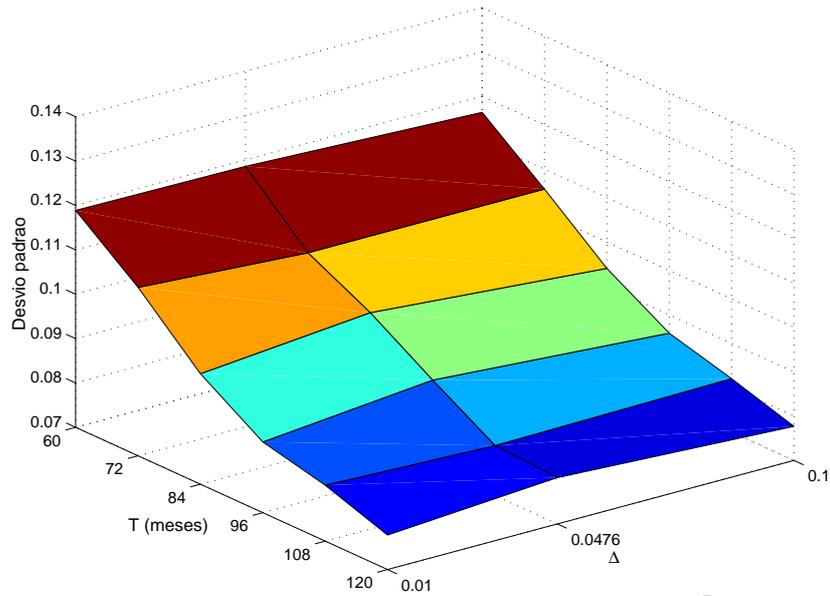


Figura 4.21: Variação do desvio padrão de $\hat{\kappa}_{mqo}$ para o modelo de Vasicek e $\kappa = 0, 3$.

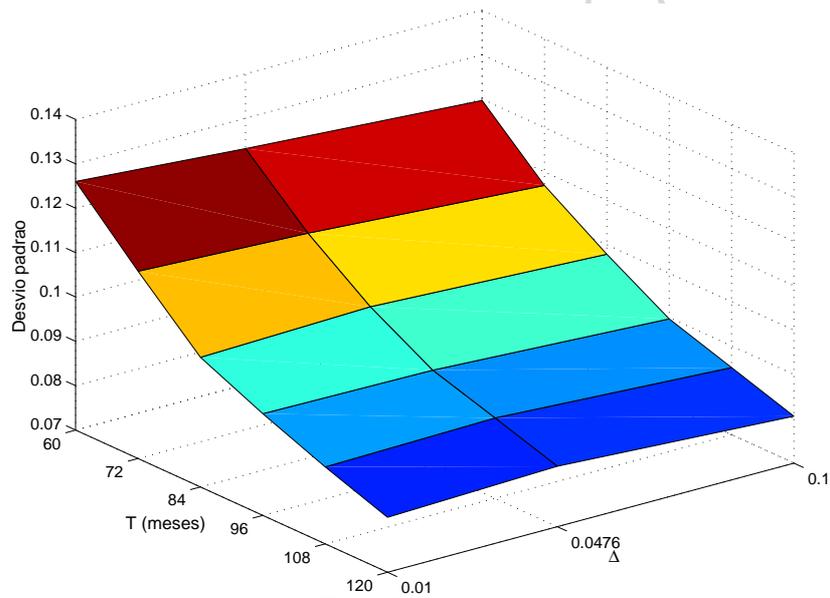


Figura 4.22: Variação do desvio padrão de $\hat{\kappa}_{mqo}$ para o modelo de CIR e $\kappa = 0, 3$.

e $\kappa = 0, 3$.

- o gráfico da figura (4.16) com o gráfico da figura (4.20): Modelo de CIR, EMVD e $\kappa = 0, 3$.
- o gráfico da figura (4.17) com o gráfico da figura (4.21): Modelo de Vasicek, EMQO

e $\kappa = 0,3$.

- o gráfico da figura (4.18) com o gráfico da figura (4.22): Modelo de CIR, EMQO e $\kappa = 0,3$.

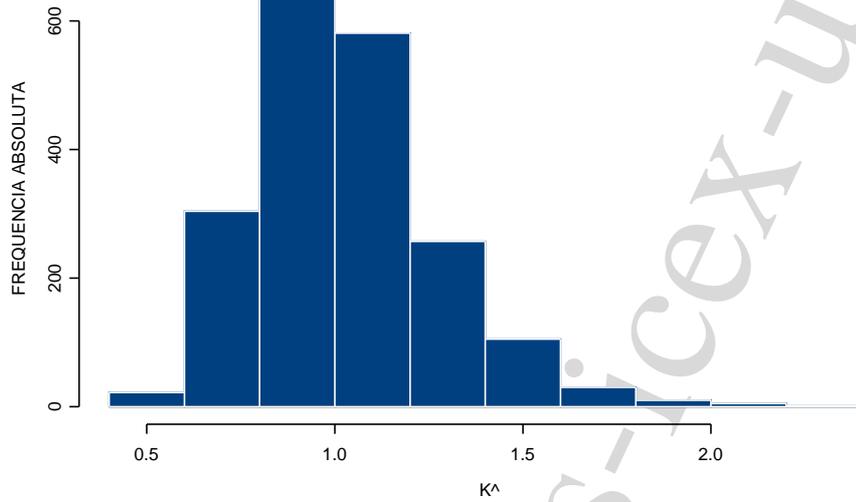


Figura 4.23: Distribuição de $\hat{\kappa}_{mvd}$ para o modelo de CIR com $\kappa = 0,8$, $\nu = 0,1$, $\sigma = 0,06$, $T = 60$ meses e $\Delta = 0,04761$.

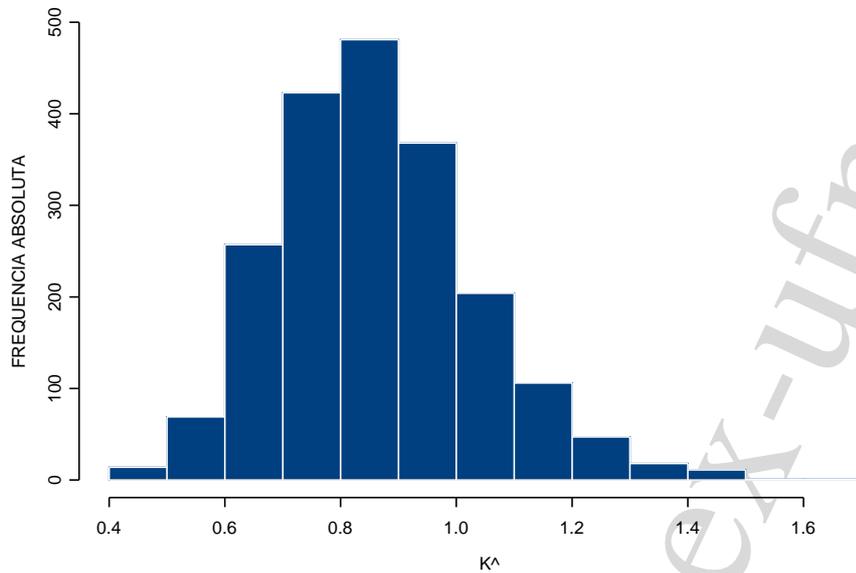


Figura 4.24: Distribuição de $\hat{\kappa}_{mqo}$ para o modelo de CIR com $\kappa = 0,8$, $\nu = 0,1$, $\sigma = 0,06$, $T = 60$ meses e $\Delta = 0,04761$

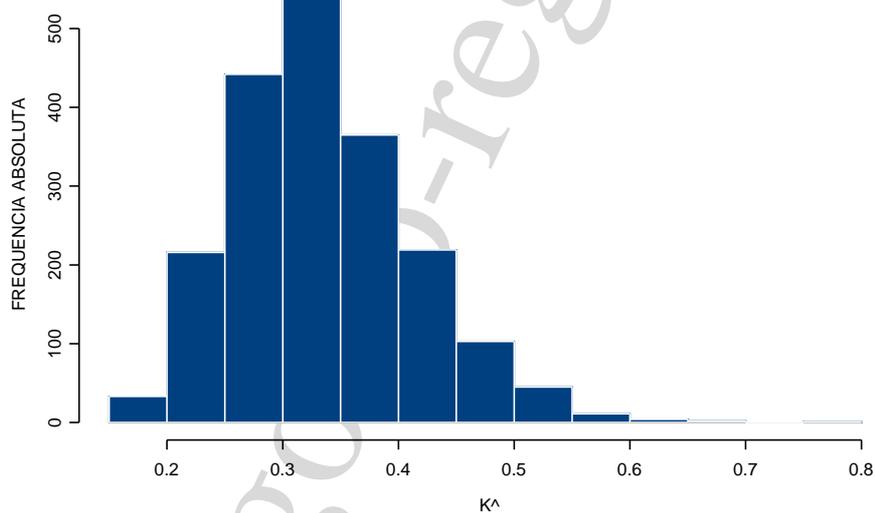


Figura 4.25: Distribuição de $\hat{\kappa}_{mvd}$ para o modelo de Vasicek com $\kappa = 0,3$, $\nu = 0,1$, $\sigma = 0,06$, $T = 60$ meses e $\Delta = 0,04761$

Capítulo 5

Aplicações

5.1 Títulos dos EUA

Os dados deste conjunto constituem 1809 observações semanais, desde 5 de janeiro de 1962 até 30 de agosto de 1996, dos rendimentos das Notas do Tesouro dos EUA com maturidade de 3 meses. As taxas foram observadas nas sexta-feiras, com exceção de feriados, caso em que as taxas das quintas-feiras foram utilizadas. Estas são taxas anualizadas e foi considerado o ano de 360 dias. Este conjunto de dados é o mesmo utilizado no trabalho de Durham e Gallant [14].

A figura (5.1) exhibe apenas a série, enquanto a figura (5.3) contrasta a mesma série das taxas observadas com a tendência, ou seja, os níveis médios, obtidos dos ajustes para ambos os modelos, CIR e Vasicek, usando os estimadores de máxima verossimilhança discretizados. Na figura (5.2) o que se observa é o gráfico de dispersão da variação diária das taxas pela taxa do dia anterior. Pode-se notar que a variabilidade das diferenças diárias não é constante, aumentando conforme o nível das taxas.

5.2 Depósito Interbancário (DI) Prefixado do Brasil

Os dados deste conjunto são dez séries diárias brasileiras de taxas de juros prefixadas de depósito interbancário (DI), anuais, cada uma contendo 497 observações no período de

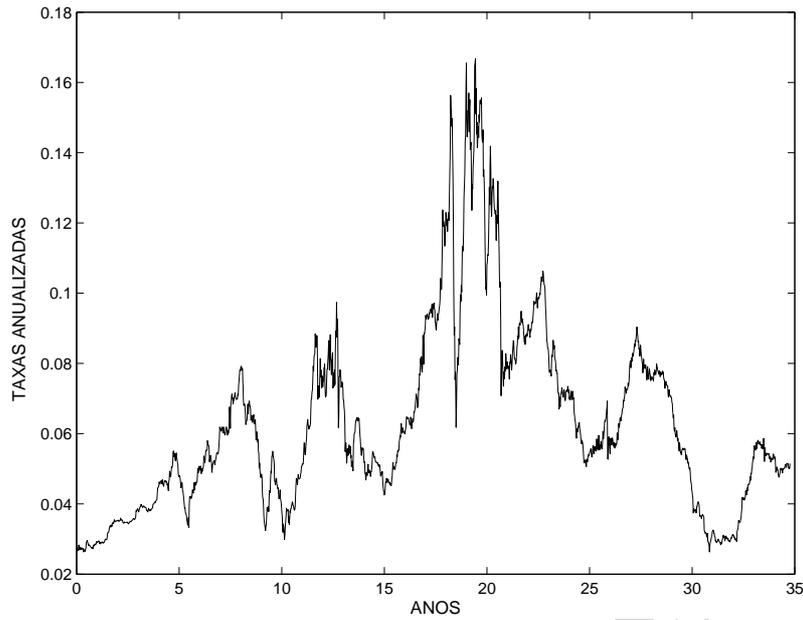


Figura 5.1: Taxas semanais dos Títulos de 3 meses dos EUA, de 05/01/1962 até 30/08/1996.

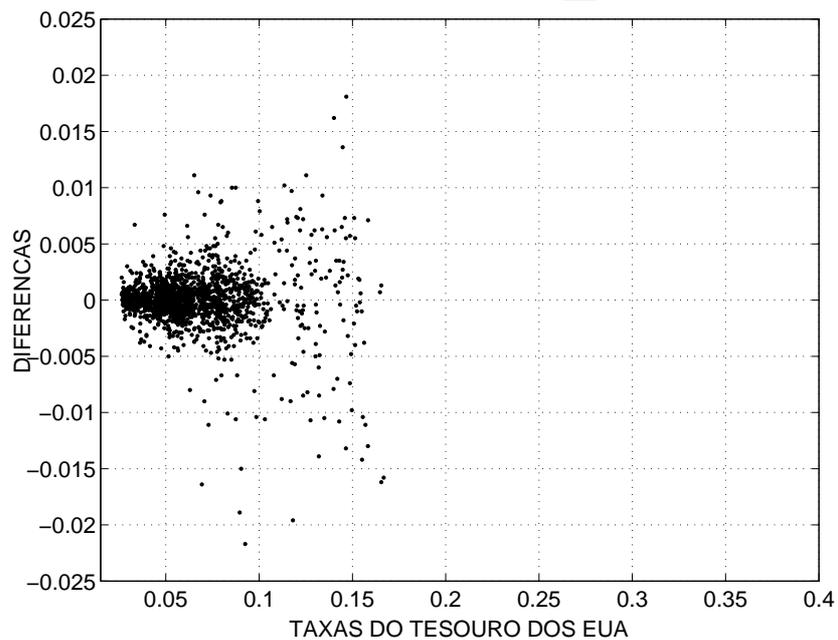


Figura 5.2: Dispersão das diferenças $Y_i = X_{i+1} - X_i$ das taxas contra as taxas X_i do tesouro dos EUA.

25/03/2003 a 22/03/2005. Vemos, na figura (5.4), três curvas de rendimento envolvendo as maturidades de 1, 3, 6, 12, 18, 24, 36, 48, 60 e 84 meses.

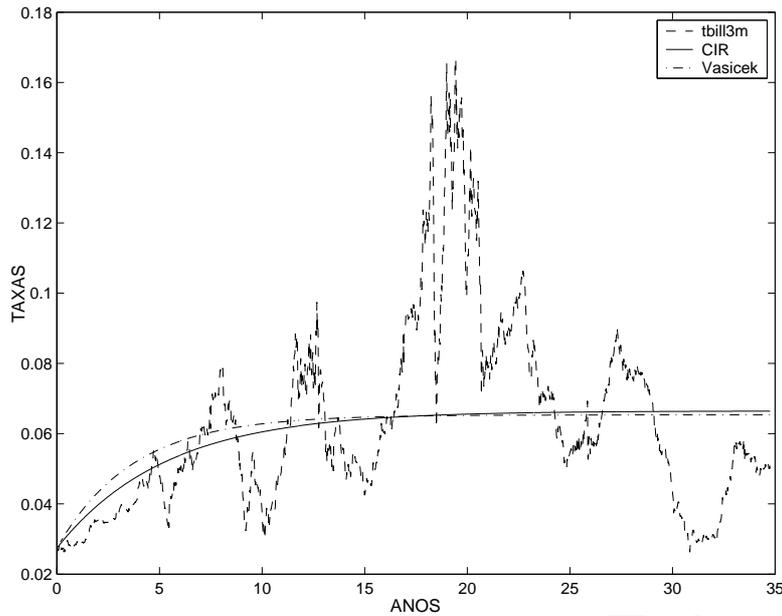


Figura 5.3: Taxas dos títulos dos EUA e tendências ajustadas pelos modelos de CIR e Vasicek.

Os comportamentos das trajetórias das taxas dos títulos de 3 meses, 60 meses e 84 meses são parecidos, acentuadamente para os dois últimos, conforme mostrado nas figuras (5.6), (5.8) e (5.10), respectivamente. Nos gráficos de dispersão das figuras (5.7) e (5.9) observa-se maior heterogeneidade nas taxas de 60 meses e de 84 meses, que também mantém comportamento parecido nesse aspecto. As taxas de 3 meses apresentam comportamento mais homogêneo que as duas primeiras, conforme ilustra (5.5).

5.3 Títulos do Canadá

Este terceiro conjunto de dados compreende, originalmente, o período de 1/1/1996 até 30/03/2006. No entanto, dias para os quais a taxa não estava disponível foram retirados da série, restando 2572 observações. O número total de dias é 2673. As taxas desses títulos com maturidade de 3 meses foram utilizadas no trabalho de Jiang e Knight [25], porém em um período diferente.

A figura (5.12) mostra a série e podemos notar que a mesma apresenta um comportamento diferente das séries americanas e brasileiras. Vemos na figura (5.11) a dispersão

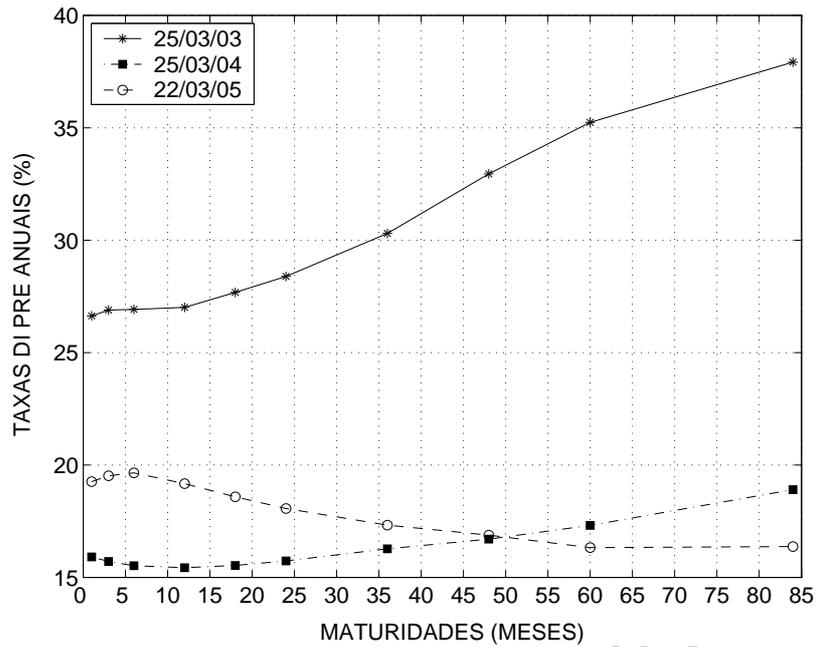


Figura 5.4: Curvas de rendimento das taxas DI diárias em três datas distintas.

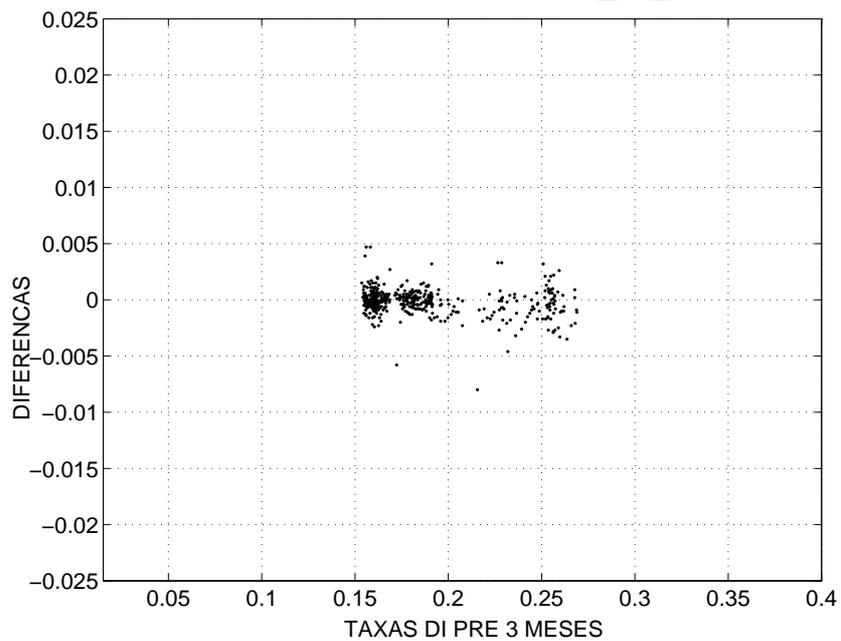


Figura 5.5: Dispersão das diferenças $Y_i = X_i - X_{i-1}$ das taxas DI de 3 meses pelas taxas X_{i-1} .

dos incrementos diários das taxas contra as taxas do dia anterior.

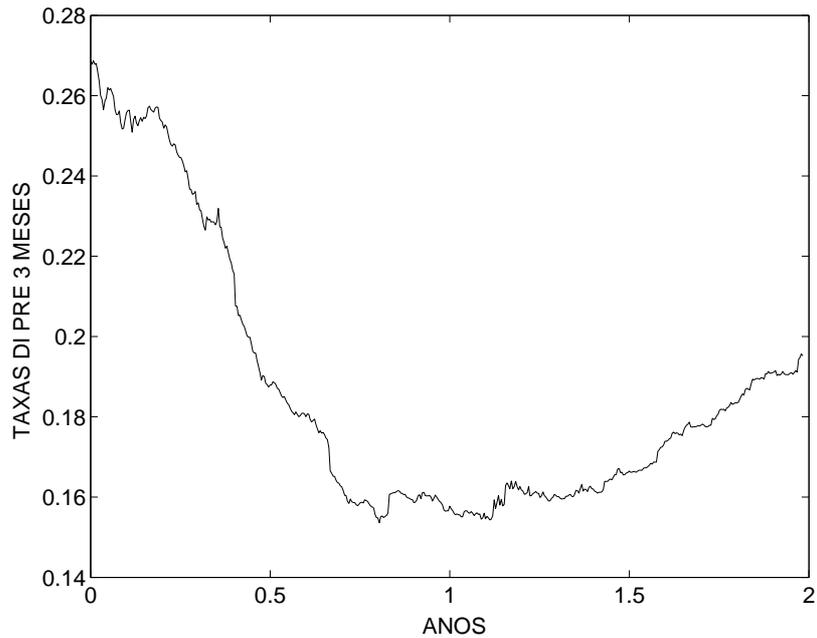


Figura 5.6: Taxas DI diárias de 3 meses brasileiras, de 23/03/2003 até 22/03/2005.

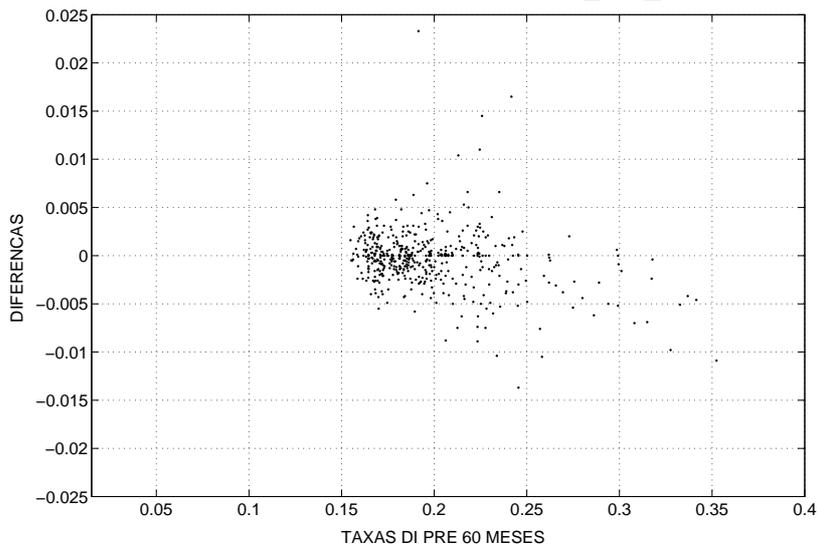


Figura 5.7: Dispersão das diferenças $Y_i = X_i - X_{i-1}$ das taxas DI de 60 meses pelas taxas X_{i-1} .

5.4 Resumo das Estimativas

Podemos notar, a partir dos gráficos de dispersão das diferenças das taxas contra as taxas, que a série dada pelas taxas diárias do Canadá é a que apresenta a volatilidade, em geral, não só menor como também mais constante. Por outro lado, esta série é a

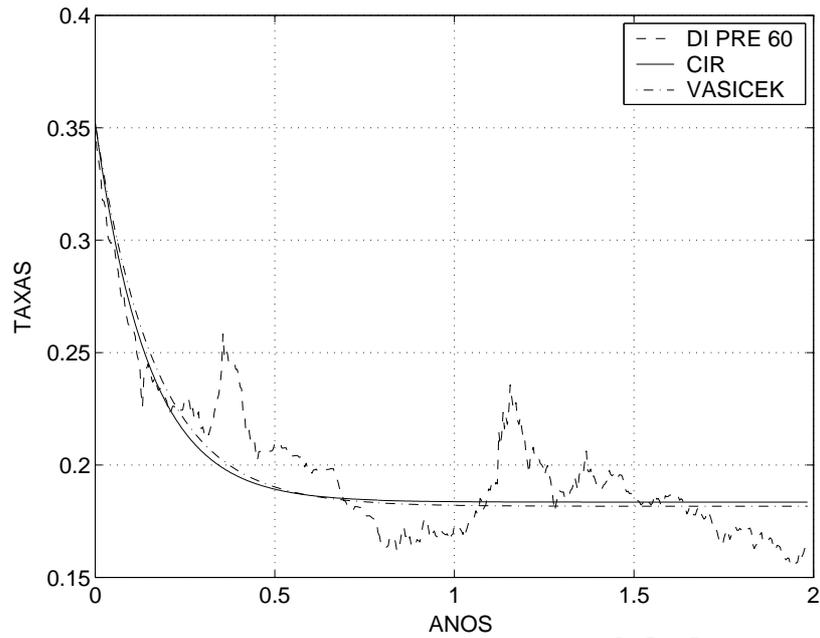


Figura 5.8: Taxas DI diárias de 60 meses e tendências ajustadas pelos modelos CIR e Vasicek.

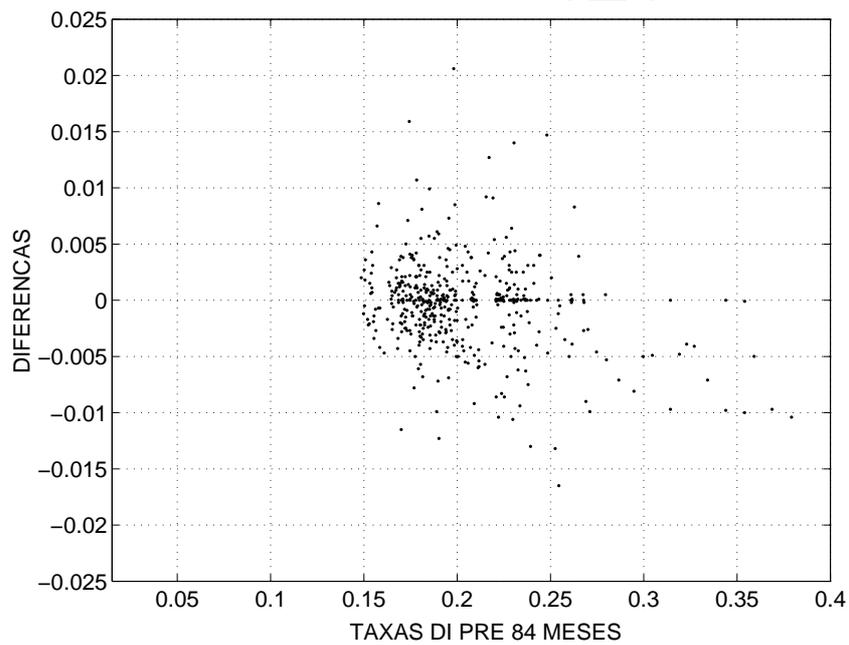


Figura 5.9: Dispersão das diferenças $Y_i = X_i - X_{i-1}$ das taxas DI de 84 meses pelas taxas X_{i-1} .

que parece apresentar o pior ajuste entre todas as séries reais utilizadas, o que pode ser explicado, considerando o gráfico desta mesma série, dado pela figura (5.12), pelo fato

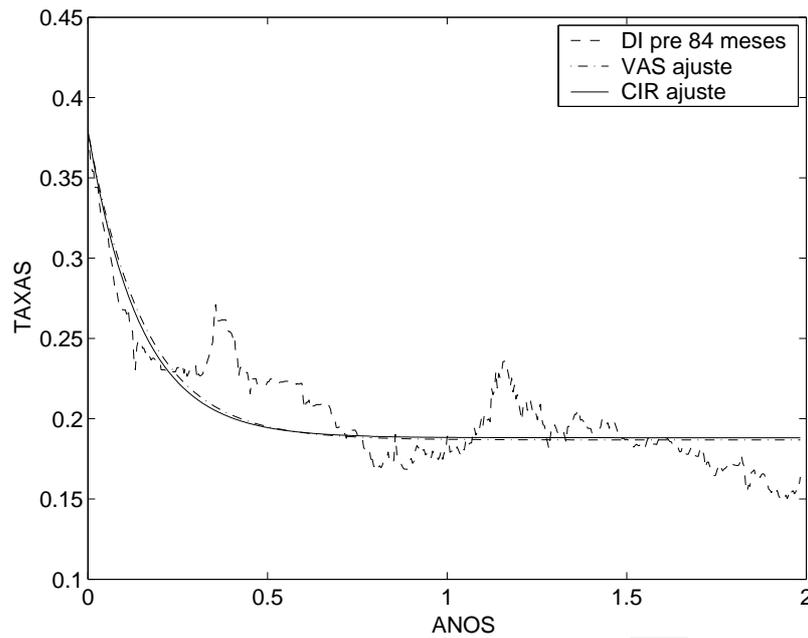


Figura 5.10: Taxas DI diárias de 84 meses e tendências ajustadas pelos modelos de CIR e Vasicek.

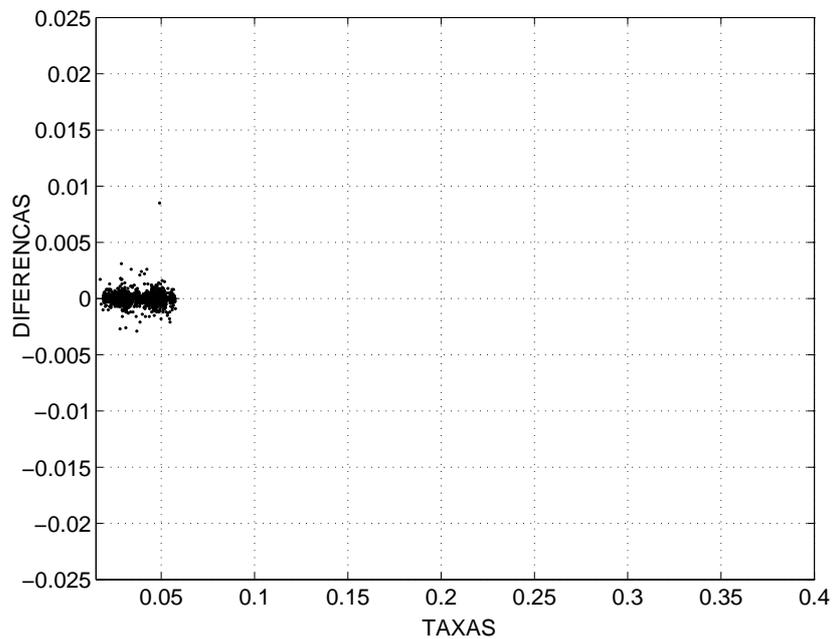


Figura 5.11: Dispersão das diferenças $Y_i = X_{i+1} - X_i$ das taxas contra as taxas X_i canadenses.

de que a tendência desta série (se existir) não é da mesma forma característica dos dois modelos utilizados.

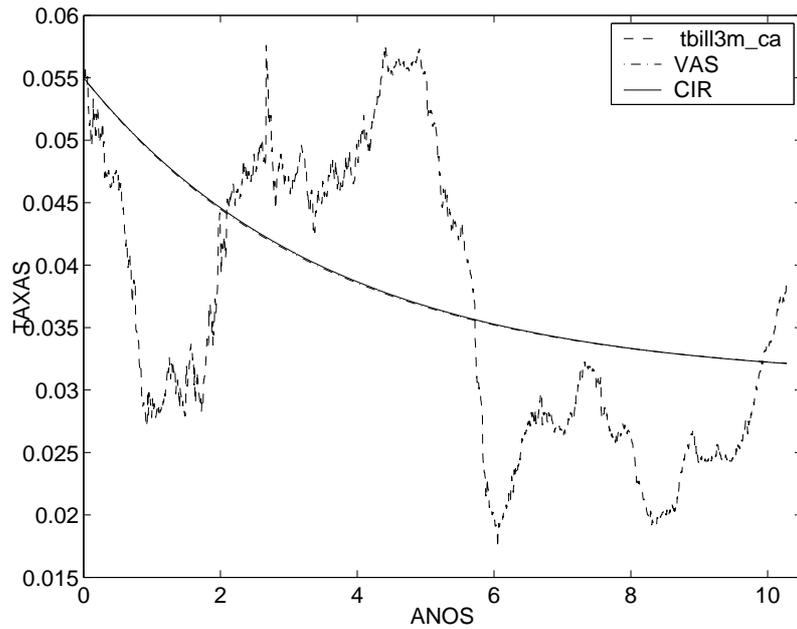


Figura 5.12: Taxas diárias dos títulos canadenses com maturidade de 3 meses, de 02/01/1996 até 30/03/2006.

Tabela 5.1: *Estimativas por EMVD para κ e ν .*

SÉRIE	N	Δ	$\hat{\kappa}$		$\hat{\nu}$	
			Vasicek	CIR	Vasicek	CIR
Tbill 03 meses	1809	0,019	0,2737122	0,1904838	0,0653499	0,0664692
DI Pre 03 meses	497	0,004	1,6138092	2,0399213	0,1627046	0,1675031
DI Pre 60 meses	497	0,004	5,9120424	6,6866580	0,1816088	0,1834737
DI Pre 84 meses	497	0,004	6,2169618	6,7423206	0,1867749	0,1881336
Tbill 03 meses CA	2572	0,004	0,2857602	0,2801387	0,0308692	0,0307537

Tabela 5.2: *Estimativas por EMQO para κ e ν .*

SÉRIE	N	Δ	$\hat{\kappa}$	$\hat{\nu}$
Tbill 03 meses	1809	0,019	0,2744174	0,0653651
DI Pre 03 meses	497	0,004	1,6898517	0,1642052
DI Pre 60 meses	497	0,004	6,1519654	0,1827553
DI Pre 84 meses	497	0,004	6,4011308	0,1878228
Tbill 03 meses CA	2572	0,004	0,2865212	0,0309022

Podemos notar, da tabela (5.1), que para as séries brasileiras as diferenças percentuais entre as estimativas de κ considerando o modelo de Vasicek e o modelo CIR diminuem conforme a maturidade dos títulos aumenta. Para os títulos de 3 meses a diferença é de aproximadamente 26%, caindo para aproximadamente 13% e depois para 8,45%, aproximadamente. Tal fato talvez seja reflexo do que a diz a teoria econômica a respeito dos títulos: a volatilidade dos títulos diminui conforme aumenta a maturidade. Mesmo em se tratando aqui de uma modelagem apenas da tendência, pode ser que tal fenômeno esteja, em parte, influenciando implicitamente a estimativa do parâmetro κ . Quanto às estimativas de ν , as diferenças percentuais são de 3%, 1% e 0,7%, aproximadamente, para os títulos com maturidade de 3 meses, 60 meses e 84 meses, respectivamente.

Quanto às estimativas por EMQO dos mesmos conjuntos de dados, tabela (5.2), encontramos $\hat{\kappa}_{mqo}$ sempre maior que o respectivo $\hat{\kappa}_{mvd}$ para o Modelo de Vasicek. Vemos também que $\hat{\nu}_{mqo}$ é sempre maior que o respectivo $\hat{\nu}_{mvd}$. Apesar de tal comportamento parecer sistemático, não se pode afirmar que o seja de fato e nem que essas diferenças são significativas.

Capítulo 6

Considerações Finais

6.1 Conclusões

O EMVD se comporta, aproximadamente, como o estimador de máxima verossimilhança de tempo contínuo usual para tais modelos de difusão. Entretanto, a simples discretização dos estimadores naturais para o caso contínuo leva a um vício de discretização, também já reportado na literatura (Cleur e Manfredi [11]), que compromete o desempenho do estimador, principalmente quando se trata de pequenas amostras.

O aspecto numérico envolvido nos estudos que fazem uso de simulação pode, em muitos casos, influenciar os resultados, desde simples problemas com a precisão dos cálculos até procedimentos adotados. Nesse sentido os métodos de geração de variáveis aleatórias (quando do uso de simulação estocástica), integração numérica e esquemas de discretização são de grande importância nesses estudos.

6.2 Propostas de Trabalhos Futuros

Como idéias para trabalhos futuros, a utilização de outros métodos de integração numérica, em particular o de Gauss-Legendre e de esquemas de discretização de ordem mais alta, como o de Taylor, serão considerados. Além disso, a questão do intervalo entre observações - passo de discretização quando em se tratando da etapa de simulação,

e espaço de tempo entre observações amostradas, quando em se tratando de aplicação a dados reais - também será avaliada com vistas à verificação de sua influência na estimação.

Neste trabalho o objetivo foi a estimativa da coeficiente de tendência de uma EDE, mas não o coeficiente de difusão, o qual será o próximo passo. Assim, será possível passar ao problema de previsão e de escolha do melhor modelo para este fim. Constatamos, tendo feito uma ampla revisão da literatura disponível, que o trabalho de Durham e Gallant [14], além de ser um dos mais recentes e de trazer abordagem nova, também tem causado certa influência nas pesquisas dessa área, tornando-se referência importante. Esta deverá ser a referência a adotarmos como base de comparação ao darmos continuidade às pesquisas.

A essência do que foi feito neste trabalho deve ser repetida num contexto completamente não-paramétrico, utilizando a metodologia dos núcleo estimadores, pois alguns trabalhos na literatura, como Ait-Sahalia [1] e [2], apontam a necessidade de não se fixar uma forma paramétrica para o coeficiente de tendência, uma vez que essa atitude influi na estimação do coeficiente de difusão, mesmo que, para este, a abordagem seja não-paramétrica.

Abstract

Due to the sophistication of the financial products in the last years, continuous time probabilistic models defined by stochastic differential equations have had the increasing attention by the financial market analysts and academics concerning this area. To surpass the limitation of some models, it is necessary to consider the way the interest rates get developed throughout the time. In this dissertation the drift coefficient estimation of the Vasicek model and the Cox-Ingersoll-Ross model is carried through the discretized maximum likelihood estimator and the ordinary least squares estimator. The necessity to know the value of the diffusion coefficient is overcome and the presence of the bias estimators is confirmed, for considering the discrete time observations for continuous time models. It is also presented the comparison of the estimators performance when changes are made in the discretisation interval and the time observation of the process. The drift coefficients of both models are estimated by the two methods, achieving the expected trajectory of the interest rates of five real data sets.

Keywords: Stochastic differential equations, Term structure of the short-term interest rates, Vasicek model, Cox-Ingersoll-Ross model, Euler discretisation scheme.

Referências Bibliográficas

- [1] AIT-SAHALIA, Y. (1996) Testing Continuous-Time Models of the Spot Interest Rate, *The Review of Financial Studies*, 9,(2), 385-426.
- [2] AIT-SAHALIA, Y. (1999) Transition Densities for Interest Rate and Other Nonlinear Diffusions, *The Journal of Finance*, 54,(4), 1361-1395.
- [3] AIT-SAHALIA, Y. (2002) Maximum Likelihood Estimation of Discretely Sampled Diffusions: A Closed-Form Approximation Approach, *Econometrica*, 70,(1), 223-262.
- [4] ASSAF NETO, A. (1999) *Mercado Financeiro*, 2a. edição, São Paulo, Editora Atlas.
- [5] BIANCHI, C. and CLEUR, E.M. (1996) Indirect Estimation of Stochastic Differential Equation Models: Some Computational Experiments, *Computational Economics*, 9, 257-274.
- [6] BILLINGSLEY, P. (1995) *Probability and Measure*, Third Edition, Wiley Series in Probability and Mathematical Statistics, New York.
- [7] BROZE, L., SCAILLET and O. ZAKOIAN, J.M. (1995a) Testing for Continuous-time Models of the Short-term Interest Rate, *Journal of Empirical Finance*, 2, 199-223.
- [8] BROZE, L., SCAILLET and O. ZAKOIAN, J.M. (1995b) Quasi Indirect Inference for Diffusion Processes, *Discussion Paper 9095*, CORE, Louvain.
- [9] CAMPOS FILHO, F.F. (2001) *Algoritmos Numéricos*, LTC, Rio de Janeiro.

- [10] CHAN, K.C., KAROLYI, G.A., LONGSTAFF, F.A. and SANDERS, A.B. (1992) An Empirical Comparison of Alternative Models of the Short-Term Interest Rate, *The Journal of Finance*, vol. 47, 3, 1209-1227.
- [11] CLEUR, E.M. and MANFREDI, P. (1999) One Dimensional SDE Models, Low Order Numerical Methods and Simulation Based Estimation: A Comparison of Alternative Estimators. *Computational Economics*, 13 (2), 177-197.
- [12] COX, J., INGERSOLL, J. e ROSS, S. (1985) A Theory of the Term Structure of Interest Rates, *Econometrica*, 53, 385-407.
- [13] DANIELSSON, J. (1994) Stochastic Volatility in Asset Prices Estimation with Simulated Maximum Likelihood, *Journal of Econometrics*, 64, 375-400.
- [14] DURHAM, G.B. and GALLANT, A.R. (2002) Numerical Techniques for Maximum Likelihood Estimation of Continuous-Time Diffusion Processes, *Journal of Business & Economic Statistics*, 20, (3), 297-316.
- [15] FAN, J. and GIJBELS, I. (1996) *Local Polynomial Modelling and Its Applications*, Chapman & Hall, London
- [16] FAN, J. and YAO, Q. (1998) Efficient Estimation of Conditional Variance Functions in Stochastic Regression, *Biometrika*, 85, 3, 645-660.
- [17] FAN, J. and ZHANG, C. (2003) A Reexamination of Diffusion Estimators with Applications to Financial Model Validation, *Journal of the American Statistical Association*, 98, 118-134.
- [18] GENON-CATALOT, V., LAREDO, C. and PICARD, D. (1992). Nonparametric estimation of the diffusion coefficient by wavelet methods, *Scand. J. Statist.*, 19, 317-335.
- [19] GIBBONS, M.R. and RAMASWAMY, K. (1993) A Test of the Cox, Ingersoll and Ross Model of the Term Structure, *The Review of Financial Studies*, 6, 3, 619-658.

- [20] GOURIEROUX, C., MONFORT, A. and RENAULT, E. (1993) Indirect Inference, *Journal of Applied Econometrics*, 8, S, 85-118.
- [21] HALL, P. and CARROLL, R. J. (1989) Variance Function Estimation in Regression: The Effect of Estimation of the Mean, *J. Roy. Statist. Soc. Ser. B*, 51, 3-14.
- [22] HARDLE, W. and TSYBAKOV, A. (1997) Local Polynomial Estimators of the Volatility Function in Nonparametric Autoregression, *J. Econometrics*, 81, 233-242.
- [23] HULL, J.C. (1997) *Options, Futures, and other Derivatives*, 3rd ed., Prentice-Hall.
- [24] JENSEN, B. and POULSEN, R. (2002) Transition Densities of Diffusion Processes: Numerical Comparison of Approximation Techniques, *The Journal of Derivatives*, Summer, 9, (4), 18-32.
- [25] JIANG, G.J. and KNIGHT, J.L. (1997) A Nonparametric Approach to the Estimation of Diffusion Processes, with an Application to a Short-Term Interest Rate Model, *Econometric Theory*, 13, 615-645.
- [26] KLOEDEN, P.E., PLATEN, E., SCHURZ, H. and SØRENSEN, M. (1996) On Effects of Discretization on Estimators of Drift Parameters for Diffusion Processes, *Journal of Applied Probability*, 33, 1601-1076.
- [27] LEPSKI, O. (1990) One Problem of Adaptive Estimation in Gaussian White Noise, *Theory Probability Applied*, 35, 459-470.
- [28] LEPSKI, O. V., MAMMEN, E. and SPOKOINY, V.G. (1997) Optimal Spatial Adaptation to Inhomogeneous Smoothness: An Approach Based on Kernel Estimates with Variable Bandwidth Selectors, *The Annals of Statistics*, 25, 929-947.
- [29] LEPSKI, O. V. and SPOKOINY, V.G. (1997) Optimal Pointwise Adaptive Methods in Nonparametric Estimation, *The Annals of Statistics*, 25, 2512-2546.
- [30] MERCURIO, D. and SPOKOINY, V. (2000) Statistical Inference for Time-inhomogeneous Volatility Models, <http://www.wias-berlin.de>, WIAS preprint 583, Berlin.

- [31] MILSTEIN, G. and NUSSBAUM, M. (1998) Diffusion Approximation for Nonparametric Autoregression, *Probability Theory and Related Fields*, 112, 535-543.
- [32] ØKSENDAL, B. (1998) *Stochastic Differential Equations: An Introduction with Applications*, 5th ed., Springer-Verlag, Berlin.
- [33] OVERBECK, L. and RYDÉN, T. (1997) Estimation in the Cox-Ingersoll-Ross Model, *Econometric Theory*, 13, 430-461.
- [34] PEDERSEN, A.G. (1995) A New Approach to Maximum Likelihood Estimation for Stochastic Differential Equations Based on Discrete Observations, *Scandinavian Journal of Statistics*, 22, 55-71.
- [35] PRESS, W. H., TEUKOLSKY, S. A., VETTERLING, W. T. and FLANNERY, B. P. (1992) *Numerical Recipes in C: The Art of Scientific Computing*, 2nd. ed., Cambridge University Press.
- [36] RUPPERT, D., WAND, M. P., HOLST, U. and HOSSJER, O. (1997) Local Polynomial Variance Function Estimation, *Technometrics*, 39, 262-273.
- [37] SANTA-CLARA, P. (1996) Simulated Likelihood Estimation of Diffusions with an Application to the Short Term Interest Rate, *Tese de Doutorado*
- [38] SMITH JR, A.A. (1993) Estimating Nonlinear Time-Series Models Using Simulated Vector Autoregressions, *Journal of Applied Econometrics*, 8, S63-S84.
- [39] SPOKOINY, V.G. (2000) Adaptive Drift Estimation for Nonparametric Diffusion Model, *The Annals of Statistics*, 28 (3), 815-836.
- [40] STANTON, R. (1997) A Nonparametric Model of Term Structure Dynamics and the Market Price of Interest Rate Risk, *The Journal of Finance*, 52, 1973-2002.
- [41] VASICEK, O. (1977) An Equilibrium Characterization of the Term Structure, *Journal of Financial Economics*, 5, 177-188.