SERVIÇO DE PÓS-GRADUAÇÃO DO ICMC-USP

Data de Depósito: 01.02.2002

Assinatura:

Inferência em processos de difusão com
observações parciais e determinação da medida
martingale equivalente na precificação de
opções

Ulisses Umbelino dos Anjos

Orientador: Prof. Dr. Marinho Gomes de Andrade Filho

Dissertação apresentada ao Instituto de Ciências Matemáticas e de Computação - ICMC-USP, como parte dos requisitos para obtenção do título de Mestre em Ciências de Computação e Matemática Computacional.

USP – São Carlos
Janeiro/2002

A Comissão Julgadora:

Prof. Dr. Marinho Gomes de Andrade Filho

A handwritten signature in blue ink, written over a horizontal line. The signature is stylized and appears to be 'Marinho Gomes de Andrade Filho'.

Prof. Dr. João Bosco Ribeiro do Val

A handwritten signature in blue ink, written over a horizontal line. The signature is stylized and appears to be 'João Bosco Ribeiro do Val'.

Prof. Dr. Marcelo Dutra Fragoso

A handwritten signature in blue ink, written over a horizontal line. The signature is stylized and appears to be 'Marcelo Dutra Fragoso'.

Abstract

In this work, an application of the stochastic differential equations to the theory of options was carried out. This theory had great impulse with the work by [Black & Scholes, 73]. In their work, amongst other premises, they consider that the log-returns of actives had a normal distribution. In this work, three models were considered, one of which is the log-normal distribution used by Black & Scholes. The other two models are the linear diffusion and the Ornstein-Uhlenbeck. For these three models, the Equivalent Martingales Measures were determined, this was made using the theorem of Cameron-Martin-Girsanov, see [Friedman, 75]. Discrete versions of these models were also analysed, which were obtained by the approximation of Euler, see [Kloeden & Platen, 95]. The objective was to compare the results for the continuous models to the results obtained for the discrete models. It was also made the analysis of the estimates of parameters of the continuous models. In this analysis, the Classic approach and the Bayesian approach were used. Firstly, a comparison was carried out for the estimates obtained for these two approaches and a further analysis of the asymptotic behaviour of these estimates was also made.

Resumo

Neste trabalho foi feita uma aplicação das Equações Diferenciais Estocásticas á teoria da Precificação de Opções. Esta teoria teve grande impulso com o trabalho [Black & Scholes, 73]. Black e Scholes em seu trabalho entre outras premissas feitas consideraram que os log-retornos dos ativos tinha uma distribuição normal. Aqui neste trabalho foram considerados três modelos, um deles é a difusão log-normal utilizada por Black-Scholes os outros dois modelos são a difusão linear e o processo de Ornstein-Uhlenbeck. Para estes três modelos foram determinadas as Medidas Martingales Equivalentes, isto foi feito utilizando o Teorema de Cameron-Martin-Girsanov, veja [Friedman, 75]. Também foram analisadas versões discretas destes modelos obtidas pela aproximação de Euler, veja [Kloeden & Platen, 95]. O objetivo foi comparar os resultados obtidos com os modelos contínuos com os resultados obtidos com os modelos discretos. Também se fez uma análise dos estimadores dos parâmetros dos modelos contínuos. Nesta análise foi utilizada a abordagem Clássica e a abordagem Bayesiana. Primeiramente se fez uma comparação das estimativas obtidas por estas duas abordagens e posteriormente uma análise do comportamento assintótico desses estimadores.

Agradecimentos

Eu gostaria de fazer três agradecimentos.

- O primeiro a Deus
 - pelos pais que Ele me deu que sempre me amaram e nunca deixaram de acreditar em mim;
 - pela esposa e pelo filho que Ele me deu que souberam suportar a pouca atenção dada a eles durante estes anos;
 - pela saúde que nunca deixou me faltar.
- Ao apoio financeiro do CNPQ sem o qual seria difícil concluir este projeto;
- Às funcionárias da biblioteca Prof. Achille Bassi que sempre foram tão gentis e atenciosas.

Penso logo existo.

Descartes

Existo onde não penso

Lacan

As verdades humanas são relativas, mutáveis e efêmeras;

A verdade de Deus é absoluta, imutável e perene.

Sumário

1 Introdução	1
1.1 Motivação	1
1.2 Objetivos	2
1.3 Organização do Trabalho	3
2 Opções: Propriedades, Limites Máximo e Mínimo	11
2.1 Derivativos	11
2.1.1 Contratos Futuros	12
2.1.2 Opções	13
2.2 Propriedades Básicas das Opções	16
2.3 Fatores que Afetam o Preço das Opções	16
2.3.1 O preço da ação e o preço de exercício	17
2.3.2 O tempo para o vencimento	17
2.3.3 A volatilidade	18
2.3.4 A taxa de juro livre de risco	19
2.3.5 Dividendos	20
2.4 Limites Máximos e Mínimos para os Preços de Opções	20
2.4.1 Os Limites Máximos para Opções de Compra e de Venda	21
2.4.2 O limite mínimo para opções de compra de ações sem dividendos	22
2.4.3 O limite mínimo para opções europeias de venda de ações sem dividendos	24

	x
2.5 O Exercício Antecipado de Opções de Compra Americana de Ações sem Dividendos.....	24
2.5.1 O Exercício Antecipado de Opções de Venda Americana de Ações sem Dividendos.....	25
2.6 A Paridade entre as Opções de Compra e Venda Europeias	26
2.7 A Relação entre os Preços de Opções de Compra e de Venda Americanas	28
3 Precificação das Opções a Tempo Discreto.....	31
3.1 Considerações Iniciais	32
3.2 Medida Martingale Equivalente (MME)	37
3.3 Avaliando Opções pela Medida Martingale Equivalente	39
3.4 Premissas do Modelo	41
3.4.1 O Modelo Binomial	42
3.4.2 Avaliando uma Opção pelo Modelo Binomial	43
3.4.3 Avaliação da Opção Europeia pelo Modelo Binomial	45
3.5 Análise dos Modelos: Versões a Tempo Discreto	46
3.5.1 Avaliação das Opções Europeias	46
4 Precificação das Opções a Tempo Contínuo	55
4.1 Considerações Iniciais	56
4.2 Premissas do Modelo	59
4.3 O Modelo Log-Normal: Black-Scholes	60
4.3.1 A Abordagem de Black-Scholes	64
4.4 Análise dos Modelos Difusão Linear e Ornstein-Uhlenbeck	66
5 Inferências dos Parâmetros de EDE Cuja Solução Forte é Conhecida ...	71
5.1 Análise Clássica dos Modelos.....	72
5.1.1 Modelo 1: $dX_t = \mu dt + \sigma dB_t$	72
5.1.2 Modelo 2: $dX_t = \mu X_t dt + \sigma dB_t$	74
5.1.3 Modelo 3: $dX_t = \mu X_t dt + \sigma X_t dB_t$	76

	xi
5.2 Análise Bayesiana dos Modelos	78
6 Inferência nas Versões Discretizadas	85
7 Resultados	89
7.1 Estimação dos Parâmetros dos Modelos Contínuos	89
7.1.1 Modelo 1: $dX_t = \mu dt + \sigma dB_t$	89
7.1.2 Modelo 2: $dX_t = \mu X_t dt + \sigma dB_t$	91
7.1.3 Modelo 3: $dX_t = \mu X_t dt + \sigma X_t dB_t$	91
7.2 Análise da Performance dos Estimadores a Tempo Contínuo	92
7.3 Análise da Performance dos Estimadores Clássico e Bayesiano	92
7.3.1 Modelo 1: $\mu_R = 0.35, \sigma_R^2 = 1$ e $\Delta t = 1$	93
7.3.2 Modelo 2: $\mu_R = -0.5, \sigma_R^2 = 0.04$ e $\Delta t = 1$	94
7.3.3 Modelo 3: $\mu_R = 3 \times 10^{-4}, \sigma_R^2 = 3.6 \times 10^{-7}$ e $\Delta t = 1$	95
7.4 Valor das Opções de Venda Europeias	96
8 Conclusão	103
9 Propostas Futuras	105
9.1 Precificação da Opção do tipo Americana	105
9.2 Inferência em Processos Onde a Densidade de Transição é Desconhecida	106
Bibliografia	109
A Abordagem Clássica	113
A.1 Reparametrização	113
A.1.1 Invariância	113
A.1.2 Relação entre a Função de Informação de θ e λ	115
B Abordagem Bayesiana	117
B.1 Inferência Bayesiana	118

	xii
B.1.1 Densidade Preditiva	118
B.1.2 Verossimilhança	119
B.1.3 Discriminação de Modelos	119
C Elementos da Análise Estocástica Discreta	123
C.1 Base Estocástica Discreta	123
C.2 Martingales	124
C.2.1 Convergência de Martingales	125
C.2.1 Decomposição de Doob	126
C.2.1 Transformações de Martingales	127
C.2.1 Tempos de Parada	128
D Elementos da Análise Estocástica Contínua	131
D.1 Considerações Iniciais	131
D.2 Martingales	132
D.3 Equações Diferenciais Estocásticas	135
D.3.1 Dificuldades na Definição da Integral Estocástica	136
D.4 Conexão Entre o Cálculo Estocástico e as Equações Diferenciais Parciais	142
D.4.1 Relação entre EDP e EDE via Teoria dos Processos de Markov	143
D.4.2 Relação entre EDP e EDE via Cálculo Estocástico	150
D.5 Transformações de Medida: A Formula de Cameron-Martin-Girsanov	153

Capítulo 1

Introdução

1.1 Motivação

Desde as publicações [Black & Scholes, 73] e [Merton, 73a] o impacto da matemática aplicada na modelagem econométrica financeira resultou em grandes avanços. Tempo e incerteza são os elementos centrais que influenciam o comportamento econômico financeiro. A complexidade da interação entre eles produziu grande desafio intelectual para o estudo das finanças. Analisar os efeitos desta interação as vezes requer sofisticadas ferramentas analíticas. De fato, conhecimento matemático avançado torna-se um pré-requisito para pesquisadores nesta área. Não obstante, toda esta complexidade matemática, a teoria financeira tem tido uma direta e significativa influência na prática financeira. Uma comparação casual das práticas correntes com aquelas de 30 anos atrás é suficiente para notar o impacto da teoria na análise do problema de modelos financeiros. Passadas quase três décadas, o modelo a tempo contínuo tem provado ser uma ferramenta versátil e produtiva no desenvolvimento de modelos para problemas de finanças. Apesar de matematicamente mais complexa, a formulação a tempo contínuo muitas vezes produz ambas soluções teóricas mais precisas e hipóteses empíricas mais refinadas.

Neste trabalho vamos estudar os processos difusivos, que são modelados por equações diferenciais estocásticas, e fazer a inferência dos parâmetros deste processo. Os processos de difusão constituem uma classe dos processos estocásticos que são caracterizados por duas propriedades: A propriedade Markoviana e a continuidade das trajetórias. Os processos de difusão que iremos trabalhar são os processos de difusão de Itô. A difusão de Itô é um processo estocástico $X(t, \omega) : [0, \infty) \times \Omega \rightarrow \mathbb{R}^n$ satisfazendo a equação diferencial estocástica da forma

$$dX_t = a(X_t; \Theta) dt + b(X_t; \Theta) dB_t \quad (1.1)$$

onde $a(X_t; \Theta)$ e $b(X_t; \Theta)$ são as funções drift e volatilidade, respectivamente, dependendo do estado de X , no tempo t , e um vetor parâmetro Θ desconhecido, e dB_t é o incremento do processo de Wiener.

1.2 Objetivos

Os objetivos deste trabalho são dois:

1. Inferência dos parâmetros da Equação Diferencial Estocástica(EDE) para dados observados em tempos discretos. Neste trabalho vamos estudar três modelos, a saber:

- (a) $dX_t = \mu dt + \sigma dB_t$ (Difusão Linear);
- (b) $dX_t = \mu X_t dt + \sigma dB_t$ (Processo de Ornstein-Uhlenbeck);
- (c) $dX_t = \mu X_t dt + \sigma X_t dB_t$ (Difusão log-normal).

onde B_t é o processo de Wiener.

É feita uma comparação entre as abordagens Clássica e Bayesiana e um estudo do comportamento assintótico dos estimadores clássico e bayesiano. Esta análise é feita utilizando a simulação de Monte Carlo.

2. Inferência dos parâmetros da versão discreta dos modelos acima obtida pela aproximação de Euler, veja [Kloeden & Platen, 95]. Assim temos que

$$\begin{aligned} dS_t &= \mu dt + \sigma dB_t \xrightarrow{\text{Euler}} S_{t_{i+1}} = S_{t_i}^1 + \mu \Delta t + \sigma \Delta B_{t_{j+1}} \\ dS_t &= S_t \mu dt + \sigma dB_t \xrightarrow{\text{Euler}} S_{t_{i+1}} = S_{t_i}^1 + S_{t_i}^1 \mu \Delta t + \sigma \Delta B_{t_{j+1}} \\ dS_t &= S_t \mu dt + S_t \sigma dB_t \xrightarrow{\text{Euler}} S_{t_{i+1}} = S_{t_i}^1 + S_{t_i}^1 \mu \Delta t + S_{t_i}^1 \sigma \Delta B_{t_{j+1}} \end{aligned} \quad (1.2)$$

portanto as versões discretas são dadas por

$$\begin{aligned} S_{t_{i+1}} &= S_{t_i}^1 + \mu \Delta t + \sigma \Delta B_{t_{j+1}} \xrightarrow{\text{Versão discreta}} S_{t+1} = S_t + \mu_{\Delta t} + \sigma_{\Delta t} Z_{t+1} \\ S_{t_{i+1}} &= S_{t_i}^1 + S_{t_i}^1 \mu \Delta t + \sigma \Delta B_{t_{j+1}} \xrightarrow{\text{Versão discreta}} S_{t+1} = S_t + S_{t+1} \mu_{\Delta t} + \sigma_{\Delta t} Z_{t+1} \\ S_{t_{i+1}} &= S_{t_i}^1 + S_{t_i}^1 \mu \Delta t + S_{t_i}^1 \sigma \Delta B_{t_{j+1}} \xrightarrow{\text{Versão discreta}} S_{t+1} = S_t^1 + S_{t+1} \mu_{\Delta t} + S_{t+1} \sigma_{\Delta t} Z_{t+1} \end{aligned} \quad (1.3)$$

onde $Z_{t+1} \sim N(0,1)$. O objetivo desta análise é poder utilizar estes modelos a tempo discreto na formulação de modelos de precificação de opções.

3. Formulação do modelo de precificação de derivativos, determinação da Medida Martingale Equivalente e determinação do valor da opção do tipo europeia para cada um destes modelos. Foram feitas as formulações na estrutura contínua e discreta. Foi feita a comparação dos resultados do valor da opção a tempo contínuo e a tempo discreto.

1.3 Organização do Trabalho

Este trabalho está organizado da seguinte maneira:

Capítulo 2:

Neste capítulo foram colocadas algumas noções sobre os mercados de Derivativos, pois o trabalho gira em torno de modelos de precificação de opções e como é visto as opções são um tipo de Derivativo. Neste capítulo também se fez um pequeno resumo das propriedades básicas das opções assim como uma breve análise dos fatores que afetam o preço das opções. Estes fatores são:

- O preço da ação S_t e o valor de exercício K da ação fornecida pela opção;
- O tempo para o vencimento T , conhecida como maturidade da opção;
- A volatilidade σ da opção;
- A taxa de juros livre de risco r ;
- Os dividendos pagos pela ação.

Aqui neste trabalho, para que os modelos tenham uma solução analítica, consideramos que o coeficiente de difusão σ conhecida como volatilidade da opção e a taxa de juros livre de risco r são constantes e que as ações não pagam dividendos durante a vida da opção.

Analizamos também os limites máximo e mínimo que as opções podem assumir, onde o limite mínimo para as opções de compra do tipo européia é

$$\max(S_T - K, 0) \quad (1.4)$$

e para a opção de venda européia

$$\max(K - S_T, 0) \quad (1.5)$$

estes dois resultados são importantes pois eles são as condições finais das equações diferenciais parciais que iremos deduzir para cada um dos modelos contínuos. Para os modelos discretos eles também são a condição final de onde estimamos o valor mínimo que a opção deverá ter. É importante ressaltar que os todos os valores das opções que iremos determinar são valores mínimos. É interessante notar que muitas vezes na prática estes valores mínimos não

são obedecidos criando-se no mercado o que chamamos de possibilidade de arbitragem, isto é, a possibilidade de se obter lucro sem risco. Infelizmente a maioria destas oportunidades de arbitragem tem curta duração e só podem ser aproveitadas pelas instituições financeiras tais como bancos e corretoras, pois elas estão livres de vários custos de transação que normalmente incidem sobre o investidor. Este é um negócio tão atrativo que estas instituições financeiras tem equipes especializadas em garimpar no mercado estas arbitragens.

Capítulo 3

Neste capítulo foi estudado o problema da precificação de opções na estrutura discreta, isto é, construindo modelos que utilizam a teoria dos processos estocásticos discretos.

A parte inicial é dedicada a exposição dos conceitos básicos que são necessários na construção e no entendimento do modelo de precificação. Alguns teoremas, lemas, corolários e proposições são dados e a maioria deles demonstrados. Nesta parte inicial é exposto o conceito que é a base para toda a construção do modelo de precificação que é a Medida Martingale Equivalente (MME). É visto que sobre a hipótese de termos um mercado completo, é possível determinar uma nova medida sobre a qual os processos dos preços das ações descontadas se comportam como um martingale e pelo fato do mercado ser completo esta medida é única. É visto também que essa medida nos fornece uma estratégia de transação que duplica o fluxo de caixa de cada derivativo do mercado, portanto com esta estratégia temos que o processo valor do derivativo $(H_t)_{t \in [0, T]}$ é igual ao processo valor da carteira $(V_t^\varphi)_{t \in [0, T]}$ quando a carteira V^φ é estabelecida sobre o conjunto das estratégias admissíveis Φ_a , estas estratégias são somente as estratégias duplicantes, assim:

$$(H_t)_{t \in [0, T]} = (V_t^\varphi)_{t \in [0, T]} \quad (1.6)$$

Portanto o valor da opção sobre estas condições é:

$$H_t = \beta_t^{-1} E_{\mathbb{Q}} [\beta_T H_T | \mathcal{F}_t] \quad (1.7)$$

para todo $t \in [0, T]$, $T \in \mathbb{Z}^+$. Onde β_t é o fator de desconto em t e H_T é a condição final que depende de cada derivativo em questão.

A segunda parte do capítulo 3 é dedicada a construção do modelo propriamente dito. Inicialmente consideramos o modelo binomial multiperíodo,

$$S_{t+1} = S_t^1 + S_t^1 \Delta b_{t+1} \quad (1.8)$$

onde $\Delta b_{t+1} \sim \text{bernoulli}(p, (1-p))$, onde

$$\Delta b_{t+1} = \begin{cases} u & \text{com probabilidade } p \\ d & \text{com probabilidade } 1 - p \end{cases} \quad (1.9)$$

onde $0 < d < u < 1$ e $0 < p < 1$ ou seja a cada passo o processo poderá subir para um valor $(1 + u) S_t^1$ ou descer para um valor $(1 + d) S_t^1$. Note que a probabilidade p não pode assumir os valores 0 e 1, pois com estes valores haverá possibilidades de arbitragem o que contraria uma das premissas do modelo: o mercado é completo. Deste modo temos que o valor da opção de compra do tipo europeu c_t é dado por

$$c_t = c_t = \beta_{T-t} S_t \sum_{j=\tau}^{T-t} \binom{T-t}{j} q^j (1-q)^{T-t-j} (1+u)^j (1+d)^{T-t-j} - \beta_{T-t} K \sum_{j=\tau}^{T-t} \binom{T-t}{j} q^j (1-q)^{T-t-j} (1+u)^j (1+d)^{T-t-j} \quad (1.10)$$

onde q é a probabilidade do processo subir para um valor $(1 + u) S_t^1$, na medida martingale equivalente $\mathbb{Q}$, K é valor de exercício da opção e τ é determinado pela condição final da opção. A equação (1.10) é conhecida como fórmula de Cox-Ross-Rubinstein, veja [Cox et al, 79].

Após estas considerações iniciais, se fez a análise dos três modelos discretos

1. $S_{t+1} = S_t + \mu_{\Delta t} + \sigma_{\Delta t} Z_{t+1}$
2. $S_{t+1} = S_t + S_t \mu_{\Delta t} + \sigma_{\Delta t} Z_{t+1}$
3. $S_{t+1} = S_t + S_t \mu_{\Delta t} + S_t \sigma_{\Delta t} Z_{t+1}$

como já dito anteriormente estes modelos podem ser deduzidos utilizando-se a aproximação de Euler dos modelos contínuos análogos.

Sobre a medida martingale equivalente $\mathbb{Q}$ foi visto que $S_{t+1} | \mathcal{F}_t$ para os três modelos acima tem a seguinte forma

$$\begin{aligned} S_{t+1} &= S_t + S_t r + \sigma_{\Delta t} Z_{t+1} \\ S_{t+1} &= S_t + S_t r + \sigma_{\Delta t} Z_{t+1} \\ S_{t+1} &= S_t + S_t r + S_t \sigma_{\Delta t} Z_{t+1} \end{aligned} \quad (1.11)$$

note que os três modelos tem os mesmos retornos esperados. Isto é exatamente o que faz a medida martingale equivalente, em outras palavras, quando os investidores “olham” para o mercado numa situação de indiferença ao risco o retorno esperados de todos os ativos que compõe o mercado é a taxa de juros livre de risco r . Deste modo, o primeiro e o segundo modelo tem o mesmo valor da opção de compra européia, dado por

$$c_t = \frac{\sigma_{\Delta t} \exp\left(-\frac{A^2}{2}\right) \sum_{j=1}^{T-t} (1+r)^{-j}}{\sqrt{2\pi}} + \left(S_t - K(1+r)^{t-T}\right) \times [1 - \Phi(A)] \quad (1.12)$$

a opção de venda é dada pela seguinte relação

$$p_t = c_t + K(1+r)^{t-T} - S_t \quad (1.13)$$

esta relação é conhecida como relação de paridade entre opção de compra e venda, esta relação será vista com mais detalhes no capítulo 2.

O modelo 3 tem o seguinte valor para a opção de compra europeia

$$c_t = S_t \exp\left[\frac{[\sigma_{\Delta t}(T-t)]^2}{2}\right] \left[1 - \Phi\left(A - \sigma\sqrt{T-t}\right)\right] - Ke^{-r(T-t)} \times [1 - \Phi(A)] \quad (1.14)$$

onde $\Phi(\cdot)$ é a função densidade acumulada da normal padrão.

Capítulo 4 :

Neste capítulo são analisados os modelos a tempo contínuo. A abordagem utilizada foi diferente daquela utilizada por [Black & Scholes, 73]. Aqui utiliza-se as ferramentas do cálculo estocástico para deduzir as fórmulas das opções. No apêndice D, apresentaram-se os principais resultados necessários ao desenvolvimento dessa parte. Este apêndice é bastante interessante principalmente a seção sobre as conexões entre as Equações Diferenciais Parciais e as Equações Diferenciais Estocásticas. É com estas ferramentas que foi mostrado que a equação diferencial parabólica de Black-Scholes, veja [Black & Scholes, 73],

$$\frac{\partial H(t, S)}{\partial t} + rS \frac{\partial H(t, S)}{\partial S} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 H(t, S)}{\partial S^2} - rH(t, S) = 0 \quad (1.15)$$

com condição final $H(T, S)$, tem representação estocástica dado por

$$H_t = e^{-r(T-t)} E_{\mathbb{Q}}[H_T | \mathcal{F}_t] \quad (1.16)$$

Na seção subsequente se fez uma breve exposição da abordagem de Black-Scholes na determinação da sua famosa fórmula de precificação da opção europeia

$$c_t = S_t \left[1 - \Phi\left(A - \sigma\sqrt{T-t}\right)\right] - Ke^{-r(T-t)} \times [1 - \Phi(A)] \quad (1.17)$$

esta fórmula é obtida resolvendo a equação (1.15) ou calculando $E_{\mathbb{Q}}[H_T | \mathcal{F}_t]$. É óbvio que é mais fácil calcular o valor esperado.

Após isto se fez a análise para os modelos 1 e 2. Da mesma maneira como ocorreu para o caso discreto aqui também se tem o mesmo valor da opção para os dois modelos

$$c_t = e^{-r(T-t)} \sqrt{\frac{\sigma^2}{2r}(e^{2r(T-t)} - 1)} \frac{\exp\left(-\frac{A^2}{2}\right)}{\sqrt{2\pi}} + (S_t^1 - Ke^{-r(T-t)}) \times [1 - \Phi(A)] \quad (1.18)$$

Capítulo 5 :

Neste capítulo se fez a inferência dos modelos a tempo contínuo. A inferência em modelos a tempo contínuo é estudada de duas perspectivas diferentes. A primeira considera a hipótese de que se pode observar continuamente um processo durante um intervalo qualquer $[0, T]$. A segunda perspectiva, é a que na prática utilizamos, considera que só podemos observar o processo em tempos discretos. No caso de modelos a tempo contínuo, os tempos entre as observações pode ser igualmente espaçados ou podem ser diferentes, e neste caso podem até ser aleatórios. Neste trabalho se fez a inferência sobre duas abordagens diferentes, a primeira é a abordagem Clássica e a segunda a Bayesiana em ambas as abordagens é considerada que o processo é observado em tempos discretos e que os intervalos entre as observações Δt é constante.

A inferência Clássica tem como principal objeto para análise dos dados a função de verossimilhança. Quando se supõe que somente observações em tempos discretos estão disponíveis, o estudo deste problema é dividido em duas partes. A primeira, considerando que a solução forte do processo é conhecida, isto é, se conhece a densidade de transição do processo. A segunda quando isto não é possível. Aqui consideramos apenas modelos que se conhece a solução forte do processo. Assim tem-se que:

1. **Modelo 1:** $dX_t = \mu dt + \sigma dB_t$ a densidade de Transição é dada por
$$X_{t_{i+1}} | X_{t_i} \sim N [X_{t_i} + \mu(t_{i+1} - t_i), \sigma^2(t_{i+1} - t_i)] \quad (1.19)$$

2. **Modelo 2:** $dX_t = \mu X_t dt + \sigma dB_t$

$$X_{t_{i+1}} | X_{t_i} \sim N \left[X_{t_i} e^{\mu \Delta t}, \frac{\sigma^2}{2\mu} (e^{2\mu \Delta t} - 1) \right] \quad (1.20)$$

3. **Modelo 3:** $dX_t = \mu X_t dt + \sigma X_t dB_t$

$$\ln(X_{t_{i+1}}) | \ln(X_{t_i}) \sim N \left[\ln(X_{t_i}) + \left(\mu - \frac{\sigma^2}{2}\right) \Delta t, \sigma^2 \Delta t \right] \quad (1.21)$$

Quando se considera o problema da inferência, no contexto bayesiano, a estimativa dos parâmetros é baseada na densidade a posteriori. Se uma solução forte do processo é possível, i.e., a Equação Diferencial Estocástica,

$$dX_t = a(X_t; \Theta) dt + b(X_t; \Theta) dB_t \quad (1.22)$$

pode ser resolvida analiticamente na forma de Itô, para $t \in [0, T]$, então a densidade de transição de X_t está disponível em forma fechada e a inferência da densidade a posteriori é direta. A densidade a posteriori é dada por

$$p(\Theta|X) \propto p(\Theta) p(X|\Theta) \quad (1.23)$$

esta é a posteriori não normalizada, como podemos ver ela é composta pela verossimilhança $p(X|\Theta)$ mais a densidade a priori do vetor de parâmetros $p(\Theta)$. Como neste trabalho são utilizadas somente grandes séries, foi utilizado somente priori não informativa.

Foi utilizada tanto na abordagem Clássica quanto na Bayesiana uma reparametrização dos modelos 2 e 3, sem as quais teríamos que utilizar algoritmos numéricos para extrair os estimadores dos parâmetros.

Capítulo 6 :

Neste capítulo se fez a análise das versões discretas dos modelos contínuos obtidos pela aproximação de Euler, veja [Kloeden & Platen, 95]. Como o objetivo da utilização destes modelos discretos é apenas a de comparar os resultados do valor da opção feitas a tempo contínuo com o valor feita a tempo discreto foi utilizada nas estimativas dos parâmetros apenas a abordagem Clássica, pois nestas versões discretas as verossimilhanças são bastante simples portanto não precisando de uma análise mais detalhada que a justifique.

Capítulo 7 :

Nesta análise foram utilizadas séries geradas. Desta maneira podemos avaliar e comparar melhor as metodologias utilizadas no processo de estimação.

Foram geradas séries com 1000 elementos para fazer a estimação dos parâmetros dos modelos contínuos. Numa segunda etapa foram geradas 1000 amostras de tamanho 180 e com elas foram feitas o estudo do comportamento assintótico dos estimadores dos modelos contínuos.

Após a etapa da estimação dos parâmetros foram geradas mais 2 amostras de tamanho 1200 para cada um dos modelos, onde os primeiros 1000 foram utilizados na estimação dos parâmetros e os 200 últimos para fazer a comparação dos resultados dos modelos a tempo contínuo com os dos modelos a tempo discreto. A taxa de juros livre de risco utilizada será de 1.51% ao mês, portanto para uma capitalização diária teremos $r = 5 \times 10^{-4}$ e para uma capitalização contínua teremos $r = 4.9988 \times 10^{-4}$. O tempo de exercício das opções será $T =$

30 dias. O valor de exercício das opções K serem escolhidos aleatoriamente de $N(S_T, 0.01S_T)$. Onde S_T é conhecido pois a série é gerada. Adotaremos $t = 0$ deste modo $(T - t) = T = 30$. Como já foi visto na parte teórica só precisaremos estimar σ , pois o retorno esperado quando se faz a avaliação do risco neutro é a taxa de juros livre de risco r .

Este trabalho é encerrado com um capítulo de conclusão (Capítulo 8) e com um capítulo de propostas futuras (Capítulo 9) que indicam a continuidade do trabalho apresentado aqui.

Capítulo 2

Opções: Propriedades, Limites Máximo e Mínimo

Nos últimos anos, o mercado de Derivativos tornou-se extremamente importante no mundo das finanças e dos investimentos. Atingimos um estágio em que se faz necessário a todos os profissionais de finanças compreender o funcionamento desse mercado como instrumentos na administração do risco de carteiras, sua utilização e fatores que afetam a formação dos preços futuros desse mercado.

Neste trabalho se fez uma aplicação das equações diferenciais estocásticas na teoria da formação de preços de derivativos. Mais especificamente, iremos trabalhar com modelos de precificação de opções.

Esta seção tem o objetivo de introduzir alguns conceitos e notações que serão utilizadas quando do desenvolvimento dos modelos. Também será mostrado os limites máximos e mínimos que as opções podem assumir.

Existem boas referências sobre derivativos, entre elas podemos citar [Bodie & Merton, 99], [Hull, 96] e [Jorion, 98].

2.1 Derivativos

Os contratos futuros e de opções são exemplos do que chamamos de derivativos ou produtos derivativos, que podem ser definidos como títulos cujos valores dependem dos valores de outras variáveis mais básicas. Uma opção da ação da Petrobrás, por exemplo, é um derivativo, porque seu valor depende do preço da ação da Petrobrás; um contrato futuro de trigo também é um derivativo, pois seu valor depende do preço do trigo; e assim por diante.

Nos últimos anos, os bancos de investimento têm sido muito criativos no desenvolvimento de novos produtos derivativos, com o intuito de satisfazer as necessidades de seus clientes. Normalmente, tais títulos não são negociados em bolsa, mas vendidos por instituições financeiras no mercado de balcão a seus clientes ou incorporados a emissão de ações ou títulos para torná-los mais atrativos aos investidores. Alguns desses títulos derivativos são semelhantes aos contratos futuros e de opções negociados em bolsa; outros são bem mais complexos. A possibilidade de desenvolver títulos derivativos novos e interessantes parece realmente ilimitada. Entre os derivativos mais conhecidos estão o contrato futuro e as opções.

2.1.1 Contratos Futuros

Um contrato futuro é o compromisso de comprar ou vender determinado ativo numa data específica no futuro, por um preço previamente estabelecido.

Sempre que duas partes concordam em negociar algum item no futuro a um preço predeterminado, elas estão participando de um contrato a termo. Muitas vezes, as pessoas assinam um contrato dessa natureza sem saber que esse é o nome dado a esse tipo de transação.

Por exemplo, você pode estar planejando uma viagem de São Paulo a Austrália para daqui a um ano. Faz sua reserva agora, e o encarregado de reservas da companhia aérea informa-lhe que você pode fechar o voo ao preço de R\$1.000,00 ao comprometer-se agora, ou pode pagar a passagem pelo preço em vigor no dia do embarque. Em ambos os casos, o pagamento será efetuado no dia em que você for viajar. Se decidir fechar com o preço de R\$1.000,00, você fez um contrato a termo com a companhia aérea.

Ao fazer um contrato a termo, você elimina o risco de o preço da passagem subir além de R\$1.000,00. Se o preço do bilhete subir para R\$1.500,00 dentro de um ano, você ficará feliz em saber que teve o bom senso de “travar” o preço em R\$1.000,00. Por outro lado, se o preço da passagem baixar para R\$500,00 no dia de sua partida, você terá de pagar os R\$1.000,00 contratados. Nesse caso, irá lamentar a decisão.

As características principais dos contratos a termo e a terminologia usada para descrevê-los são as seguintes:

- Duas partes concordam em trocar um determinado item no futuro a um preço especificado agora, o preço a termo. O preço para entrega imediata do item chama-se preço à vista (spot price);

- Dinheiro não é transacionado no momento do acordo por nenhuma das partes;
- O valor de face do contrato é a quantidade de itens especificados no contrato multiplicados pelo preço a termo.
- A parte que se compromete a comprar o item especificado assume uma posição comprada (long position) e a parte que se compromete a vender o item assume uma posição vendida (short position).

Um contrato futuro é essencialmente um contrato a termo. Contudo, há importantes diferenças entre os dois. Os Contratos futuros são negociados numa bolsa organizada e suas características são por ela padronizadas; inversamente, os contratos a termo são acordos particulares entre duas instituições financeiras ou entre uma instituição financeira e seus clientes. Uma outra diferença fundamental é que em um contrato futuro, uma posição pode ser encerrada antecipadamente enquanto que no mercado a termo, não há possibilidade de troca de posições como ocorre no mercado futuro ou opções. Isto é, o comprador e o vendedor a termo continuam com o contrato até o seu vencimento, ainda que possam liquidá-lo antecipadamente.

A maioria dos contratos transacionados nas várias bolsas do mundo pode ser categorizada como futuros de mercadorias (em que o objeto de negociação é a mercadoria) ou como futuros financeiros (em que o objeto de negociação é um ativo financeiro, como um título ou uma carteira de ações). Novos contratos são propostos a todo o instante e não há dúvida de que os mercados futuros representam até hoje uma das grandes inovações da área financeira.

2.1.2 Opções

Os contratos de opções são negociados há menos tempo que os futuros, mas também são bastante populares entre os investidores.

Há basicamente dois tipos de opções: *calls* (opções de compra) e *puts* (opções de venda). Na *opção de compra*, o detentor (titular ou comprador da opção) tem o direito de comprar um ativo em certa data por determinado preço. Na *opção de venda*, o detentor tem o direito de vender um ativo em certa data por determinado preço. O preço do contrato (valor futuro pelo qual o bem será negociado) é conhecido como *preço de exercício* (strike price ou exercise price) e sua data (o dia em que a posição será exercida) são conhecidas como *data de vencimento* (expiration date, exercise date ou maturity). Uma *opção europeia* pode ser exercida somente

na data de vencimento. Uma *opção americana* pode ser exercida a qualquer momento, até o vencimento. O valor pago por uma opção é chamado de *prêmio da opção*. As opções podem estar dentro do preço (in-the-money), estar no preço (at-the-money) ou estar fora do preço (out-of-the-money). Uma opção dentro do preço proporciona ao seu titular um fluxo de caixa positivo, caso seja exercida imediatamente. Da mesma forma, uma opção no preço proporciona um fluxo de caixa zero e uma opção fora do preço um fluxo de caixa negativo. Se S é o preço da ação e K o preço de exercício, então:

- Uma opção de compra está:
 - dentro do preço quando $S > K$;
 - no preço quando $S = K$;
 - e fora do preço quando $S < K$.
- Uma opção de venda está:
 - dentro do preço quando $S < K$;
 - no preço quando $S = K$;
 - e fora do preço quando $S > K$.

É óbvio que uma opção só será exercida se estiver dentro do preço. Na ausência de custos operacionais, uma opção será sempre exercida na data de vencimento, se não o tiver sido anteriormente.

Devemos enfatizar que o titular de uma opção tem o direito de fazer algo; esse direito, porém, não precisa ser exercido. Esse fato diferencia os contratos futuros dos de opções, pois o comprador de um contrato futuro assume um compromisso de comprar um bem por determinado preço numa data futura. Por sua vez, o detentor de uma opção de compra pode escolher se irá comprar o bem por determinado preço numa data futura. Não há custos para realizar um contrato futuro (com exceção das margens); todavia, um investidor deve pagar um preço antecipado por um contrato de opções.

Há quatro tipos de participantes nos mercados de opções:

1. compradores de opções de compra;
2. vendedores de opções de compra;

3. compradores de opções de venda;
4. vendedores de opções de venda.

Os compradores são os que possuem posições compradas (long positions); já os vendedores possuem posições vendidas (short positions).

Colocar uma opção no mercado, seja ela uma opção de compra ou de venda, é conhecido como lançar uma opção.

Tanto os mercados futuros como os de opções têm tido um êxito notável. Um dos motivos do sucesso dos mercados futuros e de opções é que eles atraem vários tipos de participantes, dentre os quais destacamos três: hedgers, especuladores e arbitradores.

Operações de **hedge** são operações que tem como finalidade proteger-se contra possíveis perdas. Tanto os mercados futuros como o de opções, foram criados originalmente para atender às necessidades dos hedgers. O mercado futuro, por exemplo, é utilizado pelos produtores que querem manter um preço para sua produção e pelos comerciantes que querem garantir um preço para obter tal produto. Os contratos futuros, então, permitem que ambas as partes atinjam seus objetivos.

Se compararmos as estratégias de **hedge** com futuros e opções, verificaremos que existe uma profunda diferença entre o uso de futuros e o uso de opções para hedging. Os *contratos futuros* são designados para neutralizar o risco, fixando o preço que o hedger irá pagar ou receber pelo objeto de negociação; os *contratos de opções*, por sua vez, além de fornecerem uma segurança para que os investidores se protejam contra oscilações adversas de preços no futuro, permitem que eles tirem proveito de oscilações favoráveis de preços.

Enquanto os **hedgers** não desejam ficar expostos a movimentos adversos no preço de um ativo, os **especuladores** querem abrir posições opostas aos dos hedgers, apostando numa alta ou numa queda de preços. De fato, os especuladores são necessários para que o hedging seja possível, caso contrário um hedger desejando eliminar parte do risco de sua carteira não poderia fazê-lo se não houvesse alguém, o especulador, disposto a assumir este risco. Uma das estratégias mais utilizadas pelos especuladores são as estratégias de “daytrade”. “Daytrade” nada mais é do que a compra e a venda do ativo em um mesmo dia ou uma operação inversa na qual o investidor vende para em seguida recomprar o mesmo ativo a um preço menor.

Os **arbitradores** formam o terceiro grupo de participantes dos mercados derivativos. A arbitragem envolve conseguir um lucro sem risco. Em realidade, a existência de arbitradores ávidos por lucro torna pouco provável que ocorra uma grande disparidade entre os preços de um ativo. Uma das premissas básicas na construção dos modelos de precificação que foram desenvolvidos é exatamente esta que no mercado não existem possibilidade de arbitragem.

2.2 Propriedades Básicas das Opções

Há duas pontas em cada contrato de opções. Numa delas está o investidor que assume a posição comprada, isto é compra a opção, na outra está o investidor que assume a posição vendida, isto é vende ou lança a opção. O lançador de uma opção recebe dinheiro adiantado, mas tem obrigações potenciais mais tarde, seu lucro/perda é oposto ao do comprador da opção. É sempre útil caracterizar as posições em opções européias em termos de valor final ou retorno para o investidor no vencimento, pois estas relações serão utilizadas mais tarde quando formos analisar a formação de preços das opções.

O custo inicial da opção, portanto não é incluído no cálculo. Se K é o valor de exercício e S_T é preço da ação em T , o retorno em T de uma posição comprada numa opção de compra européia é

$$\max(S_T - K, 0) \quad (2.24)$$

Isso reflete o fato que a opção será exercida se $S_T > K$ e de que ela não será exercida se $S_T < K$. O retorno para o titular de uma posição vendida numa opção de compra européia é:

$$\min(K - S_T, 0) \quad (2.25)$$

O retorno para o titular de uma posição comprada de uma opção de venda européia é

$$\max(K - S_T, 0) \quad (2.26)$$

e o retorno de uma posição vendida numa opção de venda européia é

$$\min(S_T - K, 0) \quad (2.27)$$

Vamos agora analisar os fatores que afetam os preços das opções de ações. Estudaremos a relação entre preços de opções européias, preços de opções americanas e preço da ação objeto. Mostraremos que nunca é ideal exercer uma opção americana de compra de uma ação que não distribui dividendos antes do vencimento.

2.3 Fatores que Afetam o Preço das Opções

Há seis fatores que afetam o preço de uma opção de ação:

1. o preço atual da ação;
2. o preço de exercício;

3. o tempo para o vencimento;
4. a volatilidade do preço da ação;
5. a taxa de juro livre de risco;
6. os dividendos esperados durante a vida da opção.

Nesta seção, consideraremos o que acontece com os preços das opções quando um desses fatores muda e todos os outros permanecem fixos.

2.3.1 O preço da ação e o preço de exercício

O retorno de uma *opção de compra* será a quantia pela qual o preço da ação exceder o preço de exercício, por exemplo, se o preço de exercício é K e no dia em que a opção é exercida o valor da ação for S_T , então teremos

- Se $S_T > K$ o investidor exercerá a opção de compra. Sendo assim comprará a ação por K e poderá vender na mesma hora por S_T tendo um lucro de $S_T - K$.
- Se $S_T < K$ o investidor não exercerá a opção de compra pois poderá comprar a ação no mercado por um preço menor que K .

Portanto, uma opção de compra têm mais valor quando aumenta o preço da ação e menos valor quando diminui o preço da ação ou podemos pensar de uma outra maneira, uma opção de compra têm mais valor quando menor o preço de exercício e menos valor quando maior o preço de exercício.

Para uma *opção de venda*, temos exatamente o inverso. O retorno de uma *opção de venda* será a quantia pela qual o preço de exercício exceder o preço da ação. Sendo assim, quanto menor o preço da ação maior o valor da opção de venda e quanto maior o preço da ação menor o valor da opção de venda. Em termos de preço de exercício temos que quanto maior o preço de exercício maior o valor da opção de venda e quanto menor o preço de exercício menor o valor da opção de venda..

2.3.2 O tempo para o vencimento

Consideraremos o efeito da data de vencimento. Tanto as opções de compra como as opções de venda *americanas* tornam-se mais valiosas conforme aumenta o prazo para o venci-

mento. Para exemplificar isto, considere duas opções que difiram entre si apenas no que se refere à data de vencimento. O titular de uma opção de longa duração tem mais oportunidades de exercício que o titular de uma opção de curta duração. A opção de longa duração, por conseguinte, deve sempre valer, pelo menos, tanto quanto a opção de curta duração.

As opções de compra e venda *européias* não se tornam necessariamente mais valiosas à medida que aumenta o tempo para o vencimento. Isso porque o titular de uma opção européia de longa duração assim como o titular de uma européia de curta duração só pode exercê-la no vencimento. Assim consideremos duas opções européias de compra de uma ação, uma com a data de vencimento em um mês e a outra em dois meses. Suponhamos que um dividendo muito grande seja esperado em seis semanas e que ele cause uma queda no preço da ação. É provável que isso faça com que a opção de curta duração tenha um valor maior do que a opção de longa duração.

2.3.3 A volatilidade

A volatilidade σ do preço de uma ação é a medida da incerteza quanto às oscilações futuras em seu preço. Quanto maior a volatilidade, maior a possibilidade de a ação ter um desempenho tanto muito bom quanto muito ruim. Para o detentor de uma ação, as duas coisas tendem a se compensar. O mesmo não é válido para o titular de uma opção de compra ou de uma opção de venda. O primeiro beneficia-se com elevações de preço, mas terá risco limitado se o preço cair, já que o máximo que poderá perder será o preço da opção. O segundo beneficia-se com quedas de preço, mas terá risco limitado se ele subir. Assim, os valores das opções de compra e de venda crescem à medida que aumenta a volatilidade

Estabelecemos que uma maior volatilidade no preço das ações implica um preço mais alto tanto para as opções de compra como para as opções de venda. Em outras palavras, um aumento na volatilidade das ações, mantendo fixos o preço atual e a taxa de retorno esperada das ações, causará um aumento no pagamento esperado das opções de venda e nas opções de compra das ações. Consequentemente, para manter o equilíbrio, o preço corrente das opções de compra e das opções de venda aumentará com a volatilidade das ações.

Para ver como isso acontece, considere o exemplo em que o preço das ações pode assumir apenas um dos dois valores dentro de um ano, R\$120,00 ou R\$80,00, cada qual com uma probabilidade de 0,5. Portanto, o valor esperado do preço ao final de um ano é

$$0,5 \times 120,00 + 0,5 \times 80,00 = R\$100,00 \quad (2.28)$$

Considere agora uma opção de compra de ações com preço de exercício $k = R\$100,00$ que vence em um ano. No vencimento, a opção de compra terá o seguinte valor H_T

$$\begin{aligned} \text{Se } S_T = R\$120,00 \text{ então } H_T &= S_T - K = R\$20,00 \\ \text{Se } S_T = R\$80,00 \text{ então } H_T &= 0 \end{aligned} \quad (2.29)$$

Portanto, o valor esperado da opção de compra é

$$0,5 \times 20,00 + 0,5 \times 0 = R\$10,00 \quad (2.30)$$

Suponha que as ações tornem-se mais voláteis sem qualquer alteração em seu preço esperado de fim de ano. Suponha, por exemplo, que os dois valores possíveis para o preço das ações ao final do ano sejam agora $R\$180,00$ ou $R\$20,00$, cada qual com a probabilidade $0,5$. O valor esperado do preço das ações ao final do ano ainda é $R\$100,00$, pois

$$0,5 \times 180,00 + 0,5 \times 20,00 = R\$100,00 \quad (2.31)$$

mas sua volatilidade é muito maior. O valor da opção de compra, entretanto, é agora

$$\begin{aligned} \text{Se } S_T = R\$180,00 \text{ então } H_T &= R\$80,00 \\ \text{Se } S_T = R\$80,00 \text{ então } H_T &= 0 \end{aligned} \quad (2.32)$$

Portanto, o valor esperado da opção de compra é

$$0,5 \times 80,00 + 0,5 \times 0 = R\$40,00 \quad (2.33)$$

um aumento de 400%. Com esse grande aumento no valor esperado da opção no fim de um ano, o preço da opção de compra aumentará seu valor atual para manter o equilíbrio, pois o seu valor atual é diretamente proporcional ao seu valor esperado futuro, mais adiante será dado mais detalhes sobre a formulação do preço da opção.

Assim, verificamos a partir desse exemplo que um aumento na volatilidade (mantendo constante o preço atual das ações) causa um aumento no valor esperado das opções de compra. De um modo geral, para manter o equilíbrio, um aumento na volatilidade provoca um aumento no preço corrente das opções.

2.3.4 A taxa de juro livre de risco

A taxa de juro livre de risco afeta o preço de uma opção de maneira menos clara. A expansão das taxas de juro na economia, tende a aumentar a taxa de crescimento esperada para o preço da ação. Todavia, diminui o valor atual de quaisquer fluxos de caixa a serem recebidos pelo titular da opção no futuro. Esses dois efeitos tendem a diminuir o valor de uma opção de

venda. Logo, os preços de uma opção de venda recuam à medida que se eleva a taxa de juro livre de risco. No caso das opções de compra, o primeiro efeito tende a aumentar seu preço e o segundo, a diminuí-lo. O primeiro efeito prevalece sobre o segundo, isto é, o preço das opções de compra sempre aumentam com a taxa de juro livre de risco. Essas teorias não pressupõem mudanças em outras variáveis. Na prática, quando crescem as taxas de juro, os preços da ação tendem a cair, o que *eleva* o preço das opções de *venda* e *diminui* os preços das opções de *compra*.

2.3.5 Dividendos

Os dividendos têm o efeito de reduzir o preço da ação na data ex-dividendo, o que é bom para o valor das opções de venda e ruim para o valor das opções de compra. Os valores das opções de compra, portanto, relacionam-se negativamente com o tamanho de quaisquer dividendos antecipados e os valores das opções de venda, positivamente.

2.4 Limites Máximos e Mínimos para os Preços de Opções

Nesta seção, derivaremos os limites máximos e mínimos dos preços das opções, que não dependem de nenhum pressuposto específico com a exceção de que a taxa de juros deve ser positiva $r > 0$. Se o preço da opção estiver acima do limite máximo ou abaixo do limite mínimo, haverá oportunidades lucrativas para os arbitadores.

Aqui damos algumas notações que serão utilizadas:

- S_t preço da ação;
- K preço de exercício ou vencimento da opção;
- T tempo para o vencimento da opção;
- r taxa de juro livre de risco;
- β_t Fator de desconto;
- σ volatilidade do preço da ação;
- C_t valor de uma opção de compra americana de uma ação;
- P_t valor de uma opção de venda americana de uma ação;
- c_t valor de uma opção de compra européia de uma ação;
- p_t valor de uma opção de venda européia de uma ação.

Um outro ponto que queremos ressaltar é sobre o regime de capitalização. O regime pode ser discreto ou contínuo. Se for discreto então o fator de desconto de um título em T será dado por

$$\beta_{T-t} = (1 + r_d)^{-(T-t)} \quad (2.34)$$

Já se o regime de capitalização for contínua então

$$\beta_{T-t} = e^{-r_c(T-t)} \quad (2.35)$$

A relação entre a taxa de juros no regime da capitalização contínua r_c e a taxa de juros no regime de capitalização discreta r_d é dada por

$$1 + r_d = e^r \quad (2.36)$$

Serão dados alguns exemplos neste capítulo usando-se o regime de capitalização contínua.

2.4.1 Os Limites Máximos para Opções de Compra e de Venda

Uma opção de *compra* americana ou europeia dá a seu titular o direito de comprar uma ação por certo preço. Aconteça o que acontecer, a opção nunca poderá valer mais do que a ação. Logo, o preço da ação é um limite máximo para o preço da opção de compra:

$$c_t \leq S_t \text{ e } C_t \leq S_t, \text{ para todo } t \in [0, T] \quad (2.37)$$

Se isso não fosse verdade, um arbitrador poderia facilmente realizar um lucro sem risco, vendendo uma opção de compra adquirindo a ação.

Uma opção de *venda* americana ou europeia proporciona a seu titular o direito de vender uma ação por K . não importa o quanto cairá o preço da ação a opção nunca poderá valer mais que K . Logo:

$$p_t \leq K \text{ e } P_t \leq K, \text{ para todo } t \in [0, T] \quad (2.38)$$

Para as opções europeias, sabemos que, no instante T , a opção valerá menos que K . Assim, ela deve valer menos que o valor atual de K :

$$p_t \leq K\beta_{T-t}, \text{ para todo } t \in [0, T] \quad (2.39)$$

Se isso não fosse verdade, um arbitrador poderia realizar um lucro sem riscos, lançando uma opção *venda* e investindo o resultado da venda à taxa de juro livre de risco.

2.4.2 O limite mínimo para opções de compra de ações sem dividendos

Um limite mínimo para o preço de uma opção europeia de compra c_t de uma ação que não paga dividendo é:

$$\begin{aligned} c_t &\geq S_t - K\beta_{T-t} \text{ ou} \\ c_T &\geq S_T - K \end{aligned} \quad (2.40)$$

Vamos dar um exemplo de uma oportunidade de arbitragem quando $c_t \leq S_t - K\beta_{T-t}$.

Consideremos que um investidor acaba de obter as seguintes cotações para uma opção europeia de compra de uma ação que não distribui dividendos, com valor de exercício K de R\$18,00 e data de vencimento em um ano:

$$\begin{aligned} \text{Preço atual da ação } &R\$20,00 \\ \text{Preço da opção de compra } &R\$3,00 \\ \text{A taxa de juro livre de risco } &10\% \text{ a.a} \end{aligned} \quad (2.41)$$

Portanto, o investidor vende a ação (estamos presumindo que ele a possua neste momento), depois compra a opção de compra da mesma ação e investe o restante à taxa de juro livre de risco que neste caso é 10%.

Essa estratégia fornece um fluxo de caixa positivo imediato de

$$R\$20,00 - R\$3,00 = R\$17,00 \quad (2.42)$$

Os R\$17,00 são investidos a 10% ao ano e aumentam para

$$17 \times e^{0,1 \times 1} = 18,79 \quad (2.43)$$

no final de um ano. Nessa época $t = T$, a opção vence. Se $S_T > K$ o investidor exercerá a opção e encerrará a posição vendida para um lucro de:

$$R\$18,79 - R\$18,00 = R\$0,79 \quad (2.44)$$

Se $S_T < K$ a ação será comprada no mercado a vista e a posição vendida será encerrada. O investidor, então, realizará lucro de:

$$18,79 - S_T \quad (2.45)$$

Visto que $S_T < 18$, isso significa pelo menos R\$0,79 de lucro.

Portanto calculando o mínimo teórico para esta opção de compra, temos

$$S_t - K\beta_{T-t} = 20 - 18 \times e^{-0,1 \times 1} = R\$3,71 \quad (2.46)$$

desta maneira agora teríamos

$$R\$20,00 - R\$3,71 = R\$16,29 \quad (2.47)$$

Investindo agora estes R\$16,29 à taxa de 10% a.a teremos

$$16,29 \times e^{0,1 \times 1} = R\$18,00 \quad (2.48)$$

Portanto a opção de compra com valor igual ou superior a R\$3,71 a possibilidade de arbitragem foi eliminada, pois agora só teremos lucro se $S_T < 18$. Note que se $S_T = 18$ então o fluxo de caixa será zero ou seja não teremos nem lucro nem prejuízo.

Formalizando o que vimos neste exemplo, considere as seguintes carteiras:

Carteira A: uma opção de compra americana mais uma quantia em dinheiro igual a $K\beta_{T-t}$;

Carteira B: uma ação.

Na carteira A, se o dinheiro for investido à taxa de juro livre de risco, ele crescerá para K no instante T . Se $S_T > K$, a opção de compra será exercida no instante T e a carteira A valerá S_T . Se $S_T < K$, a opção de compra vencerá sem valor e a carteira valerá K . Logo, no instante T , a carteira A valerá:

$$\max(S_T, K)$$

A carteira B vale S_T no instante T . Logo, a carteira A sempre vale tanto quanto a carteira B no instante T (às vezes, até mais que ela). Assim, na ausência de oportunidades de arbitragem, isso também deve ser verdade hoje. Assim:

$$c_t + K\beta_{T-t} > S_t \Rightarrow c_t > S_t - K\beta_{T-t} \quad (2.49)$$

Visto que o pior que pode ocorrer a uma opção de compra é ter valor zero no vencimento, seu valor deve ser sempre positivo, portanto:

$$c_t > \max(S_t - K\beta_{T-t}, 0)$$

ou como condição final

$$c_T > \max(S_T - K, 0) \quad (2.50)$$

Esta condição será utilizada mais adiante quando formos deduzir a fórmula para precificação da opção de compra.

2.4.3 O limite mínimo para opções européias de venda de ações sem dividendos

Para uma opção européia de venda de uma ação que não paga dividendo, ocorre exatamente o oposto do que acontece com a opção de compra. sendo assim considere as seguintes carteiras.

Carteira C: uma opção de venda européia somada a uma ação;

Carteira D: uma quantia de dinheiro igual a $K\beta_{T-t}$.

Se $S_T < K$, a opção da carteira C será exercida no instante T e a carteira passará a valer K . Se $S_T > K$, a opção de venda vencerá sem valor e a carteira valerá S_T no instante T . Logo, a carteira C valerá:

$$\max(S_T, K) \quad (2.51)$$

no instante T . Se o dinheiro for investido à taxa de juro livre de risco, a carteira D valerá K no instante T . Logo, a carteira C sempre vale tanto quanto a carteira D no instante T (às vezes, até mais que ela). Assim, na ausência de oportunidades de arbitragem, a carteira C deverá valer hoje mais que a D. Logo:

$$p_t + S_t > K\beta_{T-t} \Rightarrow p_t > K\beta_{T-t} - S_t \quad (2.52)$$

Visto que o pior que pode acontecer a uma opção de venda é ter valor zero no vencimento, seu valor deve ser sempre positivo. Isso significa que:

$$p_t > \max(K\beta_{T-t} - S_t, 0) \quad (2.53)$$

ou como condição final

$$p_T > \max(K - S_T, 0) \quad (2.54)$$

2.5 O Exercício Antecipado de Opções de Compra Americana de Ações sem Dividendos

Vamos agora mostrar que nunca é ideal exercer uma opção americana de compra de uma ação sem dividendos antecipadamente. Considere as seguintes carteiras:

Carteira A: uma opção de compra americana mais uma quantia em dinheiro igual a $K\beta_{T-t}$;

Carteira B: uma ação.

O valor do dinheiro da carteira A no vencimento da opção será K . Em algum instante t anterior, ele equivale a $Ke^{-r(T-t)}$. Se a opção de compra for exercida no instante t , o valor da carteira A será:

$$S_t - K + K\beta_{T-t} \quad (2.55)$$

para isto basta o investidor exercer a opção comprando a ação por K e vendê-la no mesmo momento por S_t . Assim seu resultado será $S_t - K$.

Note portanto que o valor da carteira A será sempre menor que S_t enquanto $t < T$, desde que $r > 0$. A carteira A, por conseguinte, valerá sempre menos que B, se a opção de compra for exercida antes do vencimento. Se a opção de compra for mantida até o vencimento, o valor da carteira B no instante $t = T$ será:

$$\max(S_t, K) \quad (2.56)$$

O valor da carteira B é igual a S_t . Haverá sempre a possibilidade de $S_t < K$, o que significa que a carteira A sempre valerá tanto quanto a carteira B e às vezes, até mais que ela quando $S_t < K$ e $t = T$. Com isso, mostramos que

$$\begin{aligned} \text{se } t < T \text{ então } \text{carteira A} &\leq \text{carteira B quando a opção é exercida antecipadamente} \\ \text{se } t = T \text{ então } \text{carteira A} &\geq \text{carteira B} \end{aligned} \quad (2.57)$$

Assim, uma opção de compra de uma ação sem dividendos não deve ser nunca exercida antes da data de vencimento. Desse modo, uma opção americana de compra de uma ação sem dividendos vale o mesmo que uma opção européia correspondente da mesma ação

$$(C_t)_{t \in [0, T]} = (c_t)_{t \in [0, T]} \quad (2.58)$$

2.5.1 O Exercício Antecipado de Opções de Venda Americana de Ações sem Dividendos

No caso das opções de venda ocorre o contrário do que mostramos na seção anterior. Pode ser ideal exercer antecipadamente uma opção de venda americana de ações sem dividendos. De

fato, a qualquer momento durante sua vida, uma opção de venda deverá ser exercida se estiver suficientemente dentro do preço. Considere as seguintes carteiras:

Carteira A: uma opção de venda americana mais uma ação;

Carteira B: uma quantia em dinheiro igual a $K\beta_{T-t}$.

Se a opção for exercida no instante $t < T$, a carteira A passará a valer K , enquanto a carteira B valerá $K\beta_{T-t}$. Portanto, a carteira A vale mais que a B. Se a opção for mantida até o vencimento, a carteira A passará a valer:

$$\max(K, S_T) \quad (2.59)$$

enquanto a carteira B valerá K . Logo, a carteira A vale tanto quanto a carteira B, e às vezes mais que ela. Com isso, temos que

$$\begin{aligned} \text{se } t < T & \text{ então carteira A} > \text{ carteira B quando a opção é exercida antecipadamente} \\ \text{se } t = T & \text{ então carteira A} \geq \text{ carteira B} \end{aligned} \quad (2.60)$$

Notemos a diferença entre essa situação e a da seção anterior. Aqui, não podemos argumentar que o exercício antecipado seja indesejável, visto que a carteira A parece mais atrativa que a carteira B, independentemente da decisão de exercício antecipado.

Como uma opção de compra, uma opção de venda fornece um seguro. Quando mantida em conjunto com a ação, ela protege seu titular contra uma queda no preço da ação abaixo de certo limite. Entretanto, ela difere de uma opção de compra pelo fato de ser ideal para o investidor adiantar esse seguro e exercer antecipadamente, para realizar o preço de exercício de imediato. Em geral, o exercício antecipado de uma opção de venda torna-se mais interessante à medida que S_t diminui, r aumenta ou σ diminui.

2.6 A Paridade entre as Opções de Compra e Venda Européias

Esta relação decorre de um princípio fundamental de avaliação de ativos em finanças este princípio é conhecido como *Lei do Preço Único*.

Princípio 2.1 (Lei do Preço Único) *Em um mercado competitivo, se dois ativos são equivalentes, eles tendem a ter o mesmo preço de mercado.*

Assim considere agora as seguintes carteiras:

Carteira A: uma opção de compra européia mais uma quantia em dinheiro igual a $K\beta_{T-t}$;

Carteira B: uma opção de venda européia mais uma ação.

Ambas valem no vencimento das opções

$$\max(S_T, K) \quad (2.61)$$

Como as opções são do tipo europeu, elas não podem ser exercidas antes da data de vencimento, portanto estas carteiras são equivalentes. Assim, estas carteiras devem, por conseguinte, ter valores idênticos hoje. Isso significa que

$$c_t + K\beta_{T-t} = p_t + S_t \quad (2.62)$$

Essa relação é conhecida como *paridade entre opções de compra e venda*. Ela revela que o valor de uma opção de compra européia, com certos preço de exercício e data de vencimento, pode ser deduzido do valor de uma opção de venda européia com os mesmos preço de exercício e data de vencimento, e vice-versa.

Se a equação (2.62) não estiver sendo obedecida, haverá oportunidades de arbitragem. Digamos que o preço da ação seja de R\$31,00, o preço de exercício, de R\$30,00, a taxa de juro livre de risco, de 10% ao ano, o preço de uma opção de compra européia para três meses, de R\$3,00 e o preço de uma opção de venda européia para três meses, de R\$2,25. Nesse caso teremos

$$\begin{aligned} \text{Carteira A} &: c_t + K\beta_{T-t} = 3 + 30 \times e^{-0,1 \times \frac{3}{12}} = R\$32,26 \\ \text{Carteira B} &: p_t + S_t = 2,25 + 31,00 = R\$33,25 \end{aligned} \quad (2.63)$$

A carteira B está superavaliada relativamente à carteira A. Note que independentemente da carteira que o investido tiver ao final dos três meses qualquer uma das carteira valerá o $\max(S_T, K)$. Deste modo, podemos pensar em duas alternativas:

1. Se um investidor possuir a carteira B a estratégia de arbitragem correta para é vender os títulos da carteira B e comprar os títulos da carteira A. A estratégia gera um fluxo de caixa positivo de

$$-3,00 + 2,25 + 31,00 = R\$30,25 \quad (2.64)$$

que se investido à taxa de juro livre de risco, o montante crescerá para

$$30,25 \times e^{0,1 \times \frac{3}{12}} = R\$31,02 \quad (2.65)$$

em três meses. Se $S_T > K$, a opção de compra será exercida; se $S_T < K$, será exercida a opção de venda. Em qualquer um dos casos, o investidor acaba adquirindo uma ação por R\$30,00, que pode ser usada para encerrar a posição vendida. O lucro líquido, portanto, é de

$$R\$31,02 - R\$30,00 = R\$1,02 \quad (2.66)$$

2. Se um investidor possuir os R\$33,25 que ele poderia utilizar para comprar tanto a carteira A com a B. A estratégia correta seria comprar a carteira A. Portanto com esta estratégia a carteira A seria adquirida por R\$32,26 restando ainda

$$R\$33,25 - R\$32,26 = R\$0,99 \quad (2.67)$$

ao final dos três meses o investidor terá a carteira A que valerá $\max(S_T, K)$ mais R\$0,99 de lucro que se investido à taxa de juro livre de risco durante os três meses, dará

$$0,99 \times e^{0,1 \times \frac{3}{12}} = R\$1,02 \quad (2.68)$$

ou seja o mesmo resultado anterior, pois a estratégia utilizada foi a mesma nos dois casos.

2.7 A Relação entre os Preços de Opções de Compra e de Venda Americanas

A paridade entre opções de venda e compra é consistente apenas para opções europeias. Contudo, é possível deduzir algumas relações entre os preços de opções americanas. Como $(P_t)_{t \in [0, T]} > (p_t)_{t \in [0, T]}$, da equação (2.62), temos

$$P_t > c_t + K\beta_{T-t} - S_t \quad (2.69)$$

e como $(c_t)_{t \in [0, T]} = (C_t)_{t \in [0, T]}$, então

$$P_t > C_t + K\beta_{T-t} - S_t \quad (2.70)$$

ou

$$C_t - P_t < S_t - K\beta_{T-t} \quad (2.71)$$

Consideremos agora as seguintes carteiras:

Carteira A: uma opção de compra americana mais uma quantia em dinheiro igual a K ;

Carteira B: uma opção de venda americana mais uma ação.

Ambas as opções têm os mesmos preço de exercício e data de vencimento. O dinheiro da carteira A é investido à taxa de juro livre de risco. Se a opção de venda não for exercida antecipadamente, a carteira B valera no instante T

$$\max(S_T, K) \quad (2.72)$$

A carteira A valerá no mesmo instante T

$$[\max(S_T, K) - K] + K\beta_{T-t} \quad (2.73)$$

Assim, a carteira A vale mais que a carteira B. Suponhamos que a opção de venda da carteira B seja exercida antecipadamente, no instante t , o que significa que a carteira B valerá K no instante t . Porém, mesmo que a opção de compra não tenha valor, a carteira A valerá $K\beta_{T-t}$ no instante t . Então, a carteira A vale mais que a carteira B em qualquer circunstância. Assim

$$C_t + K > P_t + S_t \quad (2.74)$$

ou

$$C_t - P_t > S_t - K \quad (2.75)$$

Combinando isso com a equação (2.71), obtemos

$$S_t - K < C_t - P_t < S_t - K\beta_{T-t} \quad (2.76)$$

Nota: Os resultados produzidos até agora partiram do pressuposto de que o objeto das opções são ações sem dividendos.

Capítulo 3

Precificação das Opções a Tempo Discreto

Antes de passarmos aos modelos a tempo contínuo é interessante estudarmos o problema da precificação dos derivativos na estrutura a tempo discreto, pois podemos desta maneira compreender melhor o problema e a sua solução.

Esta abordagem a tempo discreto pode ser encontrada em [Bingham & Kiesel, 98], [Cox et al, 79], [Elliott & Kopp, 99], [Hull, 97], [Mel'nikov, 99], [Bodie & Merton, 99] e [Hull, 96]. As quatro primeiras referências analisam tanto modelos de mercado numa estrutura discreta como numa estrutura contínua, as duas últimas referências são textos introdutórios, sendo que o último é um texto mais geral sobre finanças tendo apenas três capítulos voltados para o problema do apreçamento de derivativos.

A discretização dos modelos a serem estudados além de oferecer uma série de simplificações técnicas nos oferece a oportunidade de poder entender com grande clareza os aspectos tanto econômicos quanto a sua formulação matemática. Este é o motivo principal pelo qual optamos tratar os modelos numa base discreta para depois fazermos a análise a tempo contínuo.

Aqui vamos seguir o enfoque dado por [Bingham & Kiesel, 98] e [Elliott & Kopp, 99], pois eles exploram de uma maneira mais efetiva a teoria dos martingales na sua análise o que neste estudo é desejável devido a íntima relação da teoria da precificação de opções com esta teoria. Isto ficará evidenciado tão logo começemos a desenvolver os primeiros resultados.

A análise dos modelos numa estrutura discreta além de oferecer uma série de simplificações técnicas nos oferece a oportunidade de poder entender com maior clareza os aspectos tanto econômicos quanto a sua formulação matemática. Este é um dos motivos pelo qual optamos tratar os modelos numa base discreta para depois fazermos a análise a tempo contínuo, o outro

motivo é para que possamos comparar os valores determinados pelos modelos a tempo discreto com os valores dos modelos a tempo contínuo.

3.1 Considerações Iniciais

Vamos trabalhar com a base discreta $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ onde o espaço amostral Ω tem inicialmente um número finito $|\Omega|$ de pontos ω , posteriormente o espaço amostral será não enumerável quando utilizarmos as versões discretas dos modelos contínuos analisados aqui neste trabalho. Denotaremos um horizonte de tempo $T \in \mathbb{N}^+$ que será a data final para todas as atividades econômicas consideradas. Por exemplo, para um modelo de precificação de opções o horizonte de tempo tipicamente corresponde à data de vencimento da opção.

Consideremos um mercado financeiro com $d + 1$ ativos, $d \in \mathbb{Z}^+$, e admita um processo estocástico $d + 1$ -dimensional $S_t = \{S_t^i : t \in [0, T], 0 \leq i \leq d\}$ para representar a evolução temporal dos processos dos preços dos ativos. O ativo rotulado por 0 é tomado como um título sem risco, isto é, não aleatório, com processo preço $(S_t^0)_{t \in [0, T]}$, enquanto que os ativos de risco que são ações (são variáveis aleatórias), rotuladas de 1 a d , e tem seu preço determinado pelo processo $(S_t^i)_{t \in [0, T]}$. O processo S_t é adaptado a filtragem $\mathbb{F}$, tal que para cada $\{i; 0 \leq i \leq d\}$, S_t^i é $\mathcal{F}_t$ -mensurável, isto é os preços dos ativos de todos os tempos até t são conhecidos no tempo t . Frequentemente tomaremos a filtragem $\mathbb{F}$ como aquela gerada pelo processo preço S_t . Então $\mathcal{F}_t = \sigma(S_u : u \leq t)$, é a menor σ -álgebra gerada por $\mathbb{R}^{d+1}$, é tal que os processos preços $S_u = (S_u^0, S_u^1, \dots, S_u^d)$ são $\mathcal{F}_t$ -mensuráveis. Em outras palavras, no tempo t , os investidores conhecem os valores dos vetores de preço $(S_u : u \leq t)$, mas eles não têm informação sobre os valores posteriores de S . Será necessário que pelo menos um ativo tenha um processo preço estritamente positivo, isto é, para funcionar como um benchmarking (nível de referência), conhecido como numerário no modelo. Esta regra será atribuída ao processo preço $(S_t^0)_{t \in [0, T]}$.

Definição 3.1 (Numerário) *Um numerário é um processo preço $(X_t)_{t \in [0, T]}$, que é estritamente positivo para todo $t \in [0, T]$, $T \in \mathbb{Z}^+$.*

Este ativo sem risco conhecido como numerário é fornecido por uma instituição financeira, tipicamente um banco. Aqui vamos considerar para efeitos práticos que este banco seja um banco ideal.

Definição 3.2 (Banco Ideal) *Num banco ideal as taxas de depósito e empréstimo são iguais e não há taxas de serviços e transações. As taxas também independem do montante do principal.*

Se um banco ideal tem uma taxa de juros que independe do período pela qual é ela aplicada e a taxa de juros é composta de acordo com a regras usuais, ele é chamado um banco ideal constante.

Portanto, a partir destes pressupostos pode-se supor a existência de um ativo sem risco, o numérario, cujos processo preço $(S_t^0)_{t \in [0, T]}$ é dado por

$$S_t^0 = S_0^0 (1 + r)^t, t \in [0, T], T \in \mathbb{Z}^+ \quad (3.77)$$

onde r é a taxa de juros livre de risco. Além do mais, sem perda de generalidade, podemos assumir que $S_0^0 = 1$, assim

$$S_t^0 = (1 + r)^t \quad (3.78)$$

tal que o valor inicial do nosso ativo sem risco produza a unidade relativa para qual todas as outras quantidades sejam expressas. A partir de agora sempre adotaremos que $S_0^0 = 1$. Deste modo o fator de desconto é dado por

$$\beta_t = \frac{1}{S_t^0} = (1 + r)^{-t}, t \in [0, T], T \in \mathbb{Z}^+ \quad (3.79)$$

Assim o fator de desconto β_t é então a soma de dinheiro que precisamos investir em ativos no tempo 0 para obter 1 unidade no tempo t .

Definição 3.3 (Estratégia de Transação) Uma estratégia de transação (trading) φ é um processo estocástico $\mathbb{R}^{d+1}$ -dimensional $\varphi_t = (\varphi_t^0, \varphi_t^1, \varphi_t^2, \dots, \varphi_t^d)'$, $d \in \mathbb{Z}^+$ que é não antecipativo, assim φ_t^i é $\mathcal{F}_{t-1}$ -mensurável para $t \geq 1$.

Aqui φ_t^i denota a fração do ativo i contido na carteira no tempo t a ser determinado com base na informação disponível antes do tempo t , i.e., o investidor seleciona a carteira no tempo t depois de observar os preços S_{t-1} em $t - 1$. Todavia, a carteira φ_t deve ser estabelecida antes, e mantida até depois, da observação dos preços S_t em t . Os componentes φ_t^i podem assumir tanto valores positivos como valores negativos e assumimos que os ativos são perfeitamente divisíveis, isto é, são permitidas vendas de qualquer montante de um ativo financeiro. Isto nos leva ao conceito de mercado perfeito.

Definição 3.4 (Mercado Perfeito) Mercado perfeito é uma mercado sem custos de transação, no qual todos os acordos são cumpridos, com possibilidade de comprar e vender qualquer quantidade dos ativos existentes no mercado, as transações são contínuas e vendas a descoberto são permitidas.

Definição 3.5 (Valor da Carteira) O valor da carteira no tempo t é o produto escalar

$$V_t^\varphi \triangleq \varphi_t \cdot S_t = \sum_{i=0}^d \varphi_t^i S_t^i$$

com $t \in [0, T]$, $T \in \mathbb{Z}^+$, $S_t = (S_t^0, S_t^1, \dots, S_t^d)$, $\varphi_t = (\varphi_t^0, \varphi_t^1, \varphi_t^2, \dots, \varphi_t^d)'$ e $V_0^\varphi = \varphi_1 S_0$. O processo $(V_t^\varphi)_{t \in [0, T]}$ é chamado processo valor ou montante da estratégia de transação φ . O valor inicial V_0^φ é chamado de capital ou fortuna inicial.

Agora $\varphi_t S_{t-1}$ reflete o valor de mercado da carteira logo após ela ter sido estabelecida no tempo $t - 1$, enquanto que $\varphi_t S_t$ é o valor imediatamente depois que os preços no tempo t são observados, mas antes das mudanças serem feitas na carteira. Conseqüentemente

$$\varphi_t \cdot (S_t - S_{t-1}) = \varphi_t \cdot \Delta S_t \quad (3.80)$$

é a mudança no valor de mercado devido às mudanças no preço das ações que ocorreram entre $t - 1$ e t . Isto motiva a seguinte definição.

Definição 3.6 (Processo Ganho da Carteira) O processo ganho $(G_t^\varphi)_{t \in [1, T]}$ de uma estratégia de transação φ é dado por

$$G_t^\varphi \triangleq \sum_{\tau=1}^t \varphi_\tau \cdot \Delta S_\tau$$

onde $\varphi_\tau = \{\varphi_\tau^0, \varphi_\tau^1, \dots, \varphi_\tau^d\}$, $\Delta S_\tau = S_\tau - S_{\tau-1}$ e $t \in [1, T]$, $T \in \mathbb{Z}^+$.

Denotando $\tilde{S}_t = \{\beta_t S_t^0, \beta_t S_t^1, \dots, \beta_t S_t^d\}' = \{1, \tilde{S}_t^1, \dots, \tilde{S}_t^d\}'$, o vetor de preços descontados, então o processo montante descontado é dado por

$$\tilde{V}_t^\varphi = \beta_t \varphi_t \cdot S_t = \varphi_t \cdot (\beta_t S_t) = \varphi_t \cdot \tilde{S}_t \quad (3.81)$$

e o processo ganho descontado

$$\tilde{G}_t^\varphi = \sum_{\tau=1}^t \varphi_\tau \cdot \Delta \tilde{S}_\tau \quad (3.82)$$

Observe que o processo ganho descontado reflete o ganho somente da transação com os ativos de risco, pois

$$\varphi_\tau^0 (\tilde{S}_t^0 - \tilde{S}_{t-1}^0) = \varphi_\tau^0 (\beta_t S_t^0 - \beta_{t-1} S_{t-1}^0) = \varphi_\tau^0 (1 - 1) = 0 \quad (3.83)$$

Consideremos uma classe especial de estratégias de transação, as chamadas estratégias autofinanciáveis Φ .

Uma estratégia autofinanciável é uma dinâmica que requer que se ajuste o número de ações e a quantia de empréstimos a cada período, de acordo com os preços praticados na época. Esta

estratégia também exige que, após o investimento inicial, fundos adicionais não sejam aplicados pelo investidor. Daí o termo autofinanciável. A definição abaixo, mostra matematicamente o conceito da estratégia autofinanciável.

Definição 3.7 (Estratégia Autofinanciável) *Uma estratégia de transação φ é autofinanciável, $\varphi \in \Phi$ se*

$$\varphi_t \cdot S_t = \varphi_{t+1} \cdot S_t \Rightarrow \Delta \varphi_t \cdot S_t = 0$$

onde $\varphi_t = \{\varphi_t^0, \varphi_t^1, \dots, \varphi_t^d\}$, $S_t = \{S_t^0, S_t^1, \dots, S_t^d\}$ e $t \in [0, T-1]$, $T \in \mathbb{Z}^+$.

Isto significa que quando novos preços são cotados no tempo t , o investidor ajusta a carteira dele de φ_t para φ_{t+1} sem que seja feita qualquer retirada ou que novo capital seja investido. Matematicamente isto significa que para uma estratégia autofinanciável φ_t o vetor de mudanças na carteira $\Delta \varphi_t$ é ortogonal ao vetor de preço S_t , portanto o conjunto das estratégias autofinanciáveis Φ pode ser descrito mais precisamente da seguinte forma

$$\Phi = \{\varphi_t \in \mathbb{R}^d \mid \Delta \varphi_t \cdot S_t = 0, S_t \in \mathbb{R}^d, t \in [0, T], T \in \mathbb{Z}^+\} \quad (3.84)$$

Proposição 3.1 (Estratégia Autofinanciável) *Uma estratégia autofinanciável φ é se e somente se*

$$V_t^\varphi = V_0^\varphi + G_t^\varphi$$

para todo $t \in [0, T]$.

Prova: Para ver isto note que

$$\begin{aligned} V_0^\varphi + G_t^\varphi &= \varphi_1 \cdot S_0 + \sum_{\tau=1}^t \varphi_\tau \cdot \Delta S_\tau \\ &= \varphi_1 \cdot S_0 + \sum_{\tau=1}^t \varphi_\tau \cdot (S_\tau - S_{\tau-1}) \\ &= \varphi_1 \cdot S_0 + [\varphi_1 \cdot (S_1 - S_0) + \varphi_2 \cdot (S_2 - S_1)] + \dots \\ &\quad + \varphi_{t-1} \cdot (S_{t-1} - S_{t-2}) + \varphi_t \cdot (S_t - S_{t-1}) \\ &= \varphi_1 \cdot S_0 + [-\varphi_1 \cdot S_0 + (\varphi_1 - \varphi_2) \cdot S_1 + \dots + (\varphi_{t-1} - \varphi_t) \cdot S_{t-1} + \varphi_t \cdot S_t] \\ &= \varphi_t \cdot S_t + \sum_{\tau=1}^{t-1} (\varphi_\tau - \varphi_{\tau+1}) \cdot S_\tau \\ &= \varphi_t \cdot S_t \quad \text{pela definição (3.7)} \\ &= V_t^\varphi \blacksquare \end{aligned}$$

Isto significa que φ_t é autofinanciável se e somente se o valor de V_t^φ resulta unicamente como uma soma do investimento inicial V_0^φ e o processo ganho associado com a estratégia φ_t .

O próximo resultado, que é trivial em nossa estrutura de análise atual, mas requer um pequeno argumento no caso a tempo contínuo, mostra que renormalizar o preço dos ativos não tem essencialmente nenhum efeito econômico.

Proposição 3.2 (Invariância do Numerário) *Seja $(X_t)_{t \in [0, T]}$ um numerário. Uma estratégia de transação φ_t é autofinanciável com respeito a S_t se e somente se φ_t é autofinanciável com respeito a $(X_t)^{-1} S_t$.*

Prova: Desde que $(X_t)_{t \in [0, T]}$ é estritamente positivo por definição veja definição (3.1), para todo $t \in [0, T]$, temos a seguinte equivalência, que implica que

$$\begin{aligned} \varphi_t \cdot S_t &= \varphi_{t+1} \cdot S_t, \text{ para todo } t \in [0, T-1] \\ \Leftrightarrow \\ \varphi_t \cdot ((X_t)^{-1} S_t) &= \varphi_{t+1} \cdot ((X_t)^{-1} S_t), \text{ para todo } t \in [0, T-1] \blacksquare \end{aligned}$$

Deste modo a escolha de um novo benchmark não alterará a classe de estratégias de transação em consideração e deste modo não afetará o comportamento do mercado.

A partir desta proposição segue direto que:

Corolário 3.1 (Invariância do Numerário) *Uma estratégia de transação $(\varphi_t)_{t \in [0, T]}$ é autofinanciável com respeito a $(S_t)_{t \in [0, T]}$ se e somente se $(\varphi_t)_{t \in [0, T]}$ é autofinanciável com respeito ao processo preço descontado $(\tilde{S}_t)_{t \in [0, T]}$.*

Prova: Temos que $\beta_t = \frac{1}{S_t^0}$ o fator de desconto, assim da proposição anterior segue que

$$\begin{aligned} \varphi_t \cdot S_t &= \varphi_{t+1} \cdot S_t, \quad t \in [0, T-1] \\ \Leftrightarrow \\ \varphi_t \cdot (\beta_t S_t) &= \varphi_{t+1} \cdot (\beta_t S_t) \\ \varphi_t \cdot \tilde{S}_t &= \varphi_{t+1} \cdot \tilde{S}_t \quad t \in [0, T-1] \blacksquare \end{aligned}$$

Proposição 3.3 *Uma estratégia de transação φ pertence a Φ se e somente se*

$$\tilde{V}_t^\varphi = V_0^\varphi + \tilde{G}_t^\varphi$$

para todo $t \in [0, T]$.

Prova: A prova desta proposição é direta, basta usar o corolário (3.1) e a proposição (3.2) ■

Na formulação dos modelos de precificação foram utilizadas algumas suposições a cerca do mercado. A premissa principal é que no conjunto de estratégias autofinanciáveis não exista nenhuma que seja uma estratégia de arbitragem. Em termos matemáticos isto significa a existência de uma medida de neutralidade do risco esta medida é chamada de Medida Martingale Equivalente (MME). Se esta medida existe e é única dizemos que o *mercado é completo*. A determinação desta MME foi o princípio central no desenvolvimento deste trabalho.

Definição 3.8 (Estratégia de Arbitragem) *Seja $\tilde{\Phi} \subset \Phi$ um subconjunto do conjunto de estratégias autofinanciáveis. Uma estratégia $\varphi \in \tilde{\Phi}$ é chamada uma oportunidade de arbitragem ou estratégia de arbitragem com respeito a $\tilde{\Phi}$ se $P[V_0^\varphi = 0] = 1$ e o montante final de φ satisfaz*

$$P[V_T^\varphi \geq 0] = 1 \text{ e } P[V_T^\varphi > 0] \geq 0$$

Portanto uma oportunidade de arbitragem é uma estratégia autofinanciável com valor (investimento) inicial zero, que produz um montante final com probabilidade 1 e tem uma probabilidade positiva de se ter um montante final maior que zero. Observe que as oportunidades de arbitragem são sempre definidas com respeito a uma certa classe de estratégias de transação.

3.2 Medida Martingale Equivalente (MME)

Definição 3.9 (Medida Martingale Equivalente (MME)) *Uma medida de probabilidade $\mathbb{Q}$ sobre $(\Omega, \mathbb{F})$ equivalente a $\mathbb{P}$ é chamada uma medida martingale para o processo preço descontado $(\tilde{S}_t)_{t \in [0, T]}$ se o processo $(\tilde{S}_t)_{t \in [0, T]}$ é um martingale sobre $\mathbb{Q}$ com respeito a filtragem $\mathbb{F}$ e denotamos por $\mathbb{Q} \sim \mathbb{P}$. A classe dos MME será denotada por $P(\tilde{S})$.*

Proposição 3.4 *Seja o processo preço descontado $(\tilde{S}_t)_{t \in [0, T]}$, $\mathbb{Q} \in P(\tilde{S})$ uma MME para $(\tilde{S}_t)_{t \in [0, T]}$, e $\varphi \in \Phi$ alguma estratégia autofinanciável. Então o processo montante $(V_t^\varphi)_{t \in [0, T]}$ é um $\mathbb{Q}$ -martingale com respeito a filtragem $\mathbb{F}$.*

Prova: Da proposição (3.3) temos que

$$\tilde{V}_t^\varphi = V_0^\varphi + \tilde{G}_t^\varphi$$

assim

$$\begin{aligned}\tilde{V}_t^\varphi - \tilde{V}_{t-1}^\varphi &= \tilde{G}_t^\varphi - \tilde{G}_{t-1}^\varphi \\ &= \sum_{\tau=1}^t (\varphi_\tau \cdot \Delta \tilde{S}_\tau) - \sum_{\tau=1}^{t-1} (\varphi_\tau \cdot \Delta \tilde{S}_\tau) \\ \Delta \tilde{V}_t^\varphi &= (\varphi_t \cdot \Delta \tilde{S}_t)\end{aligned}$$

portanto pela definição (C.2), apêndice C, $(V_t^\varphi)_{t \in [0, T]}$ é a transformação martingale do $\mathbb{Q}$ -martingale $\tilde{S}_t$ por φ e portanto um martingale ■

Os próximos resultados mostram que a existência da MME implica em um mercado sem oportunidades de arbitragem e que um mercado sem oportunidades de arbitragem implica a existência da MME.

Proposição 3.5 *Se uma MME existe, isto é, se $P(\tilde{S}) \neq \emptyset$, então o mercado não tem oportunidades de arbitragem.*

Prova: *Se existe uma MME então para qualquer estratégia autofinanciável φ temos que pela proposição (3.3)*

$$\begin{aligned}\tilde{V}_t^\varphi &= V_0^\varphi + \tilde{G}_t^\varphi \\ &= V_0^\varphi + \sum_{i=1}^t (\varphi_i \cdot \Delta \tilde{S}_i)\end{aligned}$$

mas pela proposição (3.4) temos que para $(\tilde{S}_t)_{t \in [0, T]}$ um $\mathbb{Q}$ -martingale implica que $(\tilde{V}_t)_{t \in [0, T]}$ é um $\mathbb{Q}$ -martingale portanto aplicando o operador esperança $E_{\mathbb{Q}}(\cdot)$ em ambos os lados temos

$$E_{\mathbb{Q}}(\tilde{V}_T^\varphi | \mathcal{F}_0) = E_{\mathbb{Q}}(V_0^\varphi | \mathcal{F}_0) + \sum_{i=1}^T (\varphi_i \cdot E_{\mathbb{Q}}[(\tilde{S}_i - \tilde{S}_{i-1}) | \mathcal{F}_0])$$

$$E_{\mathbb{Q}}(\tilde{V}_T^\varphi | \mathcal{F}_0) = E_{\mathbb{Q}}(V_0^\varphi | \mathcal{F}_0) = V_0^\varphi$$

Se φ é uma estratégia de arbitragem então $V_0^\varphi = 0$ e portanto da equação acima temos que

$$E_{\mathbb{Q}}(\tilde{V}_T^\varphi | \mathcal{F}_0) = 0$$

mas pela definição (3.8)

$$E_{\mathbb{Q}}(\tilde{V}_T^\varphi | \mathcal{F}_0) > 0$$

portanto não é possível uma estratégia de arbitragem se existe a MME ■

Proposição 3.6 *Se o mercado não possui oportunidades de arbitragem, então a classe das MME é não vazia, isto é, $P(\tilde{S}) \neq \emptyset$.*

Prova: A prova deste resultado pode ser encontrada em [Bingham & Kiesel, 98].

Assim combinando estas duas últimas proposições chegamos a um resultado conhecido como teorema da não arbitragem.

Teorema 3.1 (Teorema da não Arbitragem) *Um mercado é livre de arbitragem, se e somente se existe uma MME $\mathbb{Q} \sim \mathbb{P}$ sobre o qual os processos preços descontado $(\tilde{S}_t)_{t \in [0, T]}$ são um $\mathbb{Q}$ -martingale.*

Uma outra restrição que temos com relação as estratégias autofinanciáveis é que ela não pode ser uma estratégia de arbitragem, estas estratégias são chamadas de admissíveis.

Definição 3.10 (Estratégia Admissível) *Uma estratégia de transação autofinanciável $\varphi \in \Phi$ é dita ser uma estratégia admissível se o processo montante ou valor da carteira $(V_t^\varphi)_{t \in [0, T]}$ é um $\mathbb{Q}$ -martingale e $V_t^\varphi \geq 0$ para $t \in [0, T]$. A classe das estratégias admissíveis será denotada por Φ_a .*

Um outro conceito é aquele que dá a condição para que o mercado seja um mercado completo.

Teorema 3.2 (Mercado Completo) *Um mercado sem oportunidade de arbitragem é completo se e somente se existe uma única medida martingale $\mathbb{Q}$ equivalente a $\mathbb{P}$ sobre o qual os processos preços descontado $(\tilde{S}_t)_{t \in [0, T]}$ são um $\mathbb{Q}$ -martingale.*

Combinando os teoremas (3.1) e (3.2) chegamos ao teorema fundamental da precificação de derivativos.

Teorema 3.3 (Teorema Fundamental da Precificação de Derivativos) *Em um mercado sem oportunidades de arbitragem e completo, existe uma única medida martingale equivalente $\mathbb{Q}$.*

3.3 Avaliando Opções pela Medida Martingale Equivalente

Retornemos a questão principal de todo este desenvolvimento que é a precificação das opções. Mostremos como utilizar a MME para determinar o preço justo de uma opção.

Definição 3.11 (Derivativo) *Um derivativo $(H_t)_{t \in [0, T]}$ com data de exercício (maturidade) T e valor de exercício K é um processo estocástico não negativo $\mathcal{F}_t$ -mensurável. Denotamos a*

classe de todos os derivativos por χ^+ .

Aqui estaremos interessados somente nas opções.

Definição 3.12 Dizemos que um derivativo $(H_t)_{t \in [0, T]}$ é “attainable” se existe uma estratégia de duplicação $\varphi \in \Phi_a$, tal que

$$V_T^\varphi = H_T$$

portanto a estratégia de duplicação φ gera o mesmo fluxo de caixa de H em T .

Isto significa que podemos encontrar uma estratégia φ_t , chamada de estratégia duplicante, onde φ_t é um processo não antecipativo tal que o processo valor da carteira

$$(V_t^\varphi)_{t \in [0, T]} = (\varphi_t \cdot S_t)_{t \in [0, T]} \quad (3.85)$$

satisfaz

$$(H_t)_{t \in [0, T]} = (V_t^\varphi)_{t \in [0, T]} \quad (3.86)$$

Em um mercado de seguros altamente eficiente esperamos que a lei do preço único veja princípio (2.1), seja válido a todo instante, caso contrário haverá oportunidades de arbitragem. Portanto a condição de não arbitragem implica que para um derivativo “attainable” seu preço no tempo t deve ser dado pelo valor de alguma estratégia de duplicação, assim dizemos que o derivativo é unicamente duplicado naquele caso. Esta é a idéia básica da Teoria da Precificação da Arbitragem (Arbitrage Pricing Theory-APT). Claramente a equivalência da condição de não arbitragem e a existência da MME conhecida como medida de probabilidade do risco neutro, implica na possibilidade de usar a medida do risco neutro para nossas proposta de precificação de opções.

Estas idéias foram inicialmente desenvolvidas para o caso em que temos um ativo de risco e um ativo sem risco por [Cox & Ross, 76] e em uma forma mais geral por [Harrison & Kreps, 79].

Portanto a partir destas considerações podemos agora colocar a seguinte proposição que nos responderá a pergunta inicialmente colocada.

Proposição 3.7 (Fórmula da Precificação do Risco Neutro) O processo preço da arbitragem de todo derivativo “attainable” $(H_t)_{t \in [0, T]}$ é dado pela Fórmula de Precificação do Risco Neutro

$$H_t = \beta_t^{-1} E_{\mathbb{Q}} [\beta_T H_T | \mathcal{F}_t]$$

para todo $t \in [0, T]$, $T \in \mathbb{Z}^+$.

Prova: Desde que assumamos que no mercado não existem oportunidades de arbitragem então existe a medida martingale equivalente $\mathbb{Q}$, veja teorema (3.1). Pela proposição (3.4) o processo $(\tilde{V}_t^\varphi)_{t \in [0, T]}$ de toda estratégia autofinanciável é um $\mathbb{Q}$ -martingale. Portanto, para todo derivativo “attainable” $(H_t)_{t \in [0, T]}$ com data de exercício T e estratégia replicante $\varphi \in \Phi_\alpha$, satisfaz para todo $t \in [0, T]$

$$\begin{aligned}
 \tilde{V}_T^\varphi &= \tilde{V}_t^\varphi + \tilde{G}_T^\varphi \\
 \tilde{V}_T^\varphi - \tilde{V}_t^\varphi &= \tilde{G}_T^\varphi \\
 E_{\mathbb{Q}} [\tilde{V}_T^\varphi - \tilde{V}_t^\varphi | \mathcal{F}_t] &= E_{\mathbb{Q}} [\tilde{G}_T^\varphi | \mathcal{F}_t] \\
 E_{\mathbb{Q}} [\tilde{V}_T^\varphi - \tilde{V}_t^\varphi | \mathcal{F}_t] &= 0 \text{ pela prop. (3.4)} \\
 E_{\mathbb{Q}} [\beta_t V_t^\varphi | \mathcal{F}_t] &= E_{\mathbb{Q}} [\beta_T V_T^\varphi | \mathcal{F}_t] \\
 \beta_t V_t^\varphi &= E_{\mathbb{Q}} [\beta_T V_T^\varphi | \mathcal{F}_t] \\
 V_t^\varphi &= \beta_t^{-1} E_{\mathbb{Q}} [\beta_T V_T^\varphi | \mathcal{F}_t] \\
 H_t^\varphi &= \beta_t^{-1} E_{\mathbb{Q}} [\beta_T H_T^\varphi | \mathcal{F}_t] \text{ pois } \varphi \text{ é uma estratégia duplicante} \blacksquare
 \end{aligned}$$

3.4 Premissas do Modelo

Na formulação do modelo de precificação de opções, será assumido que as seguintes condições são satisfeitas.

1. O mercado é perfeito, veja definição (3.4) ;
2. O mercado é eficiente;
3. O mercado é completo, veja teorema (3.2);
4. A ação não receberá dividendos durante a vida da opção;
5. A taxa da juros livre de risco r é constante

O **Mercado é eficiente** significa que o conhecimento das mudanças passadas do preço da ação não informa nada sobre as mudanças futuras. Assim os preços sempre refletem toda a informação relevante para os investidores e sendo assim os investidores irão mudar sua posições somente quando nova informação surgir. Ora por definição uma informação nova não pode ser predita caso contrário não seria uma nova informação. Deste modo o preço da ação reflete tudo o que é previsível enquanto a variação do preço da ação reflete o que é imprevisível. Portanto, a série de variações dos preços deve ser aleatória. Ora, a versão contínua do passeio aleatório

é o movimento browniano ou processo de Wiener. Assim, os modelos que serão utilizados na precificação das opções deverão ser funcionais do processo de Wiener.

As condições sobre dividendos e a taxa de juros são condições simplificadoras do modelo.

Algumas destas hipóteses vem sendo flexibilizadas por alguns pesquisadores. Por exemplo a taxa de juros r e o coeficiente de volatilidade σ sendo funções do tempo; uma outra consideração é a possibilidade de haver dividendos durante a vida da opção.

3.4.1 O Modelo Binomial

O modelo consiste de um ativo sem risco cujo preço no tempo t é S_t^0 , e um ativo de risco, a ação objeto da opção, cujo preço no tempo t é S_t^1 . Considerando que o processo preço $(S_t^1)_{t \in [0, T]}$ tem a seguinte dinâmica

$$S_{t+1}^1 = S_t^1 + S_t^1 \Delta b_{t+1} \quad (3.87)$$

onde $\Delta b_{t+1} \sim \text{bernoulli}(p, (1-p))$, assim

$$\Delta b_{t+1} = \begin{cases} u & \text{com probabilidade } p \\ d & \text{com probabilidade } 1-p \end{cases} \quad (3.88)$$

isto significa que o preço da ação poderá subir para um valor $S_{t+1}^1 = (1+u) S_t^1$ com uma certa probabilidade p ou poderá descer para um valor $S_{t+1}^1 = (1+d) S_t^1$ com probabilidade $1-p$, ou seja

$$\begin{aligned} \mathbb{P}\{S_{t+1}^1 = (1+u) S_t^1\} &= p \\ \mathbb{P}\{S_{t+1}^1 = (1+d) S_t^1\} &= 1-p \end{aligned} \quad (3.89)$$

onde $0 < d < u < 1$ e $0 < p < 1$. Seja $r > 0$ a taxa de juros que é paga pelo ativo sem risco S_t^0 , então um montante inicial S_0^0 aplicado neste ativo capitalizado em períodos discretos de tempo, tem a seguinte dinâmica

$$S_t^0 = S_0^0 (1+r)^t \quad (3.90)$$

como foi visto na seção (3.1), iremos sempre considerar que $S_0^0 = 1$, assim

$$S_t^0 = (1+r)^t \quad (3.91)$$

e portanto a taxa de desconto, veja seção (3.1), é dada por

$$\beta_t = (1+r)^{-t} \quad (3.92)$$

portanto o processo preço descontado é

$$\begin{aligned} S_{t+1}^1 (1+r)^{-(t+1)} &= S_t^1 (1+r)^{-t} + S_t^1 (1+r)^{-t} \Delta b_{t+1} \\ S_{t+1}^1 \beta_{t+1} &= S_t^1 \beta_t + S_t^1 \beta_t \Delta b_{t+1} \\ \tilde{S}_{t+1}^1 &= \tilde{S}_t^1 + \tilde{S}_t^1 \Delta b_{t+1} \end{aligned} \quad (3.93)$$

A dinâmica para $T - t$ passos a frente é a seguinte

$$\frac{\tilde{S}_T^1 - \tilde{S}_t^1}{\tilde{S}_t^1} = \prod_{j=t}^{T-1} \Delta b_{j+1} \sim \text{binomial}(T-t, p) \quad (3.94)$$

3.4.2 Avaliando uma Opção pelo Modelo Binomial

Considere uma opção com preço de exercício K e data de exercício T . H será o derivativo a ser avaliado no tempo T . Considere o valor corrente de H em $T - 1$, isto é, um período antes da maturidade. Podemos considerar este como o valor inicial do derivativo em um modelo de período único e supondo que a condição de não arbitragem é satisfeita e que o mercado é completo então existe uma estratégia autofinanciável φ que duplica o fluxo de caixa de H sobre o conjunto de tempo $\{T - 1, T\}$ e uma medida martingale equivalente $\mathbb{Q}$ que é única. Deste modo, temos que

$$E_{\mathbb{Q}} \left(\frac{\Delta \tilde{S}_T^1}{\tilde{S}_{T-1}^1} \middle| \mathcal{F}_{T-1} \right) = 0 \quad (3.95)$$

assim

$$\begin{aligned} E_{\mathbb{Q}} \left(\Delta \tilde{S}_T^1 \middle| \mathcal{F}_{T-1} \right) &= 0 \\ E_{\mathbb{Q}} \left(\tilde{S}_T^1 \middle| \mathcal{F}_{T-1} \right) &= \tilde{S}_{T-1}^1 \\ E_{\mathbb{Q}} \left(S_T^1 \beta_T \middle| \mathcal{F}_{T-1} \right) &= \beta_{T-1} S_{T-1}^1 \\ E_{\mathbb{Q}} \left(S_T^1 \middle| \mathcal{F}_{T-1} \right) &= \beta_{T-1}^{-1} \beta_{T-1} S_{T-1}^1 \end{aligned} \quad (3.96)$$

Da equação (3.87), temos que $S_T^1 | \mathcal{F}_{T-1}$ é dado por

$$S_T^1 = S_{T-1}^1 + S_{T-1}^1 \Delta b_T \quad (3.97)$$

portanto

$$\begin{aligned}
E_{\mathbb{Q}}(S_T^1 | \mathcal{F}_{T-1}) &= E_{\mathbb{Q}}[S_{T-1}^1 + S_{T-1}^1 \Delta b_T | \mathcal{F}_{T-1}] \\
&= S_{T-1}^1 + S_{T-1}^1 E_{\mathbb{Q}}[\Delta b_T | \mathcal{F}_{T-1}]
\end{aligned} \tag{3.98}$$

assim da equação (3.88) e considerando agora $\mathbb{Q} = (q, 1 - q)$

$$\begin{aligned}
E_{\mathbb{Q}}(S_T^1 | \mathcal{F}_{T-1}) &= S_{T-1}^1 + S_{T-1}^1 (uq + d(1 - q)) \\
E_{\mathbb{Q}}(S_T^1 | \mathcal{F}_{T-1}) &= S_{T-1}^1 [1 + uq + d(1 - q)] \\
E_{\mathbb{Q}}(S_T^1 | \mathcal{F}_{T-1}) &= S_{T-1}^1 [(1 + d) + (u - d)q]
\end{aligned} \tag{3.99}$$

da equações (3.96) e (3.99), vem que

$$\begin{aligned}
S_{T-1}^1 [(1 + d) + (u - d)q] &= \beta_T^{-1} \beta_{T-1} S_{T-1}^1 \\
(1 + d) + (u - d)q &= (1 + r)^T (1 + r)^{-(T-1)} \\
&= (1 + r)
\end{aligned} \tag{3.100}$$

assim

$$q = \frac{r - d}{u - d} \tag{3.101}$$

onde $0 < q \leq 1$. Desta relação podemos dizer que se r satisfizer

$$d < r < u \tag{3.102}$$

então existe a MME.

O fato de avaliarmos uma opção sobre a medida martingale equivalente (MME) significa que estamos assumindo uma situação de indiferença ao risco. Nesse caso, os investidores não exigem compensação por riscos e o retorno esperado de todos os títulos é a taxa de juros livre de risco.

Se considerarmos $T = 1$ teremos o modelo **Binomial Uniperíodo**. Deste modo, pela proposição (3.7), temos que o valor da opção em $t = 0$ é dado por

$$\begin{aligned}
H_0 &= \beta_0^{-1} E_{\mathbb{Q}}[\beta_T H_T | \mathcal{F}_0] \\
&= \beta_0^{-1} \beta_T E_{\mathbb{Q}}[H_T | \mathcal{F}_0] \\
&= (1 + r) E_{\mathbb{Q}}[H_T | \mathcal{F}_0]
\end{aligned} \tag{3.103}$$

com condição final H_T , que dependerá do tipo da opção.

Repetindo este processo recursivo inverso, para um t qualquer teremos:

$$\prod_{j=t}^{T-1} \frac{\Delta \tilde{S}_{j+1}}{\tilde{S}_j} = \prod_{j=t}^{T-1} \Delta b_{j+1} \sim \text{binomial}(T-t, p) \quad (3.104)$$

supondo $\Delta b_{t+1}, \dots, \Delta b_T$, independentes.

Deste modo, pela proposição (3.7), temos que o valor da opção em t é dado por

$$H_t = \beta_{T-t} E_Q [H_T | \mathcal{F}_t] \quad (3.105)$$

onde H_T agora será uma binomial.

3.4.3 Avaliação da Opção Européia pelo Modelo Binomial

Vamos determinar o preço da opção de compra, a opção de venda sai diretamente da paridade entre o preço da opção de compra e venda veja equação (2.62).

Assim temos que para $H_t = c_t$ e condição final $H_T = c_T = \max(S_T^1 - K, 0)$, veja equação (2.50), portanto

$$\begin{aligned} c_t &= \beta_{T-t} E_Q [c_T | \mathcal{F}_t] \\ &= \beta_{T-t} E_Q [\max(S_T^1 - K, 0) | \mathcal{F}_t] \\ &= \beta_{T-t} \sum_{j=0}^{T-t} \binom{T-t}{j} q^j (1-q)^{T-t-j} \max[S_t^1 (1+u)^j (1+d)^{T-t-j} - K] \end{aligned}$$

considerando τ o primeiro inteiro para o qual $S_t^1 (1+u)^\tau (1+d)^{T-t-\tau} > K$ assim

$$\begin{aligned} c_t &= \beta_{T-t} S_t^1 \sum_{j=\tau}^{T-t} \binom{T-t}{j} q^j (1-q)^{T-t-j} (1+u)^j (1+d)^{T-t-j} - \\ &\quad \beta_{T-t} K \sum_{j=\tau}^{T-t} \binom{T-t}{j} q^j (1-q)^{T-t-j} (1+u)^j (1+d)^{T-t-j} \end{aligned} \quad (3.106)$$

Esta fórmula é conhecida como a fórmula de Cox-Ross-Rubinstein (CRR), veja [Cox et al, 79]. A opção de venda européia sai direto da relação de paridade entre a opção de compra e venda, veja equação (2.62), deste modo temos que

$$p_t = c_t + K (1+r)^{t-T} - S_t^1 \quad (3.107)$$

3.5 Análise dos Modelos: Versões a Tempo Discreto

Vamos considerar três modelos discretos mais gerais que os modelos binomiais. Agora ao invés de termos apenas dois estados possíveis teremos um número não enumerável, o que é bem mais real do que termos apenas dois estados. O nosso modelo agora será um modelo onde a probabilidade de transição será uma normal. Os três modelos são

1. $S_{t+1}^1 = S_t^1 + \mu_{\Delta t} + \sigma_{\Delta t} Z_{t+1};$
2. $S_{t+1}^1 = S_t^1 + S_t^1 \mu_{\Delta t} + \sigma_{\Delta t} Z_{t+1};$
3. $S_{t+1}^1 = S_t^1 + S_t^1 \mu_{\Delta t} + S_t^1 \sigma_{\Delta t} Z_{t+1};$

onde $Z_{t+1} \sim N(0, 1)$, com

$$E[Z_{t+i} Z_{t+j}] = \begin{cases} 1 & \text{se } i = j \\ 0 & \text{se } i \neq j \end{cases} \quad (3.108)$$

$\mu_{\Delta t}$ e $\sigma_{\Delta t}$ são parâmetros diretamente proporcionais a Δt , ou seja

$$\mu_{\Delta t} = \mu \Delta t, \sigma_{\Delta t} = \sigma \sqrt{\Delta t} \text{ e } Z = \frac{\Delta B_{t_{j+1}}}{\sqrt{\Delta t}} \quad (3.109)$$

Nosso objetivo é comparar as estimativas dos modelos contínuos com estes modelos que são as versões discretas dos modelos contínuos quando aproximamos estes modelos utilizando o esquema de Euler, veja [Kloeden & Platen, 95].

3.5.1 Avaliação das Opções Europeias

Vamos determinar deste modo para cada modelo a formula de precificação da opção de compra europeia.

Modelo 1: $S_{t+1}^1 = S_t^1 + \mu_{\Delta t} + \sigma_{\Delta t} Z_{t+1}$

Sobre $\mathbb{Q} = (\mu^*, \sigma^*)$ a dinâmica de S_t é dado por

$$S_{t+1}^1 = S_t^1 + \mu^* + \sigma^* Z_{t+1} \quad (3.110)$$

sabemos que sobre a medida martingale equivalente $\mathbb{Q}$ o processo preço da ação descontado $(\tilde{S}_t)_{t \in [0, T]}$ é um martingale, portanto temos que

$$\begin{aligned}
E_{\mathbb{Q}} \left(\tilde{S}_{t+1}^1 - \tilde{S}_t^1 \middle| \mathcal{F}_t \right) &= 0 \\
\tilde{S}_t^1 &= E_{\mathbb{Q}} \left(\tilde{S}_{t+1}^1 \middle| \mathcal{F}_t \right) \\
S_t^1 \beta_t &= \beta_{t+1} E_{\mathbb{Q}} \left(S_{t+1}^1 \middle| \mathcal{F}_t \right) \\
S_t^1 (1+r)^{-t} &= (1+r)^{1-t} E_{\mathbb{Q}} \left(S_t^1 + \mu^* + \sigma^* \Delta Z_{t+1} \middle| \mathcal{F}_t \right) \\
S_t^1 (1+r) &= S_t^1 + \mu^* + \underbrace{\phi^* E_{\mathbb{Q}} (\Delta Z_{t+1} \middle| \mathcal{F}_t)}_{=0}
\end{aligned} \tag{3.111}$$

portanto

$$\mu^* = S_t^1 r \text{ e } \sigma^* = \sigma_{\Delta t} \tag{3.112}$$

deste modo sobre $\mathbb{Q}(\mu^*, \sigma^*)$ o processo $(S_t^1)_{t \in [0, T]}$ tem a seguinte dinâmica

$$S_{t+1}^1 = S_t^1 + S_t^1 r + \sigma_{\Delta t} Z_{t+1} \tag{3.113}$$

Ora sabemos que para um investidor aplicar o seu dinheiro em um ativo de risco ele espera que o retorno esperado do investimento seja maior que o retorno de um investimento feito a taxa de juros livre de risco isto é

$$E_{\mathbb{P}} \left(S_T^1 \middle| \mathcal{F}_t \right) > S_t^1 (1+r)^{T-t} \tag{3.114}$$

sobre $\mathbb{P}$ temos que $S_T^1 \middle| \mathcal{F}_t$ é dado por

$$S_T^1 = S_t^1 + (T-t) \mu_{\Delta t} + \sigma_{\Delta t} \sum_{j=1}^{T-t} (1+r)^{T-t-j} Z_{t+j} \tag{3.115}$$

como $\Delta Z_{t+1} = \Delta Z_{t+2} = \dots = \Delta Z_T = Z \sim N(0, 1)$ são independentes, então segue que

$$S_T^1 = S_t^1 + (T-t) \mu_{\Delta t} + \sigma_{\Delta t} Z \sum_{j=1}^{T-t} (1+r)^{T-t-j} \tag{3.116}$$

assim $E_{\mathbb{P}} \left(S_T^1 \middle| \mathcal{F}_t \right) > S_t^1 (1+r)^{T-t}$ implica que

$$S_t^1 + (T-t) \mu_{\Delta t} > S_t^1 (1+r)^{T-t} \tag{3.117}$$

e portanto

$$\mu_{\Delta t} > \frac{S_t^1 \left[(1+r)^{T-t} - 1 \right]}{(T-t)} \tag{3.118}$$

A diferença

$$E_{\mathbb{P}}(S_T^1 | \mathcal{F}_t) - S_t^1 (1+r)^{T-t} \quad (3.119)$$

é conhecida como prêmio pelo risco.

Sobre $\mathbb{Q}$ temos que $S_T^1 | \mathcal{F}_t$ é dado por

$$S_T^1 = S_t^1 (1+r)^{T-t} + \sigma_{\Delta t} Z \sum_{j=1}^{T-t} (1+r)^{T-t-j} \quad (3.120)$$

Deste modo o preço justo para uma opção européia é dada por

$$\begin{aligned} c_t &= (1+r)^{t-T} E_{\mathbb{Q}}[c_T | \mathcal{F}_t] \\ &= (1+r)^{t-T} E_{\mathbb{Q}}[\max[(S_T^1 - K), 0] | \mathcal{F}_t] \\ &= (1+r)^{t-T} \int_{-\infty}^{\infty} \max[(S_T^1 - K), 0] | \mathcal{F}_t \Psi(S_T^1) dS_T^1 \\ &= (1+r)^{t-T} \int_K^{\infty} (S_T^1 - K) \Psi(S_T^1) dS_T^1 \\ &= (1+r)^{t-T} \int_K^{\infty} S_T^1 \Psi(S_T^1) dS_T^1 - K \int_K^{\infty} \Psi(S_T^1) dS_T^1 \end{aligned} \quad (3.121)$$

como

$$S_T^1 \sim N \left[S_t^1 (1+r)^{T-t}, \left[\sigma_{\Delta t} \sum_{j=1}^{T-t} (1+r)^{T-t-j} \right]^2 \right] \quad (3.122)$$

então

$$S_T^1 - K = c_T \sim N \left[S_t^1 (1+r)^{T-t} - K, \left[\sigma_{\Delta t} \sum_{j=1}^{T-t} (1+r)^{T-t-j} \right]^2 \right] \quad (3.123)$$

portanto

$$c_t = (1+r)^{t-T} \int_0^{\infty} c_T \Psi(c_T) dc_T \quad (3.124)$$

fazendo a seguinte mudança de variável

$$Z = \frac{c_T - S_t^1 (1+r)^{T-t} + K}{\sigma_{\Delta t} \sum_{j=1}^{T-t} (1+r)^{T-t-j}} \sim N(0, 1) \quad (3.125)$$

deste modo

$$c_T = Z \sigma_{\Delta t} \sum_{j=1}^{T-t} (1+r)^{T-t-j} + S_t^1 (1+r)^{T-t} - K \quad (3.126)$$

assim

$$\begin{aligned}
c_t &= (1+r)^{t-T} \int_0^\infty c_T \Psi(c_T) dc_T \\
&= (1+r)^{t-T} \int_A^\infty \left[Z \sigma_{\Delta t} \sum_{j=1}^{T-t} (1+r)^{T-t-j} + S_t^1 (1+r)^{T-t} - K \right] \Psi(Z) dZ \\
&= (1+r)^{t-T} \sigma_{\Delta t} \sum_{j=1}^{T-t} (1+r)^{T-t-j} \int_A^\infty Z \Psi(Z) dZ + \\
&\quad \left(S_t^1 - K S_t (1+r)^{t-T} \right) \int_A^\infty \Psi(Z) dZ
\end{aligned} \tag{3.127}$$

onde

$$A = \frac{k - S_t^1 (1+r)^{T-t}}{\sigma_{\Delta t} \sum_{j=1}^{T-t} (1+r)^{T-t-j}} \tag{3.128}$$

portanto o preço da opção de compra européia para este modelo é dado por

$$c_t = \frac{\sigma_{\Delta t} \exp\left(-\frac{A^2}{2}\right) \sum_{j=1}^{T-t} (1+r)^{-j}}{\sqrt{2\pi}} + \left(S_t^1 - K (1+r)^{t-T} \right) \times [1 - \Phi(A)] \tag{3.129}$$

O preço da opção de venda européia sai da relação de paridade entre a opção de compra e venda, veja equação (2.62), assim temos que

$$c_t + K (1+r)^{t-T} = p_t + S_t^1 \tag{3.130}$$

portanto

$$p_t = c_t + K (1+r)^{t-T} - S_t^1 \tag{3.131}$$

Modelo 2: $S_{t+1}^1 = S_t^1 + S_t^1 \mu_{\Delta t} + \sigma_{\Delta t} \Delta Z$

Começamos determinando a medida martingale equivalente $\mathbb{Q}(\mu^*, \sigma^*)$. Deste modo temos que sobre $\mathbb{Q}$

$$S_{t+1}^1 = S_t^1 + S_t^1 \mu^* + \sigma^* Z_{t+1} = S_t^1 [1 + \mu^*] + \sigma^* Z_{t+1} \tag{3.132}$$

sabemos que sobre a medida martingale equivalente $\mathbb{Q}$ o processo preço da ação descontado $\left(\tilde{S}_t^1\right)_{t \in [0, T]}$ é um martingale, portanto temos que

$$\begin{aligned}
E_{\mathbb{Q}}(\tilde{S}_{t+1}^1 - \tilde{S}_t^1 | \mathcal{F}_t) &= 0 \\
\tilde{S}_t^1 &= E_{\mathbb{Q}}(\tilde{S}_{t+1}^1 | \mathcal{F}_t) \\
S_t^1 \beta_t &= \beta_{t+1} E_{\mathbb{Q}}(S_{t+1}^1 | \mathcal{F}_t) \\
S_t^1 (1+r)^{-t} &= (1+r)^{1-t} E_{\mathbb{Q}}(S_t^1 [1 + \mu^*] + \sigma^* \Delta Z_{t+1} | \mathcal{F}_t) \\
S_t^1 (1+r) &= S_t^1 [1 + \mu^*] + \sigma^* \underbrace{E_{\mathbb{Q}}(\Delta Z_{t+1} | \mathcal{F}_t)}_{=0}
\end{aligned}$$

portanto

$$\mu^* = r \text{ e } \sigma^* = \sigma_{\Delta t} \quad (3.133)$$

deste modo sobre $\mathbb{Q}(\mu^*, \sigma^*)$ o processo $(S_t)_{t \in [0, T]}$ tem a seguinte dinâmica

$$S_{t+1}^1 = S_t^1 + S_t^1 r + \sigma_{\Delta t} Z_{t+1} \quad (3.134)$$

Sabemos que o investidor espera que o retorno esperado do investimento seja maior que o retorno de um investimento feito a taxa de juros livre de risco isto é

$$E_{\mathbb{P}}(S_T^1 | \mathcal{F}_t) > S_t^1 (1+r)^{T-t} \quad (3.135)$$

sobre $\mathbb{P}$ temos que $S_T^1 | \mathcal{F}_t$ é dado por

$$S_T^1 = S_t^1 (1 + \mu_{\Delta t})^{T-t} + \sigma_{\Delta t} \sum_{j=1}^{T-t} (1+r)^{T-t-j} Z_{t+j} \quad (3.136)$$

como $\Delta Z_{t+1} = \Delta Z_{t+2} = \dots = \Delta Z_T = Z \sim N(0, 1)$ são independentes, então segue que

$$S_T^1 = S_t^1 (1 + \mu_{\Delta t})^{T-t} + \sigma_{\Delta t} Z \sum_{j=1}^{T-t} (1+r)^{T-t-j} \quad (3.137)$$

assim $E_{\mathbb{P}}(S_T^1 | \mathcal{F}_t) > S_t^1 (1+r)^{T-t}$ implica que

$$S_t^1 (1 + \mu_{\Delta t})^{T-t} > S_t^1 (1+r)^{T-t} \quad (3.138)$$

e portanto

$$\mu_{\Delta t} > r \quad (3.139)$$

note que aqui ao contrário do modelo anterior basta verificar se o retorno de curto prazo do ativo $\mu_{\Delta t}$ é maior que a taxa de juros livre de risco r .

Sobre $\mathbb{Q}$ temos que $S_T^1 | \mathcal{F}_t$ é dado por

$$S_T^1 = S_t^1 (1+r)^{T-t} + \sigma_{\Delta t} Z \sum_{j=1}^{T-t} (1+r)^{T-t-j} \quad (3.140)$$

Note que sobre $\mathbb{Q}$ tanto o modelo 1 como o modelo 2 tem a mesma dinâmica portanto o preço justo para uma opção europeia é o mesmo do modelo anterior

$$c_t = \frac{\sigma_{\Delta t} \exp\left(-\frac{A^2}{2}\right) \sum_{j=1}^{T-t} (1+r)^{-j}}{\sqrt{2\pi}} + \left(S_t^1 - K(1+r)^{t-T}\right) \times [1 - \Phi(A)] \quad (3.141)$$

O preço da opção de venda europeia sai da relação de paridade entre a opção de compra e venda veja equação (2.62), assim temos que

$$p_t = c_t + K(1+r)^{t-T} - S_t^1 \quad (3.142)$$

Modelo 3: $S_{t+1}^1 = S_t^1 + S_t^1 \mu_{\Delta t} + S_t^1 \sigma_{\Delta t} \Delta Z$

Sobre $\mathbb{Q} = (\mu^*, \sigma^*)$ a dinâmica de S_t^1 é dado por

$$S_{t+1}^1 = S_t^1 + S_t^1 \mu^* + S_t^1 \sigma^* Z_{t+1} = S_t^1 [1 + \mu^*] + S_t^1 \sigma^* Z_{t+1} \quad (3.143)$$

sabemos que sobre a medida martingale equivalente $\mathbb{Q}$ o processo preço da ação descontado $(\tilde{S}_t^1)_{t \in [0, T]}$ é um martingale, portanto temos que

$$\begin{aligned} E_{\mathbb{Q}}(\tilde{S}_{t+1}^1 - \tilde{S}_t^1 | \mathcal{F}_t) &= 0 \\ \tilde{S}_t^1 &= E_{\mathbb{Q}}(\tilde{S}_{t+1}^1 | \mathcal{F}_t) \\ S_t^1 \beta_t &= \beta_{t+1} E_{\mathbb{Q}}(S_{t+1}^1 | \mathcal{F}_t) \\ S_t^1 (1+r)^{-t} &= (1+r)^{1-t} E_{\mathbb{Q}}(S_t^1 [1 + \mu^*] + S_t^1 \sigma^* Z_{t+1} | \mathcal{F}_t) \\ S_t^1 (1+r) &= S_t^1 [1 + \mu^*] + S_t^1 \sigma^* \underbrace{E_{\mathbb{Q}}(\Delta Z_{t+1} | \mathcal{F}_t)}_{=0} \end{aligned}$$

portanto

$$\mu^* = r \text{ e } \sigma^* = \sigma_{\Delta t} \quad (3.144)$$

deste modo sobre $\mathbb{Q}(\mu^*, \sigma^*)$ o processo $(S_t^1)_{t \in [0, T]}$ tem a seguinte dinâmica

$$S_{t+1}^1 = S_t^1 + S_t^1 r + S_t^1 \sigma_{\Delta t} Z_{t+1} \quad (3.145)$$

Sabemos que o investidor espera que o retorno esperado do investimento seja maior que o retorno de um investimento feito a taxa de juros livre de risco isto é

$$E_{\mathbb{P}}(S_T^1 | \mathcal{F}_t) > S_t^1 (1+r)^{T-t} \quad (3.146)$$

sobre $\mathbb{P}$ temos que $S_T^1 | \mathcal{F}_t$ é dado por

$$S_T^1 = S_t^1 \prod_{i=t}^{T-1} [(1 + \mu_{\Delta t}) + \sigma_{\Delta t} Z_{i+1}] \quad (3.147)$$

como $\Delta Z_{t+1} = \Delta Z_{t+2} = \dots = \Delta Z_T = Z \sim N(0, 1)$ são independentes, então segue que

$$S_T^1 = S_t^1 [(1 + \mu_{\Delta t}) + \sigma_{\Delta t} Z]^{T-t} \quad (3.148)$$

agora sabemos que para $T - t$ suficientemente grande temos que

$$[1 + \mu_{\Delta t} + \sigma_{\Delta t} Z]^{T-t} \approx \exp[\mu_{\Delta t}(T-t) + \sigma_{\Delta t}(T-t)Z] \quad (3.149)$$

deste modo iremos utilizar esta aproximação pois ela permitirá que tenhamos uma forma fechada para o valor da opção, caso contrário teríamos que utilizar algum algoritmo numérico o que acarretaria também em um erro de aproximação. Deste modo temos que a equação 3.148, fica da seguinte maneira:

$$S_T^1 = S_t^1 \exp[\mu_{\Delta t}(T-t) + \sigma_{\Delta t}(T-t)Z] \quad (3.150)$$

portanto

$$\begin{aligned} E_{\mathbb{P}}(S_T^1 | \mathcal{F}_t) &= E_{\mathbb{P}}(S_t^1 \exp[\mu_{\Delta t}(T-t) + \sigma_{\Delta t}(T-t)Z] | \mathcal{F}_t) \\ &= S_t^1 \exp[\mu_{\Delta t}(T-t)] E_{\mathbb{P}}(\exp[\sigma_{\Delta t}(T-t)Z] | \mathcal{F}_t) \\ &= S_t^1 \exp[\mu_{\Delta t}(T-t)] \int_{-\infty}^{\infty} \frac{\exp(\sigma_{\Delta t}(T-t)Z) \exp(-\frac{Z^2}{2})}{\sqrt{2\pi}} dZ \\ &= S_t^1 \exp[\mu_{\Delta t}(T-t)] \exp\left(\frac{[\sigma_{\Delta t}(T-t)]^2}{2}\right) \\ &= S_t^1 \exp\left[\mu_{\Delta t}(T-t) + \frac{(\sigma_{\Delta t}(T-t))^2}{2}\right] \end{aligned}$$

assim $E_{\mathbb{P}}(S_T^1 | \mathcal{F}_t) > S_t^1 (1+r)^{T-t}$ implica que

$$S_t^1 \exp\left[\mu_{\Delta t}(T-t) + \frac{(\sigma_{\Delta t}(T-t))^2}{2}\right] > S_t^1 \exp[r(T-t)] \quad (3.151)$$

note que aqui estamos utilizando a mesma aproximação que foi feita acima para a expressão do lado direito, portanto

$$\mu_{\Delta t} > r - \frac{\sigma_{\Delta t}^2 (T-t)}{2} \quad (3.152)$$

Sobre $\mathbb{Q}$ temos que $S_T^1 | \mathcal{F}_t$, utilizando a aproximação acima, é dado por

$$S_T^1 = S_t^1 \exp [r (T-t) + \sigma_{\Delta t} (T-t) Z] \quad (3.153)$$

Deste modo o preço justo para uma opção de compra europeia é dada por

$$\begin{aligned} c_t &= \beta_{T-t} E_{\mathbb{Q}} (c_T | \mathcal{F}_t) \\ &= (1+r)^{t-T} E_{\mathbb{Q}} (\max [(S_T^1 - K), 0] | \mathcal{F}_t) \end{aligned}$$

Vamos primeiro calcular $E_{\mathbb{Q}} [\max [(S_T^1 - K), 0] | \mathcal{F}_t]$, deste modo temos que

$$\begin{aligned} &= \int_{-\infty}^{\infty} \max (S_t^1 \exp [r (T-t) + \sigma_{\Delta t} (T-t) Z] - K, 0) \Psi (Z) dZ \\ &= \int_A^{\infty} (S_t^1 \exp [r (T-t) + \sigma_{\Delta t} (T-t) Z] - K) \Psi (Z) dZ \end{aligned} \quad (3.154)$$

onde é a solução para

$$S_t^1 \exp [r (T-t) + \sigma_{\Delta t} (T-t) A] - K = 0 \quad (3.155)$$

portanto temos que

$$A = \frac{1}{\sigma_{\Delta t} (T-t)} \ln \left(\frac{K}{S_t^1} \right) + \frac{r}{\sigma_{\Delta t}} \quad (3.156)$$

Assim

$$\begin{aligned}
&= \frac{1}{\sqrt{2\pi}} \int_A^\infty (S_t^1 \exp[r(T-t) + \sigma_{\Delta t}(T-t)Z] - K) \exp\left(-\frac{Z^2}{2}\right) dZ \\
&= \frac{1}{\sqrt{2\pi}} S_t^1 \exp(r(T-t)) \int_A^\infty \exp(\sigma_{\Delta t}(T-t)Z) \exp\left(-\frac{Z^2}{2}\right) dZ - \\
&\quad \frac{1}{\sqrt{2\pi}} \int_A^\infty K \exp\left(-\frac{Z^2}{2}\right) dZ \\
&= S_t^1 \exp(r(T-t)) \exp\left(\frac{[\sigma_{\Delta t}(T-t)]^2}{2}\right) \int_A^\infty \exp\left(-\frac{[Z - \sigma_{\Delta t}(T-t)]^2}{2}\right) dZ \\
&\quad - K \frac{1}{\sqrt{2\pi}} \int_A^\infty \exp\left(-\frac{Z^2}{2}\right) dZ \\
&= S_t^1 \exp\left[r(T-t) + \frac{[\sigma_{\Delta t}(T-t)]^2}{2}\right] \int_{A - \sigma_{\Delta t}(T-t)}^\infty \exp\left(-\frac{x^2}{2}\right) dx \\
&\quad - K \times [1 - \Phi(A)] \\
&= S_t^1 \exp\left[r(T-t) + \frac{[\sigma_{\Delta t}(T-t)]^2}{2}\right] [1 - \Phi(A - \sigma\sqrt{T-t})] \\
&\quad - K \times [1 - \Phi(A)]
\end{aligned} \tag{3.157}$$

Portanto o preço justo da opção de compra européia é dada por

$$\begin{aligned}
c_t &= e^{-r(T-t)} E_{\mathbb{Q}} [\max[(S_T^1 - K), 0] | \mathcal{F}_t] \\
&= S_t^1 \exp\left[\frac{[\sigma_{\Delta t}(T-t)]^2}{2}\right] [1 - \Phi(A - \sigma\sqrt{T-t})] - \\
&\quad K e^{-r(T-t)} \times [1 - \Phi(A)]
\end{aligned} \tag{3.158}$$

onde $\Phi(\cdot)$ é a função densidade acumulada da normal padrão.

A opção de compra é dada por

$$p_t = c_t + K(1+r)^{t-T} - S_t^1 \tag{3.159}$$

Capítulo 4

Precificação das Opções a Tempo Contínuo

Os primeiros estudos em finanças a tempo contínuo são anteriores aos trabalhos de Merton, [Merton, 69], [Merton, 71] e [Merton, 73b]. No entanto estes são os trabalhos que marcaram o início da moderna teoria das finanças. Nestes últimos 30 anos desde estas publicações, a metodologia a tempo contínuo tem se tornado uma parte integrante da ciência econômico financeira. De fato, em certas disciplinas em finanças, tais como formação de preços de derivativos, seleção de carteiras, por exemplo, a metodologia a tempo contínuo tem provado ser o caminho mais atrativo para conduzir as pesquisas.

No início dos anos 70, Fischer Black e Myron Scholes, veja [Black & Scholes, 73], fizeram uma descoberta importante sobre a precificação de opções de ações. Essa descoberta acabou tendo enorme influência sobre a forma de precificar e hedgear opções pelos participantes do mercado.

A análise de Black e Scholes é análoga à de não arbitragem utilizada para avaliar opções, quando mudanças no preço da ação são binomiais, veja [Hull, 96] e [Hull, 97]. Desta maneira é criada uma carteira sem risco, constituída de uma posição na opção e uma posição na ação objeto. Na ausência de oportunidades de arbitragem, o retorno da carteira deve ser a taxa de juro livre de risco r .

O motivo pelo qual podemos montar uma carteira sem risco se relaciona ao fato de que tanto o preço da ação quanto o preço da opção são afetados pela mesma origem de incerteza: as oscilações de preço da ação. Em qualquer período curto de tempo, o preço de uma opção de compra está positiva e perfeitamente correlacionado com o preço da ação objeto e o preço de uma opção de venda está negativa e perfeitamente correlacionado com o preço da ação

objeto. Em ambos os casos, quando elaboramos uma carteira apropriada com a ação e a opção, o lucro ou perda da posição na ação sempre compensa o lucro ou perda da posição na opção, de modo que o valor global da carteira, no final de um período de tempo, seja conhecido com certeza. Adiante, estes comentários iniciais ficaram mais claros quando mostrarmos como foi a abordagem de Black e Scholes.

4.1 Considerações Iniciais

Dentro das possíveis estratégias de transação, foi visto anteriormente que estamos interessados somente naquelas que são autofinanciáveis, veja capítulo 2. Depois fizemos mais uma restrição em que considerávamos para nossas propostas de precificação apenas aquelas transações em que o processo valor da carteira era um $\mathbb{Q}$ -martingale, a estas estratégias chamamos de estratégias admissíveis. Vamos agora colocar estes conceitos que foram vistos a tempo discreto, na estrutura contínua.

Definição 4.1 (Processo valor da Carteira) *O processo valor da carteira $(V_t^\varphi)_{t \in [0, T]}$ é dado pelo produto escalar*

$$V_t^\varphi \triangleq \varphi_t \cdot S_t = \sum_{i=0}^d \varphi_t^i S_t^i$$

para $t \in [0, T]$, $T \in \mathbb{R}^+$, $S_t = (S_t^0, S_t^1, \dots, S_t^d)$, $\varphi_t = (\varphi_t^0, \varphi_t^1, \dots, \varphi_t^d)'$, $d \in \mathbb{Z}^+$, tendo a seguinte dinâmica

$$dV_t^\varphi = \varphi_t \cdot dS_t + S_t \cdot d\varphi_t + \underbrace{d\langle \varphi_t \cdot S_t \rangle}_{=0}$$

onde $\langle \varphi_t \cdot S_t \rangle$ denota a variação quadrática entre φ_t e S_t e que por serem independentes implica $\langle \varphi_t \cdot S_t \rangle = 0$. Assim

$$dV_t^\varphi = \varphi_t \cdot dS_t + S_t \cdot d\varphi_t$$

Definição 4.2 (Estratégia Autofinanciável) *Uma estratégia de transação φ é autofinanciável, $\varphi \in \Phi$, se*

$$S_t \cdot d\varphi_t = 0$$

onde $t \in [0, T]$, $T \in \mathbb{R}^+$, $S_t = (S_t^0, S_t^1, \dots, S_t^d)$, $d \in \mathbb{Z}^+$.

Como já foi enfatizado anteriormente, as implicações financeiras da equação acima são que todas as mudanças de posições na carteira são feitas de tal forma que não haja fluxos ou

influxos de capital da carteira, isto é, fluxo de caixa zero no final de cada período até o período $t < T$.

Deste modo temos que a dinâmica do processo valor da carteira para uma estratégia autofinanciável é dada por

$$dV_t^\varphi = \varphi_t \cdot dS_t \quad (4.160)$$

Definição 4.3 (Processo Ganho da Carteira) O processo ganho $(G_t^\varphi)_{t \in [0, T]}$ de uma estratégia de transação φ é dado por

$$G_t^\varphi \triangleq \int_0^t \varphi_s \cdot dS_s$$

onde $\varphi_t = \{\varphi_t^0, \varphi_t^1, \dots, \varphi_t^d\}$, $d \in \mathbb{Z}^+$ e $t \in [0, T]$, $T \in \mathbb{R}^+$.

Portanto a partir desta definição podemos escrever o processo valor da carteira $(V_t^\varphi)_{t \in [0, T]}$ para uma estratégia de transação autofinanciável $\varphi \in \Phi$ como

$$V_t^\varphi = V_0^\varphi + G_t^\varphi \quad (4.161)$$

Proposição 4.1 (Invariância do Numerário a Tempo Contínuo) Carteiras autofinanciáveis permanecem autofinanciáveis após uma mudança de numerário.

Prova: Seja S_t o vetor preços no tempo t e X_t o numerário, assuma que são contínuos. Temos que para uma estratégia autofinanciável φ_t

$$S_t \cdot d\varphi_t = 0$$

desde que X_t é estritamente positivo para todo $t \in [0, T]$, então podemos dividir ambos os lados por X_t assim temos que

$$X_t^{-1} S_t \cdot d\varphi_t = 0$$

portanto renormalizar o preço dos ativos não alterou a classe das estratégias autofinanciáveis. Assim aplicando a regra do produto para $X_t^{-1} S_t$, temos que

$$d(X_t^{-1} S_t) = X_t^{-1} dS_t + S_t d(X_t^{-1}) + d\langle X_t^{-1} S_t \rangle$$

Para o processo valor da carteira temos que

$$V_t^\varphi = \varphi_t \cdot S_t$$

com a seguinte dinâmica

$$dV_t^\varphi = \varphi_t \cdot dS_t + S_t \cdot d\varphi_t$$

desde modo aplicando a regra do produto para $X_t^{-1}V_t^\varphi$, temos que

$$d(X_t^{-1}V_t^\varphi) = X_t^{-1}dV_t^\varphi + V_t^\varphi d(X_t^{-1}) + d\langle X_t^{-1}V_t^\varphi \rangle$$

portanto temos que

$$\begin{aligned} d(X_t^{-1}V_t^\varphi) &= X_t^{-1}[\varphi_t \cdot dS_t + S_t \cdot d\varphi_t] + \\ &\quad (\varphi_t \cdot S_t) d(X_t^{-1}) + d\langle X_t^{-1}(\varphi_t \cdot S_t) \rangle \\ &= \varphi_t \cdot [X_t^{-1}dS_t + S_t d(X_t^{-1}) + d\langle X_t^{-1}S_t \rangle] + \\ &\quad \underbrace{X_t^{-1}S_t \cdot d\varphi_t}_{=0} \\ &= \varphi_t \cdot d(X_t^{-1}S_t) \end{aligned}$$

desde modo a carteira expressa no novo numérario permanece autofinanciável. ■

Como já foi dito anteriormente denotamos $(S_t^0)_{t \in [0, T]} = (X_t)_{t \in [0, T]}$, e o fator de desconto por

$$\beta_t = \frac{1}{S_t^0} \quad (4.162)$$

Assim o vetor preço descontado é dado por

$$\beta_t S_t = \tilde{S}_t = (\beta_t S_t^0, \beta_t S_t^1, \dots, \beta_t S_t^d) = (1, \tilde{S}_t^1, \dots, \tilde{S}_t^d) \quad (4.163)$$

Deste modo o processo valor da carteira descontado $(\tilde{V}_t^\varphi)_{t \in [0, T]}$ para uma estratégia de transação autofinanciável $\varphi \in \Phi$ é dado por

$$\tilde{V}_t^\varphi = \tilde{V}_0^\varphi + \tilde{G}_t^\varphi \quad (4.164)$$

e o processo ganho descontado por

$$\tilde{G}_t^\varphi = \int_0^t \varphi_s \cdot d\tilde{S}_s \quad (4.165)$$

Como foi visto no capítulo 2, sabemos que se no mercado não tivermos oportunidade de arbitragem então existe a medida martingale equivalente $\mathbb{Q}$ sobre a qual o vetor processo preço descontado $\tilde{S}_t$ é um martingale. Deste modo, partindo da premissa que o mercado é livre de arbitragem então o processo ganho descontado $(\tilde{G}_t^\varphi)_{t \in [0, T]}$ é um $\mathbb{Q}$ -martingale.

Proposição 4.2 (Estratégia Admissível a Tempo Contínuo) *Uma estratégia de transação φ é dita ser admissível se o processo valor da carteira descontado $(\tilde{V}_t^\varphi)_{t \in [0, T]}$ é um $\mathbb{Q}$ -martingale com valor inicial $\tilde{V}_0^\varphi$.*

Prova: A prova desta proposição é direta desde que $(\tilde{G}_t^\varphi)_{t \in [0, T]}$ é um $\mathbb{Q}$ -martingale. Partindo da hipótese que no mercado não existem oportunidades de arbitragem então o problema da precificação de uma opção fica reduzida a existência de uma estratégia autofinanciável admissível que seja uma estratégia duplicante. Estas foram as mesmas conclusões que chegamos no caso discreto. Desde modo, completamos esta parte com as seguintes definições.

Definição 4.4 *Um derivativo é dito ser “attainable” se existe pelo menos uma estratégia de transação admissível tal que*

$$(V_t^\varphi)_{t \in [0, T]} = (H_t^\varphi)_{t \in [0, T]}$$

chamamos a estratégia φ de uma estratégia duplicante para o derivativo H_t^φ .

Definição 4.5 *Um mercado financeiro é dito ser completo se todo derivativo é “attainable”. A partir da proposição acima e desta definições podemos chegar a mesma conclusão que chegamos para o caso discreto.*

Proposição 4.3 *O processo preço da arbitragem de todo derivativo “attainable” $(H_t)_{t \in [0, T]}$ é dado pela Fórmula de Precificação do Risco Neutro*

$$\begin{aligned} H_t &= \beta_t^{-1} E_{\mathbb{Q}}[\beta_T H_T | \mathcal{F}_t] \\ &= \beta_{T-t} E_{\mathbb{Q}}[H_T | \mathcal{F}_t] \end{aligned}$$

para todo $t \in [0, T]$, $T \in \mathbb{R}^+$. Com condição final H_T .

Vamos agora a partir destes resultados fazer a análise dos modelos.

4.2 Premissas do Modelo

Assim como foi feito no modelo, veja seção (3.4), anterior as premissas do modelo são:

1. O mercado é perfeito;
2. O mercado é eficiente;
3. O mercado é completo;
4. A ação não receberá dividendos durante a vida da opção;

5. A taxa de juros livre de risco r é constante

Dos modelos analisados, apenas o modelo log-normal tem utilização efetiva em finanças, os outros modelos foram estudados mais com o intuito de ser analisado o problema da inferência dos parâmetros de uma EDE.

Assim este capítulo está organizado da seguinte maneira:

- Primeiro será deduzido o modelo de Black-Scholes;
- Depois será feita a análise para os outros dois modelos;

A abordagem utilizada nestas análises é essencialmente diferente da utilizada em [Black & Scholes, 73]. Aqui utilizaremos as ferramentas do cálculo estocástico na análise do problema da precificação de opções.

4.3 O Modelo Log-Normal: Black-Scholes

Descreveremos a partir de agora a dinâmica do modelo de Black-Scholes para a precificação de uma opção. O processo está definido sobre um base estocástica $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$. O parâmetro tempo t toma valores em $[0, T]$. O modelo consiste de um ativo sem risco cujo preço no tempo t é S_t^0 , e um ativo de risco, a ação objeto da opção, cujo preço no tempo t é S_t^1 .

Seja r uma constante não negativa que representa a taxa de juros instantânea do ativo sem risco. Então o montante aplicado neste ativo sem risco capitalizado continuamente é descrito pela seguinte equação diferencial ordinária

$$dS_t^0 = rS_t^0 dt \quad (4.166)$$

como já foi dito anteriormente, veja seção (3.1), consideraremos que o montante inicial aplicado no ativo sem risco é $S_0^0 = 1$, então a equação (4.166) com esta condição inicial tem a seguinte solução

$$S_t^0 = e^{rt}, t \geq 0 \quad (4.167)$$

e portanto o fator de desconto é dado por

$$\beta_t = \frac{1}{S_t^0} = e^{-rt} \quad (4.168)$$

O processo preço do ativo de risco S_t^1 tem a seguinte dinâmica

$$dS_t^1 = \mu S_t^1 dt + \sigma S_t^1 dB_t \quad (4.169)$$

e o processo preço descontado do ativo de risco é dado por

$$\tilde{S}_t^1 = \frac{S_t^1}{S_t^0} = e^{-rt} S_t^1 \quad (4.170)$$

portanto pelo lema de Itô, temos que

$$\begin{aligned} d\tilde{S}_t^1 &= -re^{-rt} S_t^1 dt + e^{-rt} dS_t^1 \\ &= -r\tilde{S}_t^1 dt + e^{-rt} [\mu S_t^1 dt + \sigma S_t^1 dB_t] \\ &= (\mu - r)\tilde{S}_t^1 dt + \sigma\tilde{S}_t^1 dB_t \end{aligned} \quad (4.171)$$

Para determinar a medida martingale equivalente, que será a medida sobre a qual será feita a precificação da opção, iremos utilizar a fórmula de Cameron-Martin-Girsanov, veja Teorema D.6 Apêndice D, portanto temos que sobre $\mathbb{Q}$

$$B_t^* = \int_0^t \theta_s d\mathfrak{s} + B_t \quad (4.172)$$

é um processo de Wiener. Deste modo, sobre esta nova medida $\mathbb{Q}$, $\tilde{S}_t^1$ tem a seguinte dinâmica

$$\begin{aligned} d\tilde{S}_t^1 &= (\mu - r)\tilde{S}_t^1 dt + \sigma\tilde{S}_t^1 (dB_t^* - \theta_t dt) \\ &= (\mu - r - \sigma\theta_t)\tilde{S}_t^1 dt + \sigma\tilde{S}_t^1 dB_t^* \end{aligned} \quad (4.173)$$

desde que $\tilde{S}_t^1$ tem que ser um martingale local sobre $\mathbb{Q}$, devemos ter

$$\mu - r - \sigma\theta_t = 0, \quad t \in [0, T] \quad (4.174)$$

portanto da equação acima temos que

$$\theta_t = \theta = \frac{\mu - r}{\sigma} \quad (4.175)$$

assim

$$d\tilde{S}_t^1 = \sigma\tilde{S}_t^1 dB_t^* \quad (4.176)$$

e a dinâmica para S_t^1 sobre $\mathbb{Q}$ é dada por

$$\begin{aligned} dS_t^1 &= re^{rt}\tilde{S}_t^1 dt + e^{rt}d\tilde{S}_t^1 \\ &= rS_t^1 dt + e^{rt}(\sigma\tilde{S}_t^1 dB_t^*) \\ &= rS_t^1 dt + \sigma S_t^1 dB_t^* \end{aligned} \quad (4.177)$$

cuja solução é

$$S_T^1 = S_t^1 \exp \left[\left(r - \frac{\sigma^2}{2} \right) (T - t) + \sigma (B_T^* - B_t^*) \right] \quad (4.178)$$

portanto sobre $\mathbb{Q}$, S_T^1 tem distribuição log-normal. Deste modo do teorema 4.3, temos que a fórmula para uma opção é dada por

$$H_t = \beta_{T-t} E_{\mathbb{Q}} [H_T | \mathcal{F}_t] = e^{-r(T-t)} E_{\mathbb{Q}} [H_T | \mathcal{F}_t] \quad (4.179)$$

com condição final H_T . Para obter a equação diferencial de Black-Scholes usaremos a fórmula de Feynman-Kac, veja teorema D.4, deste modo consideremos primeiro o seguinte

$$G(t, S) = e^{-rT} E_{\mathbb{Q}} [H_T | \mathcal{F}_t] \quad (4.180)$$

pela fórmula de Feynman-Kac temos que $G(t, S)$ satisfaz a seguinte equação diferencial parcial

$$\frac{\partial G(t, S)}{\partial t} + rS \frac{\partial G(t, S)}{\partial S} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 G(t, S)}{\partial S^2} = 0 \quad (4.181)$$

portanto para $H_t = H(t, S) = e^{rt} G(t, S)$, temos que

$$G(t, S) = e^{-rt} H(t, S) \quad (4.182)$$

$$\frac{\partial G(t, S)}{\partial t} = e^{-rt} \frac{\partial H(t, S)}{\partial t} - r e^{-rt} H(t, S) \quad (4.183)$$

$$rS \frac{\partial G(t, S)}{\partial S} = rS e^{-rt} \frac{\partial H(t, S)}{\partial S} \quad (4.184)$$

$$\frac{1}{2} \sigma^2 S^2 \frac{\partial^2 G(t, S)}{\partial S^2} = \frac{1}{2} \sigma^2 S^2 e^{-rt} \frac{\partial^2 H(t, S)}{\partial S^2} \quad (4.185)$$

assim

$$\begin{aligned} \left[e^{-rt} \frac{\partial H(t, S)}{\partial t} - r e^{-rt} H(t, S) \right] + \left[rS e^{-rt} \frac{\partial H(t, S)}{\partial S} \right] + \left[\frac{1}{2} \sigma^2 S^2 e^{-rt} \frac{\partial^2 H(t, S)}{\partial S^2} \right] &= 0 \\ e^{-rt} \left[\frac{\partial H(t, S)}{\partial t} + rS \frac{\partial H(t, S)}{\partial S} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 H(t, S)}{\partial S^2} - rH(t, S) \right] &= 0 \end{aligned} \quad (4.186)$$

portanto

$$\frac{\partial H(t, S)}{\partial t} + rS \frac{\partial H(t, S)}{\partial S} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 H(t, S)}{\partial S^2} - rH(t, S) = 0 \quad (4.187)$$

com condição final $H(T, S)$. Esta é a conhecida equação diferencial parcial de Black-Scholes, veja [Black & Scholes, 73]. É claro que é muito mais fácil resolver a equação (4.179) do que a equação (4.187).

A opção de compra européia com a condição final $c_T = \max[(S_T^1 - K), 0]$, veja equação (2.50), é dada por

$$\begin{aligned} c_t &= e^{-r(T-t)} E_{\mathbb{Q}}[c_T | \mathcal{F}_t] \\ &= e^{-r(T-t)} E_{\mathbb{Q}}[\max[(S_T^1 - K), 0] | \mathcal{F}_t] \end{aligned} \quad (4.188)$$

Vamos primeiro calcular $E_{\mathbb{Q}}[\max[(S_T^1 - K), 0] | \mathcal{F}_t]$, deste modo temos que

$$\begin{aligned} &= \int_{-\infty}^{\infty} \max\left(S_t^1 \exp\left[\left(r - \frac{\sigma^2}{2}\right)(T-t) + \sigma(B_T - B_t)\right] - K, 0\right) \Psi(B_T - B_t) dB_T \\ &= \int_A^{\infty} \left(S_t^1 \exp\left[\left(r - \frac{\sigma^2}{2}\right)(T-t) + \sigma\sqrt{T-t}Z\right] - K\right) \Psi(Z) dZ \end{aligned} \quad (4.189)$$

onde $\Psi(Z) \sim N(0, 1)$ e A é a solução para

$$S_t^1 \exp\left[\left(r - \frac{\sigma^2}{2}\right)(T-t) + \sigma\sqrt{T-t}A\right] - K = 0 \quad (4.190)$$

portanto temos que

$$A = \frac{1}{\sigma\sqrt{T-t}} \ln\left(\frac{K}{S_t^1}\right) - \frac{1}{\sigma\sqrt{T-t}} \left(r - \frac{\sigma^2}{2}\right)(T-t)$$

Assim

$$\begin{aligned} &= \frac{1}{\sqrt{2\pi}} \int_A^{\infty} \left(S_t^1 \exp\left[\left(r - \frac{\sigma^2}{2}\right)(T-t) + \sigma\sqrt{T-t}Z\right] - K\right) \exp\left(-\frac{Z^2}{2}\right) dZ \\ &= \frac{1}{\sqrt{2\pi}} S_t^1 \exp(r(T-t)) \int_A^{\infty} \exp\left(\sigma\sqrt{T-t}Z - \frac{\sigma^2}{2}(T-t)\right) \exp\left(-\frac{Z^2}{2}\right) dZ - \\ &\quad \frac{1}{\sqrt{2\pi}} \int_A^{\infty} K \exp\left(-\frac{Z^2}{2}\right) dZ \\ &= \frac{1}{\sqrt{2\pi}} S_t^1 \exp(r(T-t)) \int_A^{\infty} \exp\left(-\left[\frac{Z^2}{2} - \sigma\sqrt{T-t}Z + \frac{\sigma^2}{2}(T-t)\right]\right) dZ - \\ &\quad K \frac{1}{\sqrt{2\pi}} \int_A^{\infty} \exp\left(-\frac{Z^2}{2}\right) dZ \\ &= S_t^1 \exp(r(T-t)) \frac{1}{\sqrt{2\pi}} \int_A^{\infty} \exp\left(-\frac{1}{2}\left[Z - \sigma\sqrt{T-t}\right]^2\right) dZ - K \times [1 - \Phi(A)] \\ &= S_t^1 \exp(r(T-t)) \frac{1}{\sqrt{2\pi}} \int_{A - \sigma\sqrt{T-t}}^{\infty} \exp\left(-\frac{1}{2}x^2\right) dx - K \times [1 - \Phi(A)] \\ &= S_t^1 \exp(r(T-t)) \left[1 - \Phi\left(A - \sigma\sqrt{T-t}\right)\right] - K \times [1 - \Phi(A)] \end{aligned} \quad (4.191)$$

Portanto o preço justo da opção de compra européia é dada por

$$\begin{aligned} c_t &= e^{-r(T-t)} E_Q [\max [(S_T^1 - K), 0] | \mathcal{F}_t] \\ &= e^{-r(T-t)} \left[S_t^1 \exp(r(T-t)) \left[1 - \Phi \left(A - \sigma \sqrt{T-t} \right) \right] - K \times [1 - \Phi(A)] \right] \end{aligned} \quad (4.192)$$

portanto

$$c_t = S_t^1 \left[1 - \Phi \left(A - \sigma \sqrt{T-t} \right) \right] - K e^{-r(T-t)} \times [1 - \Phi(A)] \quad (4.193)$$

onde $\Phi(\cdot)$ é a função densidade acumulada da normal de média 0 e variância 1. Em outras palavras é a probabilidade de que uma variável normal padrão seja menor que x . Esta é a conhecida fórmula de Black-Scholes, veja [Black & Scholes, 73].

O valor da opção de venda européia sai direto da relação de paridade entre opção de compra e venda, veja equação (2.62), ou notando que dado a condição final da opção de venda $p_t = \max(K - S_T, 0)$, veja equação (2.54), então

$$p_t = K e^{-r(T-t)} \times [1 - \Phi(A)] - S_t^1 \left[1 - \Phi \left(A - \sigma \sqrt{T-t} \right) \right] \quad (4.194)$$

4.3.1 A Abordagem de Black-Scholes

No artigo publicado por Black-Scholes, veja [Black & Scholes, 73], eles utilizaram uma abordagem que não utiliza as ferramentas do cálculo estocástico, que foi a preterida neste trabalho. Será feito nesta seção uma análise semelhante ao encontrado em [Black & Scholes, 73]. Esta análise é interessante pois utiliza-se de vários conceitos econômico-financeiros que nos leva a várias interpretações úteis e de onde surgem novas formas de se fazer hedge¹.

Black-Scholes partiu das premissas que o mercado é completo, livre de arbitragens, sem custos de transação e que a variação infinitesimal nos preços é uma difusão log-normal. Deste modo queremos montar uma carteira que duplique o fluxo de caixa de uma opção de compra. Uma carteira montada desta maneira não tem risco, pois sabemos o quanto ela valerá na data de exercício da opção T , seu valor será $\max(S_T - K, 0)$, veja seção 2.4.2.

Seja $(V_t)_{t \in [0, T]}$ o processo valor da carteira, composto de δ unidades de uma ação e de uma opção de compra desta ação. Deseja-se determinar a relação entre ação e opção, isto é, o número de ações δ para cada opção, que a carteira deve possuir de tal modo que o risco seja eliminado. Assim temos que

¹ Hedge é um termo utilizado na bolsa de valores que designa ação compensatória de um indivíduo, a fim de cobrir-se contra possíveis prejuízos em uma transação

$$V_t = \delta S_t - c_t \quad (4.195)$$

e portanto

$$dV_t = \delta dS_t - dc_t \quad (4.196)$$

a premissa é que $(S_t)_{t \in [0, T]}$ é uma difusão log-normal e que $c_t = f(S_t)$. Assim temos que

$$dS_t = \mu S_t dt + \sigma S_t dB_t \quad (4.197)$$

e pelo lema de Itô, temos que

$$dc_t = \left[\frac{\partial c}{\partial t} + \mu S \frac{\partial c}{\partial S} + \frac{\sigma^2 S^2}{2} \frac{\partial^2 c}{\partial S^2} \right] dt + \sigma S \frac{\partial c}{\partial S} dB_t \quad (4.198)$$

portanto

$$\begin{aligned} dV_t &= dS_t - \delta dc_t \\ &= [\mu S_t dt + \sigma S_t dB_t] \delta - \\ &\quad \left\{ \left[\frac{\partial c}{\partial t} + \mu S \frac{\partial c}{\partial S} + \frac{\sigma^2 S^2}{2} \frac{\partial^2 c}{\partial S^2} \right] dt + \sigma S \frac{\partial c}{\partial S} dB_t \right\} \\ &= \left[-\frac{\partial c}{\partial t} - \mu S \frac{\partial c}{\partial S} - \frac{\sigma^2 S^2}{2} \frac{\partial^2 c}{\partial S^2} + \mu S \delta \right] dt + \\ &\quad \left[\delta - \frac{\partial c}{\partial S} \right] \sigma S dB_t \end{aligned} \quad (4.199)$$

assim

$$dV_t = \left[-\frac{\partial c}{\partial t} - \frac{\sigma^2 S^2}{2} \frac{\partial^2 c}{\partial S^2} + \mu S_t \left(\delta - \frac{\partial c}{\partial S} \right) \right] dt + \left[\delta - \frac{\partial c}{\partial S} \right] \sigma S dB_t \quad (4.200)$$

note agora que a condição para que a carteira não tenha risco é que tenhamos

$$\delta = \frac{\partial c}{\partial S} \quad (4.201)$$

Este parâmetro é conhecido como o delta da opção. O delta de uma opção de ação é a razão de mudança no preço da opção em relação à mudança no preço da ação objeto. É a quantidade de ações que devemos ter para cada opção vendida, de modo a criar um hedge sem risco. A construção de um hedge sem risco é às vezes, denominada de hedge do delta.

Deste modo, substituindo $\delta = \frac{\partial c}{\partial S}$ teremos

$$dV_t = \left[-\frac{\partial c}{\partial t} - \frac{\sigma^2 S^2}{2} \frac{\partial^2 c}{\partial S^2} \right] dt \quad (4.202)$$

Em equilíbrio, a carteira hedgeada deve ter um retorno igual à taxa de juros livre de risco r , portanto

$$\begin{aligned} dV_t &= rV_t dt \\ &= r(\delta S_t - c_t) dt \\ &= \left[rS_t \frac{\partial c}{\partial S} - rc_t \right] dt \end{aligned} \quad (4.203)$$

portanto igualando as duas equações acima temos que

$$rS_t \frac{\partial c}{\partial S} - rc_t = -\frac{\partial c}{\partial t} - \frac{\sigma^2 S^2}{2} \frac{\partial^2 c}{\partial S^2} \quad (4.204)$$

e assim

$$\frac{\partial c}{\partial t} + rS \frac{\partial c}{\partial S} + \frac{\sigma^2 S^2}{2} \frac{\partial^2 c}{\partial S^2} - rc_t = 0 \quad (4.205)$$

com condição final c_T que no caso da opção de compra européia é $\max(S_T - K, 0)$. Onde T é a data da maturidade ou exercício da opção e K é o valor de exercício da opção.

Note agora que δ varia continuamente com o tempo, isto significa que para que tenhamos um hedge perfeito teremos que fazer ajustes na nossa carteira continuamente no tempo. Na prática infelizmente isto não é possível pois temos custos de transação que tornam inviável ajustes contínuos na carteira.

4.4 Análise dos Modelos Difusão Linear e Ornstein-Uhlenbeck

Modelo 1: Difusão Linear

O processo preço do ativo de risco S_t^1 tem a seguinte dinâmica

$$dS_t^1 = \mu dt + \sigma dB_t \quad (4.206)$$

e o processo preço descontado do ativo de risco é dado por

$$\tilde{S}_t^1 = \frac{S_t^1}{S_t^0} = e^{-rt} S_t^1 \quad (4.207)$$

portanto pelo lema de Itô, temos que

$$\begin{aligned}
d\tilde{S}_t^1 &= -re^{-rt}S_t^1dt + e^{-rt}dS_t^1 \\
&= -r\tilde{S}_t^1dt + e^{-rt}[\mu dt + \sigma dB_t] \\
&= \left(\mu e^{-rt} - r\tilde{S}_t^1\right)dt + \sigma e^{-rt}dB_t
\end{aligned} \tag{4.208}$$

Assim sobre $\mathbb{Q}$, temos que

$$B_t^* = \int_0^t \theta_s ds + B_t \tag{4.209}$$

é um processo de Wiener. Deste modo, sobre esta nova medida $\mathbb{Q}$, $\tilde{S}_t^1$ tem a seguinte dinâmica

$$\begin{aligned}
d\tilde{S}_t^1 &= \left(\mu e^{-rt} - r\tilde{S}_t^1\right)dt + \sigma e^{-rt}(dB_t^* - \theta_t dt) \\
&= \left(\mu e^{-rt} - r\tilde{S}_t^1 - \sigma e^{-rt}\theta_t\right)dt + \sigma e^{-rt}dB_t^*
\end{aligned} \tag{4.210}$$

desde que $\tilde{S}_t^1$ tem que ser um martingale local sobre $\mathbb{Q}$, devemos ter

$$\mu e^{-rt} - r\tilde{S}_t^1 - \sigma e^{-rt}\theta_t = 0, \quad t \in [0, T] \tag{4.211}$$

portanto da equação acima temos que

$$\theta_t = \frac{\mu e^{-rt} - r\tilde{S}_t^1}{\sigma e^{-rt}} = \frac{\mu - rS_t^1}{\sigma} \tag{4.212}$$

assim

$$d\tilde{S}_t^1 = \sigma e^{-rt}dB_t^* \tag{4.213}$$

e a dinâmica para S_t^1 sobre $\mathbb{Q}$ é dada por

$$\begin{aligned}
dS_t^1 &= re^{rt}\tilde{S}_t^1dt + e^{rt}d\tilde{S}_t^1 \\
&= rS_t^1dt + e^{rt}(\sigma e^{-rt}dB_t^*) \\
&= rS_t^1dt + \sigma dB_t^*
\end{aligned} \tag{4.214}$$

cuja solução é

$$S_T^1 = S_t^1 e^{r(T-t)} + \sigma \int_t^T e^{r(T-s)} dB_s^* \tag{4.215}$$

assim a densidade de transição sobre $\mathbb{Q}$, veja equação (5.261), é uma distribuição normal $N \left[S_t^1 e^{r(T-t)}, \frac{\sigma^2}{2r} (e^{2r(T-t)} - 1) \right]$. Deste modo da proposição (4.3), temos que a fórmula para uma opção é dada por

$$H_t = \beta_{T-t} E_{\mathbb{Q}} [H_T | \mathcal{F}_t] = e^{-r(T-t)} E_{\mathbb{Q}} [H_T | \mathcal{F}_t] \quad (4.216)$$

com condição final H_T . E pela fórmula de Feynman-Kac, veja teorema (D.4) no apêndice D, temos que $H(t, S)$ deve satisfazer a seguinte EDP

$$\frac{\partial H(t, S)}{\partial t} + rS \frac{\partial H(t, S)}{\partial S} + \frac{1}{2} \sigma^2 \frac{\partial^2 H(t, S)}{\partial S^2} - rH(t, S) = 0 \quad (4.217)$$

com condição final $H(T, S)$.

A opção de compra européia com a condição final $c_T = \max [(S_T^1 - K), 0]$, veja equação (2.50), é dada por

$$\begin{aligned} c_t &= e^{-r(T-t)} E_{\mathbb{Q}} [c_T | \mathcal{F}_t] \\ &= e^{-r(T-t)} E_{\mathbb{Q}} [\max [(S_T^1 - K), 0] | \mathcal{F}_t] \\ &= e^{-r(T-t)} \int_{-\infty}^{\infty} \max [(S_T^1 - K), 0] | \mathcal{F}_t \Psi(S_T^1) dS_T^1 \\ &= e^{-r(T-t)} \int_K^{\infty} (S_T^1 - K) \Psi(S_T^1) dS_T^1 \\ &= e^{-r(T-t)} \int_K^{\infty} S_T^1 \Psi(S_T^1) dS_T^1 - K \int_K^{\infty} \Psi(S_T^1) dS_T^1 \end{aligned}$$

como $S_T^1 \sim N \left[S_t^1 e^{r(T-t)}, \frac{\sigma^2}{2r} (e^{2r(T-t)} - 1) \right]$ então

$$S_T^1 - K = c_T \sim N \left[S_t^1 e^{r(T-t)} - K, \frac{\sigma^2}{2r} (e^{2r(T-t)} - 1) \right] \quad (4.218)$$

portanto

$$c_t = e^{-r(T-t)} \int_0^{\infty} c_T \Psi(c_T) dc_T \quad (4.219)$$

fazendo a seguinte mudança de variável

$$Z = \frac{c_T - S_t^1 e^{r(T-t)} + K}{\sqrt{\frac{\sigma^2}{2r} (e^{2r(T-t)} - 1)}} \sim N(0, 1) \quad (4.220)$$

deste modo

$$c_T = Z \sqrt{\frac{\sigma^2}{2r} (e^{2r(T-t)} - 1)} + S_t^1 e^{r(T-t)} - K \quad (4.221)$$

assim

$$\begin{aligned}
c_t &= e^{-r(T-t)} \int_0^\infty c_T \Psi(c_T) dc_T \\
&= e^{-r(T-t)} \int_A^\infty \left[Z \sqrt{\frac{\sigma^2}{2r}(e^{2r(T-t)} - 1)} + S_t^1 e^{r(T-t)} - K \right] \Psi(Z) dZ \quad (4.222)
\end{aligned}$$

$$= e^{-r(T-t)} \sqrt{\frac{\sigma^2}{2r}(e^{2r(T-t)} - 1)} \int_A^\infty Z \Psi(Z) dZ + \quad (4.223)$$

$$(S_t^1 - K e^{-r(T-t)}) \int_A^\infty \Psi(Z) dZ \quad (4.224)$$

onde

$$A = \frac{k - S_t^1 e^{r(T-t)}}{\sqrt{\frac{\sigma^2}{2r}(e^{2r(T-t)} - 1)}} \quad (4.225)$$

portanto o preço da opção de compra européia para este modelo é dado por

$$c_t = e^{-r(T-t)} \sqrt{\frac{\sigma^2}{2r}(e^{2r(T-t)} - 1)} \frac{\exp\left(-\frac{A^2}{2}\right)}{\sqrt{2\pi}} + (S_t^1 - K e^{-r(T-t)}) \times [1 - \Phi(A)] \quad (4.226)$$

A opção de venda como já foi mostrado sai da relação de paridade entre opção de compra e venda.

Modelo 2: Processo de Ornstein-Uhlenbeck

O processo preço do ativo de risco S_t^1 tem a seguinte dinâmica

$$dS_t^1 = \mu S_t^1 dt + \sigma dB_t$$

e o processo preço descontado do ativo de risco é dado por

$$\tilde{S}_t^1 = \frac{S_t^1}{S_0^1} = e^{-rt} S_t^1 \quad (4.227)$$

portanto pelo lema de Itô, temos que

$$\begin{aligned}
d\tilde{S}_t^1 &= -r e^{-rt} S_t^1 dt + e^{-rt} dS_t^1 \\
&= -r \tilde{S}_t^1 dt + e^{-rt} [\mu S_t^1 dt + \sigma dB_t] \\
&= (\mu - r) \tilde{S}_t^1 dt + e^{-rt} \sigma dB_t \quad (4.228)
\end{aligned}$$

Assim sobre $\mathbb{Q}$, temos que

$$B_t^* = \int_0^t \theta_s ds + B_t \quad (4.229)$$

é um processo de Wiener. Deste modo, sobre esta nova medida $\mathbb{Q}$, $\tilde{S}_t^1$ tem a seguinte dinâmica

$$\begin{aligned} d\tilde{S}_t^1 &= (\mu - r) \tilde{S}_t^1 dt + \sigma e^{-rt} (dB_t^* - \theta_t dt) \\ &= \left((\mu - r) \tilde{S}_t^1 - \sigma e^{-rt} \theta_t \right) dt + \sigma e^{-rt} dB_t^* \end{aligned} \quad (4.230)$$

desde que $\tilde{S}_t^1$ tem que ser um martingale local sobre $\mathbb{Q}$, devemos ter

$$(\mu - r) \tilde{S}_t^1 - \sigma e^{-rt} \theta_t = 0, \quad t \in [0, T] \quad (4.231)$$

portanto da equação acima temos que

$$\theta_t = \frac{(\mu - r) \tilde{S}_t^1}{\sigma e^{-rt}} = \frac{(\mu - r)}{\sigma} S_t^1 \quad (4.232)$$

assim

$$d\tilde{S}_t^1 = \sigma e^{-rt} dB_t^* \quad (4.233)$$

e a dinâmica para S_t^1 sobre $\mathbb{Q}$ é dada por

$$\begin{aligned} dS_t^1 &= r e^{rt} \tilde{S}_t^1 dt + e^{rt} d\tilde{S}_t^1 \\ &= r S_t^1 dt + e^{rt} (\sigma e^{-rt} dB_t^*) \\ &= r S_t^1 dt + \sigma dB_t^* \end{aligned} \quad (4.234)$$

cuja solução é

$$S_T^1 = S_t^1 e^{r(T-t)} + \sigma \int_t^T e^{r(T-s)} dB_s^* \quad (4.235)$$

assim a densidade de transição sobre $\mathbb{Q}$, veja equação (5.261), é uma distribuição normal $N \left[S_t^1 e^{r(T-t)}, \frac{\sigma^2}{2r} (e^{2r(T-t)} - 1) \right]$.

Note que esta é a mesma densidade para o modelo 1. Portanto o preço da opção de compra europeia para este modelo é a mesma do modelo 1 portanto temos que

$$c_t = e^{-r(T-t)} \sqrt{\frac{\sigma^2}{2r} (e^{2r(T-t)} - 1)} \frac{\exp \left(-\frac{A^2}{2} \right)}{\sqrt{2\pi}} + (S_t^1 - K e^{-r(T-t)}) \times [1 - \Phi(A)] \quad (4.236)$$

onde

$$A = \frac{K - S_t^1 e^{r(T-t)}}{\sqrt{\frac{\sigma^2}{2r} (e^{2r(T-t)} - 1)}} \quad (4.237)$$

Capítulo 5

Inferências dos Parâmetros de EDE Cujas Solução Forte é Conhecida

A análise Estatística dos Processos de Difusão tem recebido considerável atenção nas últimas décadas, isto se deve aos avanços da Moderna teoria das Finanças a partir de 1973, veja [Merton, 73a] e [Black & Scholes, 73], que tem como principal ferramenta de análise a teoria das Equações Diferenciais Estocásticas. Para uma visão geral do assunto, consulte [Merton, 92], [Hull, 97] ou [Karatzas, 96].

O problema da inferência em modelos contínuos pode ser estudado sobre duas hipóteses. A primeira formula o problema da inferência baseado na idéia que é possível observar completamente um processo X_t durante um intervalo de tempo $[0, T]$. A segunda baseia-se na idéia de que o processo é observado parcialmente.

Na prática, o processo é sempre observado parcialmente. Assim, estamos interessados neste segundo tipo de análise. A inferência em processos com observações parciais pode ser dividida em dois enfoques. Um é quando conhecemos a densidade de transição do processo, i.e, quando conhecemos a solução forte do processo, e a outro é quando a desconhecemos. Aqui vamos estudar a inferência em processos onde conhecemos a densidade de transição.

Vamos analisar 3 modelos a saber:

1. $dX_t = \mu dt + \sigma dB_t$ (Difusão Linear)
2. $dX_t = \mu X_t dt + \sigma dB_t$ (Processo de Ornstein-Uhlenbeck)
3. $dX_t = \mu X_t dt + \sigma X_t dB_t$ (Difusão log-normal)

Onde dB_t é o movimento browniano ou processo de Wiener.

O objetivo é fazer a inferências dos parâmetros destes modelos. Para isto vamos utilizar duas abordagens diferentes: a clássica que busca estimadores de máxima verossimilhança dos

parâmetros e utiliza a teoria assintótica para construir intervalos de confiança e a bayesiana que parte do princípio que os parâmetros são variáveis aleatórias cujo conhecimento sobre eles é traduzido em uma densidade a priori. Neste caso os estimadores bayesianos são obtidos da densidade a posteriori que é calculada através do teorema de Bayes que combina a densidade a priori e a função de verossimilhança.

Portanto este capítulo está dividido em duas partes. Na primeira parte vamos determinar os estimadores de máxima verossimilhança (EMV) e construir os seus respectivos intervalos de confiança baseados na distribuição assintótica do EMV. Na segunda parte vamos determinar os estimadores bayesianos e construir seus respectivos intervalos de credibilidade.

Nos apêndices A e B, estão os resultados que serão utilizados aqui.

5.1 Análise Clássica dos Modelos

Para cada um dos três modelos que vamos analisar, passamos por três etapas:

1. Determinar a solução forte da EDE;
2. Determinação da densidade de transição do processo utilizando a solução forte do processo através do método do jacobiano;
3. Construção da função de verossimilhança a partir da densidade de transição encontrada.

5.1.1 Modelo 1: $dX_t = \mu dt + \sigma dB_t$

Resolvendo esta EDE temos:

$$\begin{aligned} \int_{t_i}^{t_{i+1}} dX_t &= \int_{t_i}^{t_{i+1}} \mu dt + \int_{t_i}^{t_{i+1}} \sigma dB_t \\ X_{t_{i+1}} - X_{t_i} &= \mu(t_{i+1} - t_i) + \sigma(B_{t_{i+1}} - B_{t_i}) \\ X_{t_{i+1}} &= X_{t_i} + \mu(t_{i+1} - t_i) + \sigma(B_{t_{i+1}} - B_{t_i}) \end{aligned} \quad (5.238)$$

onde $B_{t_{i+1}} - B_{t_i} \sim N(0, t_{i+1} - t_i)$, portanto a densidade de transição do processo é uma distribuição normal com média $X_{t_i} + \mu(t_{i+1} - t_i)$ e variância $\sigma^2(t_{i+1} - t_i)$, ou seja,

$$X_{t_{i+1}} | X_{t_i} \sim N [X_{t_i} + \mu(t_{i+1} - t_i), \sigma^2(t_{i+1} - t_i)] \quad (5.239)$$

Neste trabalho são considerados apenas amostras observadas em intervalos fixos. Deste modo então $t_{i+1} - t_i = \Delta t$, assim a densidade de transição fica assim

$$P(X_{t_{i+1}} | X_{t_i}; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2\Delta t}} \exp\left(-\frac{[X_{t_{i+1}} - (X_{t_i} + \mu\Delta t)]^2}{2\sigma^2\Delta t}\right). \quad (5.240)$$

A função de verossimilhança é dado por:

$$\begin{aligned} P(\mathbf{X}; \mu, \sigma^2) &= \prod_{i=1}^{n-1} P(X_{t_{i+1}} | X_{t_i}; \mu, \sigma^2) \\ &= [2\pi\sigma^2\Delta t]^{\frac{n-1}{2}} \exp\left(-\frac{1}{2\sigma^2\Delta t} \sum [X_{t_{i+1}} - (X_{t_i} + \mu\Delta t)]^2\right) \end{aligned} \quad (5.241)$$

Fazendo a seguinte reparametrização:

$$\phi = \mu\Delta t \text{ e } \nu = \sigma^2\Delta t \quad (5.242)$$

temos portanto que o log da verossimilhança é dado por:

$$l(\Theta) = -\frac{n-1}{2} \log[2\pi\nu] - \frac{1}{2\nu} [\mathbf{Z} - \mathbf{Y}\phi]' [\mathbf{Z} - \mathbf{Y}\phi] \quad (5.243)$$

com $\Theta = (\phi, \nu)$ vetor de parâmetros, onde

$$\mathbf{Z} = \begin{bmatrix} X_2 - X_1 \\ X_3 - X_2 \\ \vdots \\ X_n - X_{n-1} \end{bmatrix}_{(n-1) \times 1} \quad \mathbf{Y} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}_{(n-1) \times 1} \quad (5.244)$$

Estimadores:

$$\frac{\partial l(\Theta)}{\partial \phi} = \frac{1}{\nu} \mathbf{Y}' [\mathbf{Z} - \mathbf{Y}\phi] \quad (5.245)$$

Assim o estimador de ϕ é dado por

$$\hat{\phi} = [\mathbf{Y}'\mathbf{Y}]^{-1} \mathbf{Y}'\mathbf{Z} \quad (5.246)$$

$$\frac{\partial l(\Theta)}{\partial \nu} = -\frac{n-1}{2\nu} + \frac{1}{2\nu^2} [\mathbf{Z} - \mathbf{Y}\phi]' [\mathbf{Z} - \mathbf{Y}\phi] \quad (5.247)$$

Assim o estimador de ν é dado por

$$\hat{\nu} = \frac{1}{(n-1)} [\mathbf{Z} - \mathbf{Y}\hat{\phi}]' [\mathbf{Z} - \mathbf{Y}\hat{\phi}] \quad (5.248)$$

A matriz de informação é dada por

$$\mathbf{J}(\phi, \nu) = \begin{bmatrix} \mathbf{Y}'\mathbf{Y} & \frac{1}{\nu^2} \mathbf{Y}' [\mathbf{Z} - \mathbf{Y}\phi] \\ \frac{1}{\nu^2} \mathbf{Y}' [\mathbf{Z} - \mathbf{Y}\phi] & \frac{1}{\nu^3} [\mathbf{Z} - \mathbf{Y}\phi]' [\mathbf{Z} - \mathbf{Y}\phi] - \frac{n-1}{2\nu^2} \end{bmatrix} \quad (5.249)$$

E a matriz de informação observada

$$\mathbf{J}(\hat{\phi}, \hat{\nu}) = \begin{bmatrix} \frac{\mathbf{Y}'\mathbf{Y}}{\hat{\nu}} & 0 \\ 0 & \frac{n-1}{2\hat{\nu}^2} \end{bmatrix} \quad (5.250)$$

Os estimadores para (μ, σ^2) , veja equação (A.8), são dados por

$$\hat{\mu} = \frac{\hat{\phi}}{\Delta t} \quad (5.251)$$

$$\hat{\sigma}^2 = \frac{\hat{\nu}}{\Delta t} \quad (5.252)$$

Portanto $\mathbf{J}(\hat{\mu}, \hat{\sigma}^2)$, veja equação (A.17), é dado por

$$\begin{aligned} \mathbf{J}(\hat{\mu}, \hat{\sigma}^2) &= \begin{pmatrix} \Delta t & 0 \\ 0 & \Delta t \end{pmatrix} \begin{pmatrix} \frac{\mathbf{Y}'\mathbf{Y}}{\hat{\sigma}^2 \Delta t} & 0 \\ 0 & \frac{n-1}{2\hat{\sigma}^4 \Delta t^2} \end{pmatrix} \begin{pmatrix} \Delta t & 0 \\ 0 & \Delta t \end{pmatrix} \\ \therefore \mathbf{J}(\hat{\mu}, \hat{\sigma}^2) &= \begin{pmatrix} \frac{\mathbf{Y}'\mathbf{Y}}{\hat{\sigma}^2} \Delta t & 0 \\ 0 & \frac{n-1}{2\hat{\sigma}^4} \end{pmatrix} \end{aligned} \quad (5.253)$$

Os intervalos de confiança para $\hat{\mu}$ e $\hat{\sigma}^2$, para uma cobertura $(1 - \alpha = 0.95)$ são dados por:

$$\left\{ \mu : \hat{\mu} - 1.96 \sqrt{\frac{\hat{\sigma}^2}{\Delta t} (\mathbf{Y}'\mathbf{Y})^{-1}} \leq \mu \leq \hat{\mu} + 1.96 \sqrt{\frac{\hat{\sigma}^2}{\Delta t} (\mathbf{Y}'\mathbf{Y})^{-1}} \right\} \quad (5.254)$$

$$\left\{ \sigma^2 : \hat{\sigma}^2 - 1.96 \sqrt{\frac{2\hat{\sigma}^4}{n-1}} \leq \sigma^2 \leq \hat{\sigma}^2 + 1.96 \sqrt{\frac{2\hat{\sigma}^4}{n-1}} \right\} \quad (5.255)$$

5.1.2 Modelo 2: $dX_t = \mu X_t dt + \sigma dB_t$

Resolvendo esta EDE temos:

$$dX_t - \mu X_t dt = \sigma dB_t \quad (5.256)$$

multiplicando ambos os lados por $e^{-\mu t}$, temos

$$e^{-\mu t} dX_t - \mu e^{-\mu t} X_t dt = \sigma e^{-\mu t} dB_t \quad (5.257)$$

note agora que

$$d(X_t e^{-\mu t}) = e^{-\mu t} dX_t - \mu e^{-\mu t} X_t dt \quad (5.258)$$

assim

$$d(X_t e^{-\mu t}) = \sigma e^{-\mu t} dB_t \quad (5.259)$$

integrando, temos

$$\begin{aligned}
\int_{t_i}^{t_{i+1}} d(X_t e^{-\mu t}) &= \sigma \int_{t_i}^{t_{i+1}} e^{-\mu t} dB_t \\
X_{t_{i+1}} e^{-\mu t_{i+1}} - X_{t_i} e^{-\mu t_i} &= \sigma \int_{t_i}^{t_{i+1}} e^{-\mu t} dB_t \\
X_{t_{i+1}} &= X_{t_i} e^{\mu(t_{i+1}-t_i)} + \sigma \int_{t_i}^{t_{i+1}} e^{\mu(t_{i+1}-t)} dB_t
\end{aligned} \quad (5.260)$$

Assim,

$$X_{t_{i+1}} | X_{t_i} \sim N \left[X_{t_i} e^{\mu \Delta t}, \frac{\sigma^2}{2\mu} (e^{2\mu \Delta t} - 1) \right] \quad (5.261)$$

Fazendo a seguinte reparametrização $\phi = e^{\mu \Delta t}$ e $\nu = \frac{\sigma^2}{2\mu} (e^{2\mu \Delta t} - 1)$, temos que o log da verossimilhança é dado por:

$$l(\Theta) = -\frac{n-1}{2} \log [2\pi\nu] - \frac{1}{2\nu} [\mathbf{Z} - \mathbf{Y}\phi]' [\mathbf{Z} - \mathbf{Y}\phi] \quad (5.262)$$

onde

$$\mathbf{Z} = \begin{bmatrix} X_2 \\ X_3 \\ \vdots \\ X_n \end{bmatrix}_{(n-1) \times 1} \quad \mathbf{Y} = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_{n-1} \end{bmatrix}_{(n-1) \times 1} \quad (5.263)$$

Portanto os estimadores têm a mesma forma do modelo anterior

$$\hat{\phi} = [\mathbf{Y}'\mathbf{Y}]^{-1} \mathbf{Y}'\mathbf{Z} \quad (5.264)$$

e

$$\hat{\nu} = \frac{1}{(n-1)} [\mathbf{Z} - \mathbf{Y}\hat{\phi}]' [\mathbf{Z} - \mathbf{Y}\hat{\phi}] \quad (5.265)$$

E a matriz de informação observada, igual ao do modelo anterior, dada por

$$\mathbf{J}(\hat{\phi}, \hat{\nu}) = \begin{bmatrix} \frac{\mathbf{Y}'\mathbf{Y}}{\hat{\nu}} & 0 \\ 0 & \frac{n-1}{2\hat{\nu}^2} \end{bmatrix} \quad (5.266)$$

Os estimadores para (μ, σ^2) , veja equação (A.8), são dados por

$$\hat{\mu} = \frac{1}{\Delta t} \ln(\hat{\phi}) = \frac{1}{\Delta t} \ln([\mathbf{Y}'\mathbf{Y}]^{-1} \mathbf{Y}'\mathbf{Z}) \quad (5.267)$$

$$\hat{\sigma}^2 = \frac{2\hat{\mu}\hat{\nu}}{e^{2\hat{\mu}\Delta t} - 1} = \frac{2\hat{\mu}}{(n-1)(e^{2\hat{\mu}\Delta t} - 1)} [\mathbf{Z} - \mathbf{Y}e^{\hat{\mu}\Delta t}]' [\mathbf{Z} - \mathbf{Y}e^{\hat{\mu}\Delta t}] \quad (5.268)$$

Portanto $\mathbf{J}(\hat{\mu}, \hat{\sigma}^2)$, veja equação (A.17), é dado por

$$\mathbf{J}(\hat{\mu}, \hat{\sigma}^2) = \mathbf{Q}' \mathbf{J}(\hat{\phi}, \hat{\nu}) \mathbf{Q} \quad (5.269)$$

onde

$$\mathbf{Q} = \begin{bmatrix} \Delta t e^{\hat{\mu} \Delta t} & 0 \\ \frac{1}{2} \hat{\sigma}^2 \frac{-e^{2\hat{\mu} \Delta t} + 1 + 2\Delta t e^{2\hat{\mu} \Delta t} \hat{\mu}}{\hat{\mu}^2} & \frac{1}{2\hat{\mu}} (e^{2\hat{\mu} \Delta t} - 1) \end{bmatrix} \quad (5.270)$$

Nota: Os intervalos de confiança para este modelo não são exibidos pois a matriz $\mathbf{J}^{-1}(\hat{\mu}, \hat{\sigma}^2)$ é muito difícil de ser determinada da maneira como foi feita no modelo anterior.

5.1.3 Modelo 3: $dX_t = \mu X_t dt + \sigma X_t dB_t$

Resolvendo esta EDE temos:

$$\frac{1}{X_t} dX_t = \mu dt + \sigma dB_t \quad (5.271)$$

integrando vem que

$$\begin{aligned} \int_{t_i}^{t_{i+1}} \frac{1}{X_t} dX_t &= \int_{t_i}^{t_{i+1}} \mu dt + \int_{t_i}^{t_{i+1}} \sigma dB_t \\ &= \mu (t_{i+1} - t_i) + \sigma (B_{t_{i+1}} - B_{t_i}) \end{aligned} \quad (5.272)$$

A integral do lado esquerdo é encontrada utilizando o lema de Itô, veja [Oksendal, 92], assim temos

$$\int_{t_i}^{t_{i+1}} \frac{1}{X_t} dX_t = [\ln(X_{t_{i+1}}) - \ln(X_{t_i})] - \frac{\sigma^2}{2} (t_{i+1} - t_i) \quad (5.273)$$

Deste modo temos que

$$\ln(X_{t_{i+1}}) = \ln(X_{t_i}) + (\mu - \frac{\sigma^2}{2}) \Delta t + \sigma (B_{t_{i+1}} - B_{t_i}) \quad (5.274)$$

Portanto a densidade de transição do processo é dada por

$$\ln(X_{t_{i+1}}) | \ln(X_{t_i}) \sim N \left[\ln(X_{t_i}) + (\mu - \frac{\sigma^2}{2}) \Delta t, \sigma^2 \Delta t \right] \quad (5.275)$$

Fazendo a seguinte reparametrização

$$\phi = \mu \Delta t - \sigma^2 \frac{\Delta t}{2} \text{ e } \nu = \sigma^2 \Delta t \quad (5.276)$$

temos que o log da verossimilhança é dado por:

$$l(\Theta) = -\frac{n-1}{2} \log[2\pi\nu] - \frac{1}{2\nu} [\mathbf{Z} - \mathbf{Y}\phi]' [\mathbf{Z} - \mathbf{Y}\phi] \quad (5.277)$$

onde

$$\mathbf{Z} = \begin{bmatrix} \ln(X_2) - \ln(X_1) \\ \ln(X_3) - \ln(X_2) \\ \vdots \\ \ln(X_n) - \ln(X_{n-1}) \end{bmatrix}_{(n-1) \times 1} \quad \mathbf{Y} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}_{(n-1) \times 1} \quad (5.278)$$

Estimadores:

$$\hat{\phi} = [\mathbf{Y}'\mathbf{Y}]^{-1}\mathbf{Y}'\mathbf{Z} \quad (5.279)$$

e

$$\hat{\nu} = \frac{1}{(n-1)}[\mathbf{Z} - \mathbf{Y}\hat{\phi}]'[\mathbf{Z} - \mathbf{Y}\hat{\phi}] \quad (5.280)$$

E a matriz de informação observada, dada por

$$\mathbf{J}(\hat{\phi}, \hat{\nu}) = \begin{bmatrix} \frac{\mathbf{Y}'\mathbf{Y}}{\hat{\nu}} & 0 \\ 0 & \frac{n-1}{2\hat{\nu}^2} \end{bmatrix} \quad (5.281)$$

Os estimadores para (μ, σ^2) , veja (A.8), são dados por

$$\hat{\sigma}^2 = \frac{\hat{\nu}}{\Delta t} \quad (5.282)$$

$$\hat{\mu} = \frac{\hat{\phi}}{\Delta t} + \frac{\hat{\sigma}^2}{2} \quad (5.283)$$

Portanto $\mathbf{J}(\hat{\mu}, \hat{\sigma}^2)$, veja (A.17), é dado por

$$\mathbf{J}(\hat{\mu}, \hat{\sigma}^2) = \mathbf{Q}'\mathbf{J}(\hat{\phi}, \hat{\nu})\mathbf{Q}$$

$$\mathbf{Q} = \begin{bmatrix} \Delta t & -\frac{\Delta t}{2} \\ 0 & \Delta t \end{bmatrix} \quad (5.284)$$

Assim

$$\mathbf{J}(\hat{\mu}, \hat{\sigma}^2) = \begin{bmatrix} \frac{\Delta t}{\hat{\sigma}^2}(\mathbf{Y}'\mathbf{Y}) & -\frac{\Delta t}{2\hat{\sigma}^2}(\mathbf{Y}'\mathbf{Y}) \\ -\frac{\Delta t}{2\hat{\sigma}^2}(\mathbf{Y}'\mathbf{Y}) & -\frac{\Delta t}{4\hat{\sigma}^2}(\mathbf{Y}'\mathbf{Y}) + \frac{(n-1)}{2\hat{\sigma}^4} \end{bmatrix} \quad (5.285)$$

Os intervalos de confiança para $\hat{\mu}$ e $\hat{\sigma}^2$, para uma cobertura $(1 - \alpha = 0.95)$ são dados por:

$$\left\{ \mu : \hat{\mu} - 1.96\sqrt{\frac{\hat{\sigma}^2}{\Delta t}(\mathbf{Y}'\mathbf{Y})^{-1}} \leq \mu \leq \hat{\mu} + 1.96\sqrt{\frac{\hat{\sigma}^2}{\Delta t}(\mathbf{Y}'\mathbf{Y})^{-1}} \right\} \quad (5.286)$$

$$\left\{ \sigma^2 : \hat{\sigma}^2 - 1.96\sqrt{\frac{2\hat{\sigma}^4}{n-1} - \frac{4\hat{\sigma}^2}{\Delta t}(\mathbf{Y}'\mathbf{Y})^{-1}} \leq \sigma^2 \leq \hat{\sigma}^2 + 1.96\sqrt{\frac{2\hat{\sigma}^4}{n-1} - \frac{4\hat{\sigma}^2}{\Delta t}(\mathbf{Y}'\mathbf{Y})^{-1}} \right\} \quad (5.287)$$

5.2 Análise Bayesiana dos Modelos

No apêndice B, colocamos um pequeno resumo sobre alguns pontos importantes da inferência bayesiana que foram utilizados nesta parte do trabalho.

A principal vantagem encontrada na utilização da inferência bayesiana é quando reparametrizamos o modelo. Como vimos na seção anterior dependendo da reparametrização utilizada, obter intervalos de confiança para as estimativas pode ser uma tarefa não muito simples. No caso bayesiano isto é simplificado utilizando simulação de monte carlo para gerar valores diretamente da densidade a posteriori dos parâmetros do modelo reparametrizado e utilizando a função inversa para obter a densidade a posteriori dos parâmetros originais do modelo.

Como utilizamos grandes série neste trabalho a priori escolhida foi a que utiliza a matriz de informação observada $J(\hat{\Theta})$

$$J(\hat{\Theta}) = \left[-\frac{\partial^2 \log(p(\mathbf{X}|\Theta = \hat{\Theta}))}{\partial \Theta^2} \right] \quad (5.288)$$

Deste modo a posteriori será dada por

$$p(\Theta|\mathbf{X}) \propto J(\hat{\Theta}) p(\mathbf{X}|\Theta) \quad (5.289)$$

onde $p(\mathbf{X}|\Theta)$ é a verossimilhança. As distribuições marginais para os parâmetros $\Theta = (\theta_1, \dots, \theta_k)$, de onde extrairemos as informações requeridas a respeito dos parâmetros são obtidas integrando-se a posteriori $p(\mathbf{X}|\Theta)$ com relação aos outros parâmetros

$$\begin{aligned} p(\theta_1|\mathbf{X}) &= \int \dots \int p(\Theta|\mathbf{X}) d\theta_2 \dots d\theta_k \\ &\vdots \\ p(\theta_k|\mathbf{X}) &= \int \dots \int p(\Theta|\mathbf{X}) d\theta_1 \dots d\theta_{k-1} \end{aligned} \quad (5.290)$$

A seguir é feita a análise dos modelos.

Modelo 1: $dX_t = \mu dt + \sigma dB_t$

A verossimilhança é dada por, veja seção 5.1:

$$\begin{aligned} p(\mathbf{X}|\Theta) &= [2\pi\nu]^{-\frac{n-1}{2}} \exp \left[-\frac{1}{2\nu} [\mathbf{Z} - \mathbf{Y}\phi]' [\mathbf{Z} - \mathbf{Y}\phi] \right] \\ &\propto (\nu)^{-\frac{n-1}{2}} \exp \left[-\frac{1}{2\nu} [\mathbf{Z} - \mathbf{Y}\phi]' [\mathbf{Z} - \mathbf{Y}\phi] \right] \end{aligned} \quad (5.291)$$

onde $\Theta = (\phi, \nu)$ vetor de parâmetros, e

$$\mathbf{Z} = \begin{bmatrix} X_2 - X_1 \\ X_3 - X_2 \\ \vdots \\ X_n - X_{n-1} \end{bmatrix}_{(n-1) \times 1} \quad \mathbf{Y} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}_{(n-1) \times 1} \quad (5.292)$$

Uma priori não informativa para os parâmetros (ϕ, ν) é dada por, veja (B.4)

$$p(\Theta) \propto \frac{1}{\nu} \quad (5.293)$$

Assim a posteriori é dada por, veja (B.7)

$$\begin{aligned} p(\Theta|\mathbf{X}) &\propto p(\Theta)p(\mathbf{X}|\Theta) \\ &\propto (\nu)^{-\frac{n-1}{2}-1} \exp \left[-\frac{1}{2\nu} [\mathbf{Z} - \mathbf{Y}\phi]' [\mathbf{Z} - \mathbf{Y}\phi] \right] \end{aligned} \quad (5.294)$$

Note que

$$[\mathbf{Z} - \mathbf{Y}\phi]' [\mathbf{Z} - \mathbf{Y}\phi] = (\mathbf{Z} - \hat{\mathbf{Z}})' (\mathbf{Z} - \hat{\mathbf{Z}}) + (\phi - \hat{\phi})' (\mathbf{Y}'\mathbf{Y}) (\phi - \hat{\phi}) \quad (5.295)$$

onde

$$\hat{\phi} = [\mathbf{Y}'\mathbf{Y}]^{-1} \mathbf{Y}'\mathbf{Z} \quad (5.296)$$

$$\hat{\mathbf{Z}} = \mathbf{Y}\hat{\phi} \quad (5.297)$$

assim

$$p(\Theta|\mathbf{X}) \propto (\nu)^{-\frac{n-1}{2}-1} \exp \left[-\frac{1}{2\nu} \left[(\mathbf{Z} - \hat{\mathbf{Z}})' (\mathbf{Z} - \hat{\mathbf{Z}}) + (\phi - \hat{\phi})' (\mathbf{Y}'\mathbf{Y}) (\phi - \hat{\phi}) \right] \right] \quad (5.298)$$

A distribuição marginal a posteriori para ν é dada por

$$\begin{aligned} p(\nu|\mathbf{X}) &= p(\Theta|\mathbf{X}) d\phi \\ &\propto \nu^{-\frac{n-1}{2}-1} \exp \left(-\frac{1}{2\nu} [(\mathbf{Z} - \hat{\mathbf{Z}})' (\mathbf{Z} - \hat{\mathbf{Z}})] \right) \times \\ &\quad \int \exp \left(-\frac{1}{2\nu} [(\phi - \hat{\phi})' (\mathbf{Y}'\mathbf{Y}) (\phi - \hat{\phi})] \right) d\phi \end{aligned} \quad (5.299)$$

e

$$\int \exp \left(-\frac{1}{2\nu} [(\phi - \hat{\phi})' (\mathbf{Y}'\mathbf{Y}) (\phi - \hat{\phi})] \right) d\phi = [2\pi\nu |\mathbf{Y}'\mathbf{Y}|^{-1}]^{\frac{1}{2}} \quad (5.300)$$

portanto

$$\begin{aligned}
p(\nu|\mathbf{X}) &\propto \nu^{-\frac{n-1}{2}-1} \exp\left(-\frac{1}{2\nu} [(\mathbf{Z}-\hat{\mathbf{Z}})'(\mathbf{Z}-\hat{\mathbf{Z}})]\right) \times [2\pi\nu |\mathbf{X}'\mathbf{X}|^{-1}]^{\frac{1}{2}} \\
&\propto \nu^{-(\frac{n}{2}-1+1)} \exp\left(-\frac{1}{2\nu} [(\mathbf{Z}-\hat{\mathbf{Z}})'(\mathbf{Z}-\hat{\mathbf{Z}})]\right)
\end{aligned} \tag{5.301}$$

assim

$$\nu \sim IG\left(\frac{n}{2}-1, \frac{(\mathbf{Z}-\hat{\mathbf{Z}})'(\mathbf{Z}-\hat{\mathbf{Z}})}{2}\right) \tag{5.302}$$

Onde IG é a distribuição gama inversa. O estimador bayesiano considerando uma função de perda quadrática é dado pelo valor esperado a posteriori, assim sendo, temos que

$$\hat{\nu}_b = E(\nu|\mathbf{X}) = \frac{(\mathbf{Z}-\hat{\mathbf{Z}})'(\mathbf{Z}-\hat{\mathbf{Z}})}{(n-4)} \tag{5.303}$$

e a variância do estimador $\hat{\nu}_b$ é dado por

$$var(\hat{\nu}_b) = var(\nu|\mathbf{X}) = \frac{(\mathbf{Z}-\hat{\mathbf{Z}})'(\mathbf{Z}-\hat{\mathbf{Z}})(\mathbf{Z}-\hat{\mathbf{Z}})'(\mathbf{Z}-\hat{\mathbf{Z}})}{0.25(n-4)^2(n-6)}$$

Para σ^2 temos que $\nu = \sigma^2 \Delta t$, assim a distribuição marginal para σ^2 é dada por:

$$\sigma^2 \sim IG\left(\frac{n}{2}-1, \frac{(\mathbf{Z}-\hat{\mathbf{Z}})'(\mathbf{Z}-\hat{\mathbf{Z}})}{2\Delta t}\right) \tag{5.304}$$

portanto

$$\hat{\sigma}_b^2 = E(\sigma^2|\mathbf{X}) = \frac{(\mathbf{Z}-\hat{\mathbf{Z}})'(\mathbf{Z}-\hat{\mathbf{Z}})}{(n-4)\Delta t} \tag{5.305}$$

A variância do estimador $\hat{\sigma}^2$ é dada por

$$var(\hat{\sigma}_b^2) = Var(\sigma^2|\Theta) = \frac{(\mathbf{Z}-\hat{\mathbf{Z}})'(\mathbf{Z}-\hat{\mathbf{Z}})(\mathbf{Z}-\hat{\mathbf{Z}})'(\mathbf{Z}-\hat{\mathbf{Z}})}{0.25(n-4)^2(n-6)\Delta t^2} \tag{5.306}$$

Para o intervalo de credibilidade de σ^2 temos que

$$p(a \leq \sigma^2 \leq b) = 1 - \alpha$$

$$p(\sigma^2 \leq a) = p\left(\frac{1}{\sigma^2} \geq \frac{1}{a}\right) = 1 - p\left(\frac{1}{\sigma^2} \leq \frac{1}{a}\right) = \frac{\alpha}{2}$$

$$p(\sigma^2 \leq b) = p\left(\frac{1}{\sigma^2} \geq \frac{1}{b}\right) = 1 - p\left(\frac{1}{\sigma^2} \leq \frac{1}{b}\right) = 1 - \frac{\alpha}{2} \tag{5.307}$$

assim $\frac{1}{a} = G_{1-\frac{\alpha}{2}}$ e $\frac{1}{b} = G_{\frac{\alpha}{2}}$, portanto, $a = (G_{1-\frac{\alpha}{2}})^{-1}$ e $b = (G_{\frac{\alpha}{2}})^{-1}$ onde G é a distribuição gama com parâmetros $G\left(\frac{n}{2}-1, \frac{(\mathbf{Z}-\hat{\mathbf{Z}})'(\mathbf{Z}-\hat{\mathbf{Z}})}{2\Delta t}\right)$. Assim o intervalo de credibilidade para σ^2 é dado por

$$\left\{ \sigma^2 : (G_{1-\frac{\alpha}{2}})^{-1} \leq \sigma^2 \leq (G_{\frac{\alpha}{2}})^{-1} \right\} \quad (5.308)$$

A distribuição marginal a posteriori para ϕ é dada por

$$\begin{aligned} p(\phi|\mathbf{X}) &= p(\Theta|\mathbf{X}) d\nu \\ &\propto \int \nu^{-\frac{n-1}{2}-1} \exp \left(-\frac{1}{2\nu} \left[(\mathbf{Z} - \hat{\mathbf{Z}})'(\mathbf{Z} - \hat{\mathbf{Z}}) + (\phi - \hat{\phi})'(\mathbf{Y}'\mathbf{Y})(\phi - \hat{\phi}) \right] \right) d\nu \\ &\propto \Gamma \left(\frac{n-1}{2} \right) \left[\frac{(\mathbf{Z} - \hat{\mathbf{Z}})'(\mathbf{Z} - \hat{\mathbf{Z}}) + (\phi - \hat{\phi})'(\mathbf{Y}'\mathbf{Y})(\phi - \hat{\phi})}{2} \right]^{-\frac{n-1}{2}} \\ &\propto \Gamma \left(\frac{n-2+1}{2} \right) \left[\frac{n-2}{n-2} \frac{(\mathbf{Z} - \hat{\mathbf{Z}})'(\mathbf{Z} - \hat{\mathbf{Z}})}{2} + \frac{(\phi - \hat{\phi})'(\mathbf{Y}'\mathbf{Y})(\phi - \hat{\phi})}{2} \right]^{-\frac{n-2+1}{2}} \\ &\propto \Gamma \left(\frac{n-2+1}{2} \right) \left[1 + \frac{(\phi - \hat{\phi})'V^{-1}(\phi - \hat{\phi})}{n-2} \right]^{-\frac{n-2+1}{2}} \end{aligned} \quad (5.309)$$

onde

$$V = (\mathbf{Y}'\mathbf{Y})^{-1} \left[\frac{(\mathbf{Z} - \hat{\mathbf{Z}})'(\mathbf{Z} - \hat{\mathbf{Z}})}{(n-2)} \right] \quad (5.310)$$

e $\hat{\phi}$ é o estimador clássico para ϕ , veja (5.246). Portanto

$$\phi \sim t_{n-2}(\hat{\phi}, V) \quad (5.311)$$

e o estimador para ϕ é dado por

$$\hat{\phi}_b = E(\phi|\Theta) = \hat{\phi} \quad (5.312)$$

note que o estimador bayesiano considerando priori não informativa coincide com o estimador clássico.

A variância de ϕ é dada por

$$\begin{aligned} Var(\hat{\phi}_b) &= Var(\phi_b|\mathbf{X}) = \frac{n-2}{n-4} V \\ &= \frac{(\mathbf{Y}'\mathbf{Y})^{-1} [(\mathbf{Z} - \hat{\mathbf{Z}})'(\mathbf{Z} - \hat{\mathbf{Z}})]}{n-4} \end{aligned} \quad (5.313)$$

Para μ temos que $\phi = \mu\Delta t$, assim a distribuição marginal para μ é dada por:

$$\mu \sim t_{n-2} \left(\hat{\mu}, \frac{V}{\Delta t^2} \right) \quad (5.314)$$

portanto

$$\hat{\mu}_b = E(\mu|\mathbf{X}) = \hat{\mu} \quad (5.315)$$

onde $\hat{\mu}$ é o estimador clássico para μ , veja (5.251).

A variância do estimador $\hat{\mu}_b$ é dada por

$$\text{var}(\hat{\mu}_b) = \text{var}(\mu|\mathbf{X}) = \frac{n-2}{n-4} \frac{V}{\Delta t^2} = \frac{(\mathbf{Y}'\mathbf{Y})^{-1} [(\mathbf{Z} - \hat{\mathbf{Z}})'(\mathbf{Z} - \hat{\mathbf{Z}})]}{(n-4) \Delta t^2} \quad (5.316)$$

Para o intervalo de credibilidade de μ temos que

$$p\left(-a \leq \frac{\mu - E(\mu|\mathbf{X})}{\sqrt{\text{var}(\mu|\mathbf{X})}} \leq a\right) = 1 - \alpha \quad (5.317)$$

$$p\left(E(\mu|\mathbf{X}) - a\sqrt{\text{var}(\mu|\mathbf{X})} \leq \mu \leq E(\mu|\mathbf{X}) + a\sqrt{\text{var}(\mu|\mathbf{X})}\right) \quad (5.318)$$

onde

$$\frac{\mu - E(\mu|\mathbf{X})}{\sqrt{\text{var}(\mu|\mathbf{X})}} \sim t_{n-2} \quad (5.319)$$

portanto

$$a = t_{n-2, 1-\frac{\alpha}{2}} \quad (5.320)$$

assim o intervalo de credibilidade para μ é dado por

$$\left\{ \mu : \hat{\mu}_b - t_{n-2, 1-\frac{\alpha}{2}} \sqrt{\frac{n-2}{n-4} \frac{V}{\Delta t^2}} \leq \mu \leq \hat{\mu}_b + t_{n-2, 1-\frac{\alpha}{2}} \sqrt{\frac{n-2}{n-4} \frac{V}{\Delta t^2}} \right\} \quad (5.321)$$

Nota: Os modelos 2 e 3 terão estas mesmas distribuições marginais a posteriori, pois após a reparametrização ficaram com a mesma forma do modelo 1.

Modelo 2: $dX_t = \mu X_t dt + \sigma dB_t$

A verossimilhança é dada por, veja seção 5.1:

$$\begin{aligned} p(\mathbf{X}|\Theta) &= [2\pi\nu]^{-\frac{n-1}{2}} \exp\left[-\frac{1}{2\nu}[\mathbf{Z} - \mathbf{Y}\phi]'[\mathbf{Z} - \mathbf{Y}\phi]\right] \\ &\propto (\nu)^{-\frac{n-1}{2}} \exp\left[-\frac{1}{2\nu}[\mathbf{Z} - \mathbf{Y}\phi]'[\mathbf{Z} - \mathbf{Y}\phi]\right] \end{aligned} \quad (5.322)$$

onde

$$\mathbf{Z} = \begin{bmatrix} X_2 \\ X_3 \\ \vdots \\ X_n \end{bmatrix}_{(n-1) \times 1} \quad \mathbf{Y} = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_{n-1} \end{bmatrix}_{(n-1) \times 1}$$

Reparametrização:

$$\phi = e^{\mu \Delta t} \quad (5.323)$$

$$\nu = \frac{\sigma^2}{2\mu} (e^{2\mu \Delta t} - 1) \quad (5.324)$$

Distribuições marginais a posteriori para ϕ e ν :

$$\phi \sim t_{n-2}(\hat{\phi}, V) \quad (5.325)$$

$$\nu \sim IG\left(\frac{n}{2} - 1, \frac{(\mathbf{Z} - \hat{\mathbf{Z}})'(\mathbf{Z} - \hat{\mathbf{Z}})}{2}\right) \quad (5.326)$$

onde $\hat{\phi}$ e V têm a mesma forma do modelo anterior. As distribuições marginais a posteriori para μ e σ^2 serão obtidas através de simulação. Assim temos:

1. Gere um valor para $\phi^{(i)}$ e $\nu^{(i)}$ de suas respectivas distribuições marginais a posteriori, equações (5.325) e (5.326);
2. Determine os valores de $\mu^{(i)}$ e $\sigma^{2(i)}$ pelas inversas das reparametrizações, assim temos

$$\mu^{(i)} = \frac{1}{\Delta t} \ln(\phi^{(i)})$$

$$\sigma^{2(i)} = \frac{2\mu^{(i)} \nu^{(i)}}{e^{2\mu^{(i)} \Delta t} - 1}$$

3. Repita os passos 1 e 2 até conseguir uma amostra de tamanho razoável. Estas serão as distribuições a posteriori de μ e σ^2 .

Modelo 3: $dX_t = \mu X_t dt + \sigma X_t dW_t$

A verossimilhança é dada por, veja seção 5.1:

$$p(\mathbf{X}|\Theta) = [2\pi\nu]^{-\frac{n-1}{2}} \exp\left[-\frac{1}{2\nu}[\mathbf{Z} - \mathbf{Y}\phi]'[\mathbf{Z} - \mathbf{Y}\phi]\right]$$

$$\propto (\nu)^{-\frac{n-1}{2}} \exp\left[-\frac{1}{2\nu}[\mathbf{Z} - \mathbf{Y}\phi]'[\mathbf{Z} - \mathbf{Y}\phi]\right] \quad (5.327)$$

onde

$$\mathbf{Z} = \begin{bmatrix} \ln(X_2) - \ln(X_1) \\ \ln(X_3) - \ln(X_2) \\ \vdots \\ \ln(X_n) - \ln(X_{n-1}) \end{bmatrix}_{(n-1) \times 1} \quad \mathbf{Y} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}_{(n-1) \times 1} \quad (5.328)$$

Reparametrização:

$$\phi = \mu \Delta t - \frac{\nu}{2} \text{ e } \nu = \sigma^2 \Delta t \quad (5.329)$$

Distribuições marginais a posteriori para ϕ e ν :

$$\phi \sim t_{n-2}(\hat{\phi}, V) \quad (5.330)$$

$$\nu \sim IG\left(\frac{n}{2} - 1, \frac{(\mathbf{Z} - \hat{\mathbf{Z}})'(\mathbf{Z} - \hat{\mathbf{Z}})}{2}\right) \quad (5.331)$$

onde $\hat{\phi}$ e V têm a mesma forma do modelo anterior. Para σ^2 temos que $\nu = \sigma^2 \Delta t$, assim a distribuição marginal para σ^2 é dada por:

$$\sigma^2 \sim IG\left(\frac{n}{2} - 1, \frac{(\mathbf{Z} - \hat{\mathbf{Z}})'(\mathbf{Z} - \hat{\mathbf{Z}})}{2\Delta t}\right) \quad (5.332)$$

portanto

$$\hat{\sigma}_b^2 = E(\sigma^2 | \mathbf{X}) = \frac{(\mathbf{Z} - \hat{\mathbf{Z}})'(\mathbf{Z} - \hat{\mathbf{Z}})}{(n-4)\Delta t} \quad (5.333)$$

Assim o intervalo de credibilidade para σ^2 é dado por

$$\left\{ \sigma^2 : (G_{1-\frac{\alpha}{2}})^{-1} \leq \sigma^2 \leq (G_{\frac{\alpha}{2}})^{-1} \right\} \quad (5.334)$$

A distribuição marginal a posteriori para μ será obtidas através de simulação, da mesma maneira que o modelo anterior. Assim temos:

1. Gere um valor para $\phi^{(i)}$ e $\nu^{(i)}$ de suas respectivas distribuições marginais a posteriori, equações (5.330) e (5.331);
2. Determine o valor de $\mu^{(i)}$ pela inversa de sua reparametrização, assim temos

$$\mu^{(i)} = \frac{1}{\Delta t} \left(\phi^{(i)} + \frac{\nu^{(i)}}{2} \right) \quad (5.335)$$

3. Repita os passos 1 e 2 até conseguir uma amostra de tamanho razoável.

Capítulo 6

Inferência nas Versões Discretizadas

Neste capítulo iremos fazer a inferência dos modelos discretos que como já foi dito anteriormente podem ser obtidos pelo esquema de discretização de Euler, veja [Kloeden & Platen, 95]. Portanto utilizaremos o esquema de Euler para estimar os parâmetros dos modelos discretos. Como nosso objetivo é apenas comparar os resultados do valor da opção a tempo discreto com o valor da opção a tempo contínuo iremos considerar nesta parte apenas os estimadores Clássicos. O motivo é que nesta forma discreta a verossimilhança dos modelos é bastante simples e como o tamanho das amostra é consideravelmente grande os estimadores Clássico e Bayesiano com priori não informativa terão valores similares.

Modelo 1: $X_{t_{i+1}} = X_{t_i} + \mu_{\Delta t} + \sigma_{\Delta t} Z$

Aproximação de Euler:

$$X_{t_{i+1}} = X_{t_i} + \mu \Delta t + \sigma (B_{t_{i+1}} - B_{t_i}) \quad (6.336)$$

Assim

$$X_{t_{i+1}} | X_{t_i} \sim N [X_{t_i} + \mu \Delta t, \sigma^2 \Delta t] \quad (6.337)$$

Nota: Lembrando sempre que $\mu_{\Delta t} = \mu \Delta t$ e $\sigma_{\Delta t}^2 = \sigma^2 \Delta t$.

Portanto o log da função de verossimilhança é dada por

$$l(\Theta) = -\frac{n-1}{2} \log[2\pi\sigma^2\Delta t] - \frac{1}{2\sigma^2\Delta t} [\mathbf{Z} - \mathbf{X}\Phi]' [\mathbf{Z} - \mathbf{X}\Phi] \quad (6.338)$$

onde

$$\mathbf{Z} = \begin{bmatrix} X_2 - X_1 \\ X_3 - X_2 \\ \vdots \\ X_n - X_{n-1} \end{bmatrix}_{(n-1) \times 1} \quad \mathbf{X} = \begin{bmatrix} \Delta t \\ \Delta t \\ \vdots \\ \Delta t \end{bmatrix}_{(n-1) \times 1} \quad \Phi = \mu \quad (6.339)$$

Estimadores:

$$\frac{\partial l(\Theta)}{\partial \Phi} = \frac{1}{\sigma^2 \Delta t} \mathbf{X}' [\mathbf{Z} - \mathbf{X}\Phi] \quad (6.340)$$

$$\hat{\Phi} = [\mathbf{X}'\mathbf{X}]^{-1} \mathbf{X}'\mathbf{Z} \quad (6.341)$$

$$\frac{\partial l(\Theta)}{\partial \sigma^2} = -\frac{n-1}{2\sigma^2} + \frac{1}{2\sigma^4 \Delta t} [\mathbf{Z} - \mathbf{X}\hat{\Phi}]' [\mathbf{Z} - \mathbf{X}\hat{\Phi}] \quad (6.342)$$

$$\hat{\sigma}^2 = \frac{1}{(n-1)\Delta t} [\mathbf{Z} - \mathbf{X}\hat{\Phi}]' [\mathbf{Z} - \mathbf{X}\hat{\Phi}] \quad (6.343)$$

A hessiana é dada por

$$\mathbf{I}(\Phi, \sigma^2) = \begin{bmatrix} -\frac{\mathbf{X}'\mathbf{X}}{\sigma^2 \Delta t} & -\frac{1}{\sigma^4 \Delta t} \mathbf{X}' [\mathbf{Z} - \mathbf{X}\hat{\Phi}] \\ -\frac{1}{\sigma^4 \Delta t} \mathbf{X}' [\mathbf{Z} - \mathbf{X}\hat{\Phi}] & \frac{n-1}{2\sigma^4} - \frac{1}{\sigma^6 \Delta t} [\mathbf{Z} - \mathbf{X}\hat{\Phi}]' [\mathbf{Z} - \mathbf{X}\hat{\Phi}] \end{bmatrix} \quad (6.344)$$

E a matriz de informação de fisher observada

$$\mathbf{J}(\hat{\Phi}, \hat{\sigma}^2) = \begin{bmatrix} \hat{\sigma}^2 \Delta t (\mathbf{X}'\mathbf{X})^{-1} & 0 \\ 0 & \frac{2\hat{\sigma}^4}{n-1} \end{bmatrix} \quad (6.345)$$

Note que a densidade de transição obtida pela aproximação de Euler coincide com a verdadeira densidade de transição, veja equação (5.239), portanto não faremos uso de outras metodologias para este modelo.

Modelo 2: $X_{t_{i+1}} = X_{t_i} + \mu_{\Delta t} X_{t_i} + \sigma_{\Delta t} Z$

Aproximação de Euler:

$$X_{t_{i+1}} = X_{t_i} + \mu X_{t_i} \Delta t + \sigma (B_{t_{i+1}} - B_{t_i}) \quad (6.346)$$

Assim

$$X_{t_{i+1}} | X_{t_i} \sim N [X_{t_i} + \mu X_{t_i} \Delta t, \sigma^2 \Delta t] \quad (6.347)$$

Portanto o log da função de verossimilhança é dada por

$$l(\Theta) = -\frac{n-1}{2} \log[2\pi\sigma^2 \Delta t] - \frac{1}{2\sigma^2 \Delta t} [\mathbf{Z} - \mathbf{X}\hat{\Phi}]' [\mathbf{Z} - \mathbf{X}\hat{\Phi}] \quad (6.348)$$

onde

$$\mathbf{Z} = \begin{bmatrix} X_2 - X_1 \\ X_3 - X_2 \\ \vdots \\ X_n - X_{n-1} \end{bmatrix}_{(n-1) \times 1} \quad \mathbf{X} = \begin{bmatrix} X_1 \Delta t \\ X_2 \Delta t \\ \vdots \\ X_{n-1} \Delta t \end{bmatrix}_{(n-1) \times 1} \quad \Phi = \mu \quad (6.349)$$

Estimadores: Aqui também temos os mesmos estimadores. Observe que, apesar de os estimadores terem a mesma forma matricial, as matrizes $\mathbf{X}$ e $\mathbf{Z}$ têm formas diferentes. Deste modo,

$$\hat{\Phi} = [\mathbf{X}'\mathbf{X}]^{-1}\mathbf{X}'\mathbf{Z} \quad (6.350)$$

$$\hat{\sigma}^2 = \frac{1}{(n-1)\Delta t} [\mathbf{Z} - \mathbf{X}\hat{\Phi}]' [\mathbf{Z} - \mathbf{X}\hat{\Phi}] \quad (6.351)$$

E a matriz de informação de fisher observada

$$\mathbf{J}(\hat{\Phi}, \hat{\sigma}^2) = \begin{bmatrix} \hat{\sigma}^2 \Delta t (\mathbf{X}'\mathbf{X})^{-1} & 0 \\ 0 & \frac{2\hat{\sigma}^4}{n-1} \end{bmatrix} \quad (6.352)$$

Modelo 3: $X_{t_{i+1}} = X_{t_i} + \mu_{\Delta t} X_{t_i} + \sigma_{\Delta t} X_{t_i} Z$

Aproximação de Euler:

$$X_{t_{i+1}} = X_{t_i} + \mu X_{t_i} \Delta t + \sigma X_{t_i} (B_{t_{i+1}} - B_{t_i}) \quad (6.353)$$

Assim

$$X_{t_{i+1}} | X_{t_i} \sim N [X_{t_i} + \mu X_{t_i} \Delta t, \sigma^2 X_{t_i}^2 \Delta t] \quad (6.354)$$

Portanto o log da função de verossimilhança é dada por

$$l(\Theta) = -\frac{n-1}{2} \log[2\pi\sigma^2\Delta t] - \mathbf{C}'\mathbf{1} - \frac{1}{2\sigma^2\Delta t} [\mathbf{Z} - \mathbf{X}\hat{\Phi}]' [\mathbf{Z} - \mathbf{X}\hat{\Phi}] \quad (6.355)$$

onde

$$\begin{aligned}
\mathbf{Z} &= \begin{bmatrix} \frac{X_2 - X_1}{X_1} \\ \frac{X_3 - X_2}{X_2} \\ \vdots \\ \frac{X_n - X_{n-1}}{X_{n-1}} \end{bmatrix}_{(n-1) \times 1} & \mathbf{X} &= \begin{bmatrix} \Delta t \\ \Delta t \\ \vdots \\ \Delta t \end{bmatrix}_{(n-1) \times 1} \\
\mathbf{C} &= \begin{bmatrix} \ln(X_1) \\ \ln(X_2) \\ \vdots \\ \ln(X_{n-1}) \end{bmatrix}_{(n-1) \times 1} & \mathbf{1} &= \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}_{(n-1) \times 1} & \Phi &= \mu
\end{aligned} \tag{6.356}$$

Estimadores: As mesmas formas dos modelos anteriores. Deste modo,

$$\hat{\Phi} = [\mathbf{X}'\mathbf{X}]^{-1}\mathbf{X}'\mathbf{Z} \tag{6.357}$$

$$\hat{\sigma}^2 = \frac{1}{(n-1)\Delta t} [\mathbf{Z} - \mathbf{X}\hat{\Phi}]' [\mathbf{Z} - \mathbf{X}\hat{\Phi}] \tag{6.358}$$

E a matriz de informação de fisher observada

$$\mathbf{J}(\hat{\Phi}, \hat{\sigma}^2) = \begin{bmatrix} \hat{\sigma}^2 \Delta t (\mathbf{X}'\mathbf{X})^{-1} & 0 \\ 0 & \frac{2\hat{\sigma}^4}{n-1} \end{bmatrix} \tag{6.359}$$

Capitulo 7

Resultados

7.1 Estimación dos Parâmetros dos Modelos Contínuos

7.1.1 Modelo 1: $dX_t = \mu dt + \sigma dB_t$

Amostra 1: $\mu_R = 0.2, \sigma_R^2 = 1$ e $\Delta t = 1$

Método	$\hat{\mu}$	std	$IC95\%$	$\hat{\sigma}^2$	std	$IC95\%$
Clássico	0.1769	0.0318	0.1145 - 0.2394	1.0131	0.0453	0.9243 - 1.1020
Bayesiano	0.1769	0.0319	0.1144 - 0.2395	1.0162	0.0456	0.9307 - 1.1093

Nota: IC no método Clássico significa intervalo de confiança, no método bayesiano intervalo de credibilidade.

Amostra 2: $\mu_R = 0.2, \sigma_R^2 = 1$ e $\Delta t = 0.1$

Método	$\hat{\mu}$	<i>std</i>	IC95%	$\hat{\sigma}^2$	<i>std</i>	IC95%
Clássico	0.1448	0.0997	-0.0507 +0.3403	0.9940	0.0445	0.9068 +1.0811
Bayesiano	0.1448	0.0999	-0.0512 +0.3408	0.9969	0.0447	0.9131 +1.0884

Amostra 3: $\mu_R = 0.2, \sigma_R^2 = 0.25$ e $\Delta t = 1$

Método	$\hat{\mu}$	Var	IC 95%	$\hat{\sigma}^2$	Var	IC 95%
Clássico	0.1885	0.0159	0.1573 +0.2197	0.2533	0.0113	0.2311 +0.2755
Bayesiano	0.1885	0.0159	0.1572 +0.2198	0.2540	0.0114	0.2327 +0.2773

Amostra 4: $\mu_R = 0.2, \sigma_R^2 = 0.25$ e $\Delta t = 0.1$

Método	$\hat{\mu}$	<i>std</i>	IC95%	$\hat{\sigma}^2$	<i>std</i>	IC95%
Clássico	0.1724	0.0499	0.0746 +0.2702	0.2485	0.0111	0.2267 +0.2703
Bayesiano	0.1724	0.0500	0.0744 +0.2704	0.2493	0.0112	0.2283 +0.2721

Amostra 5: $\mu_R = 0.2, \sigma_R^2 = 0.01$ e $\Delta t = 1$

Método	$\hat{\mu}$	<i>std</i>	IC95%	$\hat{\sigma}^2$	<i>std</i>	IC95%
Clássico	0.1977	0.0032	0.1914 +0.2040	0.0102	4.56×10^{-4}	0.0093 +0.0111
Bayesiano	0.1977	0.0032	0.1914 +0.2040	0.0102	4.58×10^{-4}	0.0094 +0.0112

Amostra 6: $\mu_R = 0.2, \sigma_R^2 = 0.01$ e $\Delta t = 0.1$

Método	$\hat{\mu}$	<i>std</i>	IC95%	$\hat{\sigma}^2$	<i>std</i>	IC95%
Clássico	0.1945	0.0100	0.1749 +0.2140	0.0099	4.45×10^{-4}	0.0091 +0.0108
Bayesiano	0.1945	0.0100	0.1749 +0.2141	0.0100	4.48×10^{-4}	0.0091 +0.0109

7.1.2 Modelo 2: $dX_t = \mu X_t dt + \sigma dB_t$

Amostra 1: $\mu_R = -0.5, \sigma_R^2 = 0.04$ e $\Delta t = 1$

Método	$\hat{\mu}$	std	$IC95\%$	$\hat{\sigma}^2$	std	$IC95\%$
Clássico	-0.4937	0.0411	-0.5742 - 0.4131	0.0400	0.0023	0.0356 - 0.0445
Bayesiano	-0.4944	0.0406	-0.5755 - 0.4168	0.0402	0.0022	0.0360 - 0.0448

Amostra 2: $\mu_R = -0.5, \sigma_R^2 = 0.04$ e $\Delta t = 0.1$

Método	$\hat{\mu}$	std	$IC95\%$	$\hat{\sigma}^2$	std	$IC95\%$
Clássico	-0.6528	0.1181	-0.8843 - 0.4213	0.0412	0.0019	0.0375 - 0.0450
Bayesiano	-0.6537	0.1166	-0.8847 - 0.4306	0.0414	0.0019	0.0378 - 0.0452

7.1.3 Modelo 3: $dX_t = \mu X_t dt + \sigma X_t dB_t$

Amostra 1: $\mu_R = 3.10^{-4}, \sigma_R^2 = 3.6.10^{-7}$ e $\Delta t = 1$

Método	$\hat{\mu}$	std	$IC95\%$	$\hat{\sigma}^2$	std	$IC95\%$
Clássico	2.8618	0.1910	2.4874 - 3.2362	3.6449	0.1631	3.3253 - 3.9646
Bayesiano	2.8624	0.1922	2.4852 - 3.2376	3.6559	0.1640	3.3485 - 3.9911

Amostra 2: $\mu_R = 3.10^{-4}, \sigma_R^2 = 3.6.10^{-7}$ e $\Delta t = 0.1$

Método	$\hat{\mu}$	std	$IC95\%$	$\hat{\sigma}^2$	std	$IC95\%$
Clássico	2.6685	0.6102	1.4724 - 3.8645	3.7202	0.1665	3.3939 - 4.0464
Bayesiano	2.6779	0.6141	1.4534 - 3.8834	3.7314	0.1674	3.4176 - 4.0735

7.2 Estudo do Comportamento Assintótico dos Estimadores Clássico e Bayesiano nos Modelos a Tempo Contínuo

O objetivo desta seção é comparar a performance dos estimadores clássico e bayesiano. Vamos considerar neste estudo duas medidas.

$$R(\hat{\Theta}, \Theta) = E[\hat{\Theta} - \Theta]^2 = \text{var}(\hat{\Theta}) - b^2(\hat{\Theta}, \Theta) \quad (7.360)$$

onde

$$b(\hat{\Theta}, \Theta) = E(\hat{\Theta}) - \Theta \quad (7.361)$$

é o vício de $\hat{\Theta}$ e $\hat{\Theta}$ é um estimador para Θ . A outra medida a ser considerada é a seguinte

$$EAP = \left[\frac{1}{M} \sum \left(\frac{\hat{\Theta}_i - \Theta}{\Theta} \right) \right] \times 100\% \quad (7.362)$$

Como este estudo é feito com amostras geradas conhecemos Θ . M é o número de amostras ou realizações do processo.

Este estudo foi realizado utilizando a simulação de Monte Carlo para obter o comportamento assintótico dos estimadores. Foram geradas mil realizações do mesmo processo com a mesma condição inicial. Assim uma estimativa para $R(\hat{\Theta}, \Theta)$ é dada por

$$\hat{R} = s^2 - \hat{b}^2 \quad (7.363)$$

onde

$$s^2 = \frac{1}{M} \sum_{i=1}^M (\hat{\Theta}_i - \bar{\hat{\Theta}})^2 \quad (7.364)$$

$$\bar{\hat{\Theta}} = \frac{1}{M} \sum_{i=1}^M \hat{\Theta}_i \quad (7.365)$$

e

$$\hat{b} = \bar{\hat{\Theta}} - \Theta \quad (7.366)$$

Além destas medidas de performance iremos determinar também

$$\bar{I} = \frac{1}{M} \sum_{i=1}^M |\limsup - \liminf| \quad (7.367)$$

onde $\bar{I}$ é o intervalo de confiança ou de credibilidade médio e

$$std = \sqrt{\frac{1}{M} \sum_{i=1}^M [(\limsup - \liminf) - \bar{I}]^2}$$

(7.368)

é o desvio padrão de $\bar{I}$. Onde $\liminf$ e $\limsup$ são os limites inferior e superior dos intervalos de confiança ou de credibilidade.

7.2.1 **Modelo 1:** $\mu_R = 0.35, \sigma_R^2 = 1$ e $\Delta t = 1$

7.3.2 Modelo 2: $\mu_R = -0.5, \sigma_R^2 = 0.04$ e $\Delta t = 1$

Parâmetro: μ							
Método	$\hat{\mu}$	s^2	$\hat{b}$	$\hat{R}$	EAP (%)	SI (95%)	std
Clássico	-0.5129	0.0108	-0.0129	0.0106	15.80	0.3943	0.0659
Bayesiano	-0.5185	0.0112	-0.0185	0.0108	16.12	0.3943	0.0659

Parâmetro: σ^2							
Método	$\hat{\sigma}^2$	$s^2 \times 10^{-5}$	$\hat{b} \times 10^{-4}$	$\hat{R} \times 10^{-5}$	EAP (%)	SI (95%)	std
Clássico	0.0403	2.70	2.54	2.69	10.32	0.0213	0.0035
Bayesiano	0.0412	2.88	12.00	2.73	10.77	0.0213	0.0035

7.3.3 Modelo 3: $\mu_R = 3 \times 10^{-4}, \sigma_R^2 = 3.6 \times 10^{-7} \text{ e } \Delta t = 1$

7.4 Valor das Opções de Venda Europeias

Modelo 1 - Amostra 1 $\mu_R = 0.02, \sigma_R = 0.02$

Modelo	$\hat{\mu}$	$\hat{\sigma}$	EQM	std
Discreto	0.0185	0.0203	0.0368	0.0043
Contínuo	0.0185	0.0203	0.0203	0.0024

Modelo 1 - Amostra 2 $\mu_R = 0.02, \sigma_R = 0.02$				
Modelo	$\hat{\mu}$	$\hat{\sigma}$	EQM	std
Discreto	0.0166	0.3032	0.1665	0.0165
Contínuo	0.0166	0.3032	0.0805	0.0103

Modelo 2 - Amostra 1 $\mu_R = 0.001, \sigma_R = 0.03$

Modelo	$\hat{\mu}$	$\hat{\sigma}$	EQM	std
Discreto	9.3486×10^{-4}	0.0306	0.0486	0.0038
Contínuo	9.3529×10^{-4}	0.0307	0.0458	0.0027

Modelo 2 - Amostra 2 $\mu_R = 0.001, \sigma_R = 0.1$

Modelo	$\hat{\mu}$	$\hat{\sigma}$	EQM	std
Discreto	7.8946×10^{-4}	0.0817	0.1295	0.0140
Contínuo	7.8978×10^{-4}	0.817	0.0695	0.0086

Modelo 3 - Amostra 1 $\mu_R = 0.001, \sigma_R = 0.002$

Modelo	$\hat{\mu}$	$\hat{\sigma}$	EQM	std
Discreto	9.2497×10^{-4}	0.0020	0.0675	0.0063
Contínuo	9.2454×10^{-4}	0.0020	0.0563	0.0051

Modelo 3 - Amostra 2 $\mu_R = 0.001, \sigma_R = 0.005$				
Modelo	$\hat{\mu}$	$\hat{\sigma}$	EQM	std
Discreto	8.1160×10^{-4}	0.0051	0.1536	0.0189
Contínuo	8.1193×10^{-4}	0.0051	0.1439	0.0180

Capítulo 8

Conclusão

No modelo 1 foram utilizadas seis amostras com o objetivo de avaliar a precisão das estimativas dos parâmetros quando variamos os valores de μ , σ^2 e Δt . Após realizadas as estimativas podemos concluir que:

1. Fixos μ e σ^2 a diminuição do intervalo entre as observações Δt , mantido é claro o mesmo número de observações, reduz a precisão da estimativa de μ , no entanto a precisão da estimativa de σ^2 permanece inalterada, compare as amostras 1 e 2, 3 e 4 e, 5 e 6. Com isto constatamos que:

$$\hat{\sigma}^2 \rightarrow \sigma^2 \text{ quando } n \rightarrow \infty \quad (8.369)$$

mas

$$\hat{\mu} \rightarrow \mu \text{ quando } T \rightarrow \infty \quad (8.370)$$

onde $T = (n - 1) \Delta t$.

2. Variando a relação $\frac{\mu}{\sigma}$ obtivemos um aumento na precisão da estimativa do drift quando aumentamos o valor da relação $\frac{\mu}{\sigma}$, compare as amostras 1, 3 e 5 e, 2, 4 e 6. Portanto para que tenhamos uma boa precisão na estimativa do drift é interessante ter-mos pelo $\frac{\mu}{\sigma} \geq 2$.

Na segunda etapa da análise do problema da inferência realizamos um estudo com o intuito de analisar a performance dos estimadores Clássico e Bayesiano. Quando iniciamos este trabalho tínhamos um sentimento que a abordagem bayesiana nos modelos onde se fez a reparametrização obteria melhores resultados, contudo como podemos verificar os resultados foram semelhantes. Talvez a única vantagem da metodologia bayesiana, nos modelos considerados aqui, é a facilidade de obtenção dos resultados. Provavelmente, se a amostra analisada for pequena e tivermos alguma informação a priori a abordagem bayesiana poderá apresentar melhores resultados, mas esta situação não foi considerada neste trabalho.

Na segunda parte deste trabalho é feita uma comparação dos resultados obtidos com os modelos contínuos com os modelos discretos na precificação de opções. Como pode-se notar os resultados obtidos são praticamente iguais. Isto se deve ao seguinte fato:

Calculando $E[S_T | \mathcal{F}_t]$ para os modelos 1 e 2 temos que:

- A tempo discreto $E_{\mathbb{Q}}[S_T | \mathcal{F}_t] = S_t^1 (1 + r)^{T-t}$
- A tempo contínuo $E_{\mathbb{Q}}[S_T | \mathcal{F}_t] = S_t^1 e^{r(T-t)}$

Ora como já foi dito anteriormente para $T - t$ grande $S_t^1 (1 + r)^{T-t} \approx S_t^1 e^{r(T-t)}$. Deste modo a pequena diferença foi devido a isto. Neste trabalho utilizou-se $T - t = 30$ dias, se houvessemos utilizado ao invés de 30 dias 60 dias a diferença seria menor. Portanto quanto maior for $T - t$ menor será a diferença entre os modelos 1 e 2 a tempo contínuo e a tempo discreto.

Para o modelo 3 temos que

- A tempo discreto $E_{\mathbb{Q}}[S_T | \mathcal{F}_t] = S_t^1 \exp[r(T-t)] \exp\left[\frac{\sigma_{\Delta t}^2 (T-t)^2}{2}\right]$. Como $\sigma_{\Delta t} = \sigma \sqrt{\Delta t}$ e como utilizamos $\Delta t = 1$ então neste caso $\sigma_{\Delta t} = \sigma$, portanto

$$E_{\mathbb{Q}}[S_T | \mathcal{F}_t] = S_t^1 \exp[r(T-t)] \exp\left[\frac{\sigma^2 (T-t)^2}{2}\right] \quad (8.371)$$

- A tempo contínuo $E_{\mathbb{Q}}[S_T | \mathcal{F}_t] = S_t^1 \exp[r(T-t)]$. Deste modo, a diferença entre os modelos contínuo e discreto está no fator $\exp\left[\frac{\sigma^2 (T-t)^2}{2}\right]$. Os valores utilizados para σ foram 0.0020 e 0.0050 e $T - t = 30$ assim temos que

$$\exp\left[\frac{0.0020^2 30^2}{2}\right] = 1.0018 \quad \text{e} \quad \exp\left[\frac{0.0050^2 30^2}{2}\right] = 1.0113 \quad (8.372)$$

Portanto, só teremos uma diferença significativa entre os modelos contínuo e discreto para valores de $\sigma \geq 0.015$ ou para $T - t \geq 90$, para estes valores a diferença será de 10% o que já é uma diferença significativa em favor dos modelos contínuos.

Capítulo 9

Propostas Futuras

Existem muitas possibilidades a seguir dentre elas gostaríamos de citar duas. A primeira é referente ao problema da precificação da opção do tipo americana. As segunda é referente a inferência em processos onde não podemos determinar a densidade de transição do processo.

9.1 Precificação da Opção do tipo Americana

Como já foi visto antes, nunca é ótimo exercer uma opção de compra americana antecipadamente, veja seção 2.5, portanto

$$(C_t)_{t \in [0, T]} = (c_t)_{t \in [0, T]}$$

Assim consideremos agora como avaliar a opção de venda americana. O ponto essencial quando pretendemos avaliar uma opção de venda americana é que devemos determinar a *Região de Parada*, onde é ótimo parar e exercer a opção antecipadamente, e a *Região de Continuação*, onde devemos proceder como a opção europeia e continuar. O ponto característico principal das opções americanas e a principal dificuldade em lidar com elas, é que em geral não é possível precificar uma opção sem que tenhamos ao mesmo tempo que encontrar a região de parada e a região de continuação. Além do mais, estas regiões não podem ser determinadas definitivamente por meio de fórmulas explícitas, mas sim por meio de um procedimento recursivo, caso a caso.

Matematicamente, a tarefa é descrever a região de continuação, isto é, a região onde o exercício antecipado não é ótimo. Este é o clássico *Problema da Parada Ótima*, como foi mencionado acima, a solução por fórmulas explícitas não é possível, na verdade, é provavel-

mente não praticável. Para um tratamento mais profundo deste assunto a tempo discreto, veja [Shiryayev, 95] e a tempo contínuo [Myneni, 92].

Uma das maneiras de solucionar o problema da Parada Ótima é utilizando envelope de Snell, veja [Snell, 52]. A outra maneira é relacionando a solução deste problema com a solução do problema de fronteira livre. Tal problema consiste de uma equação diferencial parcial com a condição de Dirichlet. Pode ser mostrado, veja [Moerbeke, 76], que na região de continuação

$$\frac{\partial H(t, x)}{\partial t} + a(x) \frac{\partial H(t, x)}{\partial x} + \frac{(b(x))^2}{2} \frac{\partial^2 H(t, x)}{\partial x^2} = 0 \quad (9.371)$$

9.2 Inferência em Processos Onde a Densidade de Transição é Desconhecida

Quando desconhecemos a densidade de transição, teremos que utilizar algum método para aproximar esta densidade. Uma metodologia possível é dada por [Yoshida, 92]. Em seu estudo ele utilizou a função de verossimilhança construída a partir de uma aproximação da derivada de Random-Nikodym que é obtida considerando que o processo pode ser observado continuamente sobre um intervalo $[0, T]$.

Outra maneira é utilizar as idéias propostas no artigo [Pedersen, 95a]. Neste trabalho, ele utilizou o esquema de Euler, veja [Kloeden & Platen, 95], para aproximar a verdadeira densidade de transição. A idéia é utilizar pra gerar pontos intermediários entre cada dois pontos observados utilizando para isso a densidade aproximada pelo esquema de Euler e a partir disto encontrar uma nova aproximação $p^{(N)}(t, X_t, t+1, X_{t+1}; \Theta)$ da verdadeira densidade de transição $p(s, x, t, y; \Theta)$, onde N é o número de estados intermediários que serão introduzidos entre X_t e X_{t+1} onde

$$p^{(N)}(t, X_t, t+1, X_{t+1}; \Theta) = E[X_{t+1} | X_t, N] \quad (9.372)$$

Em seu artigo, veja [Pedersen, 95a], ele provou que $\lim_{N \rightarrow \infty} p^{(N)}(t, X_t, t+1, X_{t+1}; \Theta) \rightarrow p(s, x, t, y; \Theta)$.

O problema da metodologia de Yoshida é que a precisão vai depender do tamanho do intervalo entre as observações Δt . Já a metodologia de Pedersen o problema é o número de trajetórias que deveremos simular até que consigamos uma precisão razoável e após isto ainda teremos o problema da maximização da função de verossimilhança que será o produto de uma somatória de funções o que acarretará sérios problema na hora de maximizar a função.

Uma possível saída para contornar estes problemas seria conseguir uma função fechada que aproximasse a verossimilhança. Um trabalho recente ainda a ser publicado [Ait-Sahalia] desenvolveu um método para solucionar este problema. A idéia é fazer transformações no modelo original de tal forma que o modelo resultante tenha uma densidade de transição o mais próximo da distribuição normal padrão. O motivo disto é que se pretende fazer uma expansão utilizando os polinômios de Hermite e como se sabe esta expansão é tanto melhor quanto mais próxima for a função da distribuição normal padrão. Portanto, tendo isto em mente foram feitas duas transformações na equação original para que fosse alcançado este objetivo. Estas transformações foram feitas com segue. Dada a equação diferencial estocástica

$$dX_t = a(X_t; \Theta) dt + b(X_t; \Theta) dB_t \quad (9.373)$$

O primeiro passo na construção da solução do problema consiste em fazer a seguinte transformação de $X_t \rightarrow Y_t$, da seguinte maneira

$$Y_t = \gamma(X_t, \Theta) = \int^{X_t} b^{-1}(u; \Theta) du \quad (9.374)$$

Então aplicando o lema de Itô, veja D.2, temos que

$$dY_t = a_Y(Y_t; \Theta) dt + dB_t \quad (9.375)$$

onde

$$a_Y(Y_t; \Theta) = \frac{a[\gamma^{-1}(Y_t; \Theta); \Theta]}{b[\gamma^{-1}(Y_t; \Theta); \Theta]} - \frac{1}{2} \frac{\partial b[\gamma^{-1}(Y_t; \Theta); \Theta]}{\partial X} \quad (9.376)$$

O segundo passo é fazer a seguinte transformação

$$Z_t = Z(\Delta t, y | y_0; \Theta) = (\Delta t)^{-\frac{1}{2}} [y - y_0 - a_Y(y_0; \Theta) \Delta t] \quad (9.377)$$

onde $\Delta t = t - t_0$ e Z_t é chamado de incremento "pseudo-normalizado" de Y_t .

Portanto Z_t como definido em (9.377) tem uma densidade de transição suficientemente próxima da densidade normal padrão. Assim podemos fazer uma série de expansões em torno da função densidade normal utilizando os polinômios de Hermite. Denotando por $P_Y(\Delta t, y | y_0; \Theta)$ a densidade condicional de $Y_{t+\Delta t} | Y_t$, temos que a densidade $P_Z(\Delta t, z | y_0; \Theta)$ é dada por

$$P_Z(\Delta t, z | y_0; \Theta) = (\Delta t)^{\frac{1}{2}} P_Y\left(\Delta t, (\Delta t)^{\frac{1}{2}} z + y_0 + a_Y(y_0; \Theta) \Delta t | y_0; \Theta\right) \quad (9.378)$$

Uma vez feita uma sequência de aproximações para $P_Z(\Delta t, z | y_0; \Theta)$ utilizando os polinômios de Hermite, podemos determinar a sequência de aproximações para $P_Y(\Delta t, y | y_0; \Theta)$

invertendo 9.378, assim temos

$$P_Y (\Delta t, y | y_0; \Theta) = (\Delta t)^{-\frac{1}{2}} P_Z \left(\Delta t, (\Delta t)^{-\frac{1}{2}} [y - y_0 - a_Y (y_0; \Theta) \Delta t] \middle| y_0; \Theta \right) \quad (9.379)$$

Deste modo podemos determinar a densidade de interesse $P_X (\Delta t, x | x_0; \Theta)$, aplicando a formula do jacobiano, temos que

$$P_X (\Delta t, x | x_0; \Theta) = b^{-1} (x; \Theta) P_Y (\Delta t, \gamma (x; \Theta) | \gamma (x_0; \Theta); \Theta) \quad (9.380)$$

As condições para que estas transformações possam ser feitas bem como a prova da convergência destas aproximações estão no artigo citado de Aït-Sahalia. Deste modo temos uma expressão fechada para o estimador de máxima verossimilhança.

Bibliografia

- [Aït-Sahalia] AÏT-SAHALIA, Y. Maximum likelihood estimation of discretely sampled diffusions, *Econometrica*, (no prelo).
- [Billingsley,79] BILLINGSLEY, P. *Probability and measure*. New York: John Wiley & Sons, 1979.
- [Bingham & Kiesel,98] BINGHAM, N. H.; KIESEL, R. *Risk-neutral valuation: pricing and hedging of financial derivatives*. London: Springer Verlag, 1998.
- [Black & Scholes, 73] BLACK, F.; SCHOLES, M. The pricing of options and corporate liabilities. *Journal of Political Economy*, v.81, n.3, p.163 – 175, 1973.
- [Bodie & Merton, 99] BODIE, Z.; MERTON, R. *Finanças*. Porto Alegre: Bookman, 1999.
- [Box & Tiao, 73] BOX, G. E. P.; TIAO, G. C. *Bayesian inference in statistical analysis*. Massachusetts: Addison-Wesley, 1973.
- [Cameron & Martin, 44] CAMERON, R. H.; MARTIN, W. T. Transformations of Wiener integrals under translations. *Annals of Mathematics*, v.45, n.2, p.386 – 396, 1944.
- [Cameron & Martin, 45] CAMERON, R. H.; MARTIN, W. T. Transformations of Wiener integrals by a nonlinear transformations. *Trans. Amer. Math. Soc.*, v.66, p.253 – 283, 1949.
- [Cameron & Martin, 49] CAMERON, R. H.; MARTIN, W. T. The transformations of Wiener integrals under a general class of linear transformations, *Trans. Amer. Math. Soc.*, v.48, p.184 – 219, 1945.
- [Casella & Berger, 90] CASELA, G.; BERGER, R. *Statistical inference*. Belmont: Duxbury, 1990.
- [Cox & Ross, 76] COX, J. C.; ROSS, S. A. The valuation of options for alternative stochastic processes. *Journal of Financial Economics*, v.3, p.145 – 166, 1976.
- [Cox et al, 79] COX, J. C.; ROSS, S. A.; RUBINSTEIN, M. Option pricing: a simplified approach. *Journal of Financial Economics*, v.7, p.229 – 263, 1979.
- [Doob, 53] DOOB, J. L. *Stochastic processes*. New York: John Wiles & Sons, 1953.
- [Doob, 94] DOOB, J. L. *Measure theory*. New York: Springer Verlag, 1994.
- [Elliott & Kopp, 99] ELLIOTT, R. J.; KOPP, P. E. *Mathematics of financial markets*. New York: Springer Verlag, 1999.
- [Friedman, 75] FRIEDMAN, A. *Stochastic differential equations and applications*. New York: Academic Press, 1975. (v.1)

- [Geisser & Eddy, 79] GEISSER, S.; EDDY, W. F. A predictive approach to model selection. *Journal of American Statistical Association*, v.74, p.153 – 160, 1979.
- [Gelman et al, 95] GELMAN, A.; CARLIN, J. B.; STERN, H. S.; RUBIN, D. B. *Bayesian data analysis*. London: Chapman & Hall, 1995.
- [Gihman & Skorohod, 72] GIHMAN, I. I.; SKOROHOD, A. V. *Stochastic differential equations*. Berlin: Springer Verlag, 1972.
- [Gillespie, 92] GILLESPIE, D. T. *Markov processes: an introduction for physical scientists*. San Diego: Academic Press, 1992.
- [Girsanov, 60] GIRSANOV, I. V. On transforming a certain class of stochastic processes by absolutely continuous substitution of measures. *Theory of Probability and Its Applications*, v.5, n. 3, p.285 – 301, 1960.
- [Halmos, 88] HALMOS, P. *Measure theory*. 4. ed. New York: Springer Verlag, 1988.
- [Harrison & Kreps, 79] HARRISON, J. M.; KREPS, D. M. Martingales and arbitrage in multiperiod securities markets. *Journal of Econom. Theory*, v.20, p.381 – 408, 1979.
- [Hull, 96] HULL, J. *Introdução aos mercados futuros e de opções*. 2. ed. São Paulo: Bolsa de Mercadorias & Futuros, 1996.
- [Hull, 97] HULL, J. *Options, futures, and other derivatives*. 3. ed. London: Prentice Hall, 1997.
- [Ikeda & Watanabe, 81] IKEDA, N.; WATANABE, S. *Stochastic differential equations and diffusion processes*. New York: Elsevier North-Holland, 1981.
- [Itô, 44] ITÔ, K. Stochastic integral. *Proc. Imp. Acad. Tokyo*, v.20, p.519 – 524, 1944.
- [Itô, 46] ITÔ, K. On a stochastic integral equation. *Proc. Imp. Acad. Tokyo*, v.22, p.519 – 524, 1946.
- [Jacob & Shiriyayev, 87] JACOB, J.; SHIRYAYEV, A. N. *Limit theorems for stochastic processes*. New York: Springer Verlag, 1987.
- [Johnson, 67] JOHNSON, R. A. An asymptotic expansion for posterior distributions. *Ann. Math. Statist.*, v.38, p.1899 – 1906, 1967.
- [Johnson, 70] JOHNSON, R. A. Asymptotic expansions associated with posterior distributions. *Ann. Math. Statist.*, v.41, p.851 – 864, 1970.
- [Jorion, 98] JORION, P. *Value at risk*. São Paulo: Bolsa de Mercadorias & Futuros, 1998.
- [Kalbfleisch, 85] KALBFLEISCH, J. G. *Probability and statistical inference*. 2. ed. New York: Springer-Verlag, 1985. (v.2)
- [Karatzas & Shreve, 91] KARATZAS, I.; SHREVE, S. E. *Brownian motion and stochastic calculus*. 2. ed. New York: Springer Verlag, 1991.
- [Karatzas, 96] KARATZAS, I. *Lecture on the mathematics of finance*. Providency: American Mathematical Society, 1996.
- [Kloeden & Platen, 95] KLOEDEN, P. E.; PLATEN, E. *Numerical solution of stochastic differential equation*. 2. ed. New York: Springer-Verlag, 1995.
- [Mel'nikov, 99] MEL'NIKOV, A. V. *Financial markets: stochastic analysis and the pricing of derivative securities*. Providency: American Mathematical Society, 1999.
- [Merton, 69] MERTON, T. Lifetime portfolio selection under uncertainty: the continuous time case. *Review of Economics and Statistics*, v.51, n.247 – 257, 1969.
- [Merton, 71] MERTON, T. Optimum consumption and portfolio rules in a continuous time model. *Journal of Economics Theory*, v.3, p.373 – 413, 1971.

- [Merton, 73a] MERTON, T. Theory of rational option pricing. *Bell Journal of Economics*, v.4, p.141 – 183, 1973.
- [Merton, 73b] MERTON, T. An intertemporal capital asset pricing model. *Econometrica*, v.41, p.8678 – 8687, 1973.
- [Merton, 92] MERTON, T. *Continuous-time finance*. 2. ed. Cambridge: Blackwell Publishers, 1992.
- [Myneni, 92] MYNENI, R. The pricing of american option. *Ann. Appl. Probab.*, v.2, p.1 – 23, 1992.
- [Moerbeke, 76] MOERBEKE, P. On optimal stopping and free boundary problem. *Arch. Rational Mech. Anal.*, v.60, p.101 – 148, 1976.
- [Neveu, 79] NEVEU, J. *Discrete-parameter martingales*. Amsterdam: North-Holland, 1975.
- [Oksendal, 92] OKSENDAL, B. *Stochastic differential equations*. an introduction with applications. 3. ed. New York: Springer Verlag, 1992.
- [Pedersen, 95a] PEDERSEN, A. R. A new approach to maximum likelihood estimation for stochastic differential equations based on discrete observations. *Scandinavian Journal of Statistics*, v.22, p.55 – 71, 1995.
- [Shiryayev, 95] SHIRYAYEV, A. N. Towards the theory of pricing of options of both european and american types: discrete time. *Theory Probab. Appl.*, v.39, n.1, p.14 – 60, 1995.
- [Snell, 52] SNELL, J. L. Applications of martingales systems theorems. *Trans. Amer. Math. Soc.*, v.73, p.293 – 312, 1952.
- [Wong, 71] WONG, E. *Stochastic processes in information and dynamical systems*. New York: McGraw-Hill, 1971.
- [Yeh, 73] YEH, J. *Stochastic processes and the wiener integral*. New York: Marcel Dekker, 1973.
- [Yoshida, 92] YOSHIDA, N. Estimation for diffusion processes from discrete observation. *Journal of Multivariate Analysis*, v.41, p.220 – 242, 1992.

Apendice A

Abordagem Clássica

Faremos aqui um pequeno resumo de alguns resultados que foram utilizados. Para maiores detalhes ver por exemplo [Casela & Berger, 90], [Kalbfleisch, 85].

A.1 Reparametrização

Neste trabalho foram feitas duas reparametrizações, uma no modelo 2 e outra no modelo 3. O objetivo destas reparametrizações são dois:

1. Simplificar a estimação dos parâmetros;
2. Determinar os intervalos de confiança assintóticos dos estimadores.

Vamos fazer um breve comentário sobre dois resultados que utilizamos quando trabalhamos com um modelo reparametrizado.

A.1.1 Invariância

Suponha que o modelo probabilístico para um experimento depende de um parâmetro desconhecido θ . O modelo então consiste de toda uma família de distribuição probabilidade, uma para cada valor de θ no espaço de parâmetros Ω , ou seja, para cada valor de θ pertencendo à Ω , temos uma distribuição teórica. Todavia, poderíamos parametrizar esta mesma família de distribuições de muitas outras formas. Por exemplo, ao invés de parametrizarmos por θ poderíamos utilizar λ , onde $\lambda = f(\theta)$. Para deixar mais claro o que estamos falando consideremos a seguinte função densidade de probabilidade

$$f(x) = \frac{1}{\theta} e^{-\frac{x}{\theta}}, \quad 0 < x < \infty \quad (\text{A.1})$$

Assim a distribuição parametrizada por $\theta = 1$ é

$$f(x) = e^{-x}, \quad 0 < x < \infty \quad (\text{A.2})$$

e a distribuição parametrizada por $\theta = 2$ é

$$f(x) = \frac{1}{2}e^{-\frac{x}{2}} \quad 0 < x < \infty \quad (\text{A.3})$$

Agora observe que podemos parametrizar a família de distribuições de (A.1) usando λ ao invés de θ

$$f(x) = \lambda e^{-\lambda x} \quad 0 < x < \infty \quad (\text{A.4})$$

onde $\lambda = \frac{1}{\theta}$. Assim, agora as distribuições (A.2) e (A.3) são parametrizadas por $\lambda = 1$ e $\lambda = 0.5$. Portanto, temos a escolha de parametrizar uma família de distribuições por θ ou por λ ou por qualquer outra função um-a-um de θ . Usualmente tentamos selecionar uma parametrização tal que o parâmetro representa alguma característica interessante da distribuição e a expressão matemática seja razoavelmente simples.

Assim, por exemplo, quando dizemos que o estimador de máxima verossimilhança de θ é $\hat{\theta} = 2$, implica que a distribuição parametrizada por $\theta = 2$ é a mais provável distribuição e isto significará também que a distribuição parametrizada por $\lambda = 0.5$ será a distribuição mais provável. Desde que o método para parametrizar as distribuições é muito arbitrário, seria desejável que as plausibilidades atribuídas a uma distribuição não dependesse de um particular método de parametrização que tenha sido escolhido. Em outras palavras, as plausibilidades atribuídas deveriam ser invariantes sobre transformações um-a-um dos parâmetros.

Uma importante propriedade dos métodos de verossimilhança é que eles são invariantes sobre transformações um-a-um, para mais detalhes sobre o método da verossimilhança e transformações de parâmetros veja [Kalbfleisch, 85]. Assim, suponha que $\theta = g(\lambda)$, onde g é inversível, e seja $\mathbf{P}(E; \theta)$ a probabilidade de observar o evento E . Substituindo $\theta = g(\lambda)$ nesta expressão nos fornece a probabilidade de E como uma função de λ . Isto segue que $\mathbf{L}(\theta)$, a função de verossimilhança de θ , e $\mathbf{L}_*(\lambda)$, a função de verossimilhança de λ , são relacionados como segue:

$$\mathbf{L}_*(\lambda) = \mathbf{L}(\theta) \quad (\text{A.5})$$

Consequentemente ambas as funções têm o mesmo valor máximo, e

$$\mathbf{RL}_*(\lambda) = \mathbf{RL}(\theta) \quad (\text{A.6})$$

onde $\mathbf{R}$ é a função de verossimilhança relativa

$$RL(\Theta) = \frac{L(\Theta)}{\sup L(\Theta)} \quad (\text{A.7})$$

Se λ_1 é algum possível valor de λ e $\theta_1 = g(\lambda_1)$ é o valor correspondente de θ , então λ_1 e θ_1 tem a mesma verossimilhança relativa. Portanto a verossimilhança relativa não depende da escolha para trabalhar com o parâmetro θ ou λ .

Isto segue que, se $\hat{\lambda}$ é o EMV de λ , então

$$\hat{\theta} = g(\hat{\lambda}) \quad (\text{A.8})$$

é o EMV de θ .

A.1.2 Relação entre a Função de Informação de θ e λ

Suponha que fizemos uma reparametrização de θ para λ via uma transformação um-a-um $\theta = g(\lambda)$. Pela propriedade da invariância, veja seção (A.1.1), o log da verossimilhança de λ é

$$l_*(\lambda) = l(\theta) \quad (\text{A.9})$$

A função escore de λ é

$$S_*(\lambda) = \frac{dl_*(\lambda)}{d\lambda} = \frac{dl(\theta)}{d\theta} \frac{d\theta}{d\lambda} = S(\theta) \frac{d\theta}{d\lambda} \quad (\text{A.10})$$

A função informação de λ é

$$\begin{aligned} J(\lambda) &= -\frac{dS_*(\lambda)}{d\lambda} = -\frac{d}{d\lambda} \left[S(\theta) \frac{d\theta}{d\lambda} \right] \\ &= -\frac{dS(\theta)}{d\lambda} \frac{d\theta}{d\lambda} - S(\theta) \frac{d^2\theta}{d\lambda^2} \\ &= -\frac{dS(\theta)}{d\theta} \frac{d\theta}{d\lambda} \frac{d\theta}{d\lambda} - S(\theta) \frac{d^2\theta}{d\lambda^2} \\ J(\lambda) &= J(\theta) \left(\frac{d\theta}{d\lambda} \right)^2 - S(\theta) \frac{d^2\theta}{d\lambda^2} \end{aligned} \quad (\text{A.11})$$

No máximo temos $S(\hat{\theta}) = 0$, e portanto

$$J(\hat{\lambda}) = \hat{q}^2 J(\hat{\theta}) \quad (\text{A.12})$$

onde $\hat{q}$ é o valor de $\frac{d\theta}{d\lambda}$ no máximo.

No caso com dois parâmetros, temos:

$$(\mu, \sigma^2) \mapsto (\phi, \nu) \quad (\text{A.13})$$

Aqui, no entanto, temos a matriz de informação $J(\phi, \nu)$ dos parâmetros (ϕ, ν) e queremos determinar $J(\mu, \sigma^2)$, assim

$$J(\mu, \sigma^2) = Q' J(\phi, \nu) Q - S \partial Q \quad (\text{A.14})$$

onde

$$Q = \begin{pmatrix} \frac{\partial \phi}{\partial \mu} & \frac{\partial \phi}{\partial \sigma^2} \\ \frac{\partial \nu}{\partial \mu} & \frac{\partial \nu}{\partial \sigma^2} \end{pmatrix} \quad (\text{A.15})$$

$$S = \begin{pmatrix} \frac{\partial l(\phi, \nu)}{\partial \phi} \\ \frac{\partial l(\phi, \nu)}{\partial \nu} \end{pmatrix} \quad (\text{A.16})$$

No máximo temos $S = 0$, e portanto

$$J(\hat{\mu}, \hat{\sigma}^2) = Q' J Q \quad (\text{A.17})$$

Apendice B

Abordagem Bayesiana

Na abordagem bayesiana, uma etapa importante na análise do problema é decidir que densidade a priori será utilizada. Muitas vezes esta decisão é um problema bastante delicado, a sua escolha deve ser criteriosa pois se não for bem feita pode comprometer todo o restante da análise. Como em séries financeiras geralmente temos disponível grandes conjuntos de dados, nosso problema será simplificado pela escolha de uma priori não informativa. Existem muitas maneiras de definir uma priori não informativa. Uma regra muito conhecida é a regra de Jeffreys, veja [Box & Tiao, 73] e [Gelman et al, 95], essa regra estabelece uma priori não informativa dada por

$$p(\Theta) \propto |I(\Theta)|^{\frac{1}{2}} \quad (\text{B.1})$$

onde Θ é o vetor de parâmetros e $I(\Theta)$ é a medida de informação de Fisher sobre Θ

$$I(\Theta) = -E_{X|\Theta} \left[\frac{\partial^2 \log(p(X|\Theta))}{\partial \Theta^2} \right] \quad (\text{B.2})$$

onde $p(X|\Theta)$ é a função de verossimilhança. O problema desta priori é que nem sempre é possível determiná-la. Uma outra maneira de definir uma priori não informativa e que é bem interessante quando se tem uma amostra grande é considerar a matriz de informação observada $J(\hat{\Theta})$ que no caso em que o tamanho da amostra (n) é grande $J(\hat{\Theta})$ converge para $I(\Theta)$ sempre que $p(X|\Theta)$ obedece certas condições de regularidade, veja [Johnson, 67] e [Johnson, 70], e para um n suficientemente grande o estimador de máxima verossimilhança é aproximadamente normal. Assim temos

$$J(\hat{\Theta}) = \left[-\frac{\partial^2 \log(p(X|\Theta = \hat{\Theta}))}{\partial \Theta^2} \right] \quad (\text{B.3})$$

e portanto

$$p(\Theta) \propto |J(\hat{\Theta})|^{\frac{1}{2}} \quad (\text{B.4})$$

Neste trabalho utilizaremos esta regra para priori não informativa.

B.1 Inferência Bayesiana

As conclusões estatísticas bayesianas sobre o vetor de parâmetros Θ , ou valores futuros $\tilde{X}$ do processo, são as afirmações probabilísticas. Estas afirmações probabilísticas estão condicionadas aos dados observados X , que em nossa notação são escritos simplesmente como $p(\Theta|X)$ ou $p(\tilde{X}|X)$.

Afim de fazer afirmações probabilísticas sobre Θ dado X , devemos iniciar com um modelo que nos forneça a distribuição conjunta de probabilidade para θ e X . A distribuição conjunta ou função densidade pode ser escrito com um produto de duas outras densidades que é muitas vezes referidas como distribuição a priori $p(\Theta)$ e a distribuição amostral $p(X|\Theta)$

$$p(\Theta, X) = p(\Theta) p(X|\Theta) \quad (B.5)$$

Condicionado simplesmente nos valores conhecidos dos dados X , usando a propriedade básica da probabilidade condicional conhecida como regra de Bayes, obtem-se a densidade a posteriori

$$p(\Theta|X) = \frac{p(\Theta, X)}{p(X)} = \frac{p(\Theta) p(X|\Theta)}{p(X)} \quad (B.6)$$

onde $p(X) = \sum_{\Theta} p(\Theta) p(X|\Theta)$ no caso de Θ discreto ou $p(X) = \int p(\Theta) p(X|\Theta) d\Theta$ no caso de Θ contínuo. Uma forma equivalente de (B.6) omite o fator $p(X)$, que não depende de Θ e, com X fixo (ao contrário da inferência clássica em que X é variável), pode deste modo ser considerado constante, produzindo a densidade a posteriori não normalizada

$$p(\Theta|X) \propto p(\Theta) p(X|\Theta) \quad (B.7)$$

Estas simples equações expressam os mais importantes fatos da estrutura da inferência bayesiana. A tarefa fundamental de alguma aplicação específica é desenvolver o modelo $p(\Theta, X)$ e efetuar os cálculos necessários para sumariar $p(\Theta|X)$ de maneira apropriada.

B.1.1 Densidade Preditiva

Para fazer inferência sobre o comportamento futuro de X , muitas vezes chamado de inferência preditiva, segue-se uma lógica similar. Antes dos dados serem considerados, a distribuição do desconhecido mais observável X é

$$p(\mathbf{X}) = \int p(\Theta, \mathbf{X}) d\Theta = \int p(\Theta) p(\mathbf{X}|\Theta) d\Theta \quad (\text{B.8})$$

Esta é muitas vezes chamada de distribuição marginal de $\mathbf{X}$, mas um nome mais informativo é distribuição preditiva a priori: priori porque não está condicionada na observação prévia do processo, e preditiva porque ela é a distribuição para uma quantidade que é observável.

Depois que os dados forem observados, podemos prever valores futuros $\tilde{\mathbf{X}}$ do processo. A distribuição de $\tilde{\mathbf{X}}$ é chamada de distribuição preditiva a posteriori e é dado por

$$\begin{aligned} p(\tilde{\mathbf{X}}|\mathbf{X}) &= \int p(\tilde{\mathbf{X}}, \Theta|\mathbf{X}) d\Theta = \int p(\tilde{\mathbf{X}}|\Theta, \mathbf{X}) p(\Theta|\mathbf{X}) d\Theta \\ &= \int p(\tilde{\mathbf{X}}|\Theta) p(\Theta|\mathbf{X}) d\Theta \end{aligned} \quad (\text{B.9})$$

A última linha mostra a distribuição preditiva a posteriori como uma média das predições condicionada sobre a distribuição a posteriori de Θ . Note também que $\mathbf{X}$ e $\tilde{\mathbf{X}}$ são condicionalmente independentes neste modelo dado Θ .

B.1.2 Verossimilhança

Usando a regra de Bayes com um dado modelo probabilístico significa que os dados $\mathbf{X}$ afetam a inferência a posteriori (B.7) somente através da função $p(\mathbf{X}|\Theta)$, que, quando considerado como uma função de Θ , para $\mathbf{X}$ fixo, é chamado de função de verossimilhança. Neste caminho a inferência bayesiana obedece o que é algumas vezes chamado de princípio da verossimilhança, que diz que para uma dada amostra, quaisquer dois modelos probabilísticos $p(\mathbf{X}|\Theta)$ que tenham a mesma função de verossimilhança fornece a mesma inferência para Θ .

O princípio da verossimilhança é razoável somente dentro da estrutura do modelo ou família de modelos adotados para uma particular análise. Na prática, alguém raramente pode estar confiante que o modelo escolhido é o modelo correto. De fato, um estatístico bayesiano aplicado é alguém que está pronto a aplicar a regra de Bayes sobre uma variedade de modelos possíveis para escolher o mais adequado.

B.1.3 Discriminação de Modelos

Uma vez que os modelos probabilísticos escolhidos para a análise foram construídos e calculadas todas as distribuições a posteriori de todos os parâmetros envolvidos devemos passar a uma etapa crucial na abordagem bayesiana: A checagem do melhor modelo. A inferência bayesiana assume que a toda estrutura de um modelo probabilístico pode fornecer falsas con-

clusões quando o modelo é inválido. Uma boa análise bayesiana, portanto, deve incluir no mínimo algum controle de adequação do ajuste do modelo aos dados e a plausibilidade do modelo aos propósitos para o qual será usado. Este problema é algumas vezes discutido como um problema de sensibilidade da distribuição a priori, mas na prática a função de verossimilhança é tipicamente a suspeita. Este é o caso onde mais de um modelo fornece uma ajuste adequado aos dados.

B.1.1 Ordenada Preditiva Condicionada (CPO)

Existem várias técnicas utilizadas na discriminação de modelos. A técnica que utilizaremos neste trabalho é a que usa a distribuição preditiva a posteriori, esse critério de seleção é conhecido como ordenada preditiva condicionada (CPO), veja [Geisser & Eddy, 79]. Assim, dado o vetor de dados observados $\mathbf{X} = (x_1, \dots, x_n)$ temos $\mathbf{X}(r) = (x_1, \dots, x_{r-1}, x_{r+1}, \dots, x_n)$, $r = 1, \dots, n$ e assim

$$p(x_r | \mathbf{X}(r)) = \int p(x_r | \mathbf{X}(r), \Theta) p(\Theta | \mathbf{X}(r)) d\Theta \quad (\text{B.10})$$

A densidade preditiva $p(x_r | \mathbf{X}(r))$ é usada para comparação com x_r , $r = 1, \dots, n$, no sentido que o modelo é adequado se x_r pode ser considerado uma observação aleatória de $p(x_r | \mathbf{X}(r))$. Assim

$$\alpha_r = p(x_r | \mathbf{X}(r)) \quad (\text{B.11})$$

é a ordenada preditiva condicionada (CPO). Valores pequenos de α_r indicam falta de ajuste do modelo para a r -ésima observação. Muitos valores α_r pequenos indicam dúvida sobre a adequabilidade do modelo. Assim como critério de escolha do melhor modelo, escolhe-se aquele com maior valor de α , onde

$$\alpha = \prod_{r=1}^n \alpha_r \quad (\text{B.12})$$

Note no entanto que para processos estocásticos a CPO como foi definida não pode ser utilizada pois neste caso a sequência é essencial. Assim teremos que fazer uma pequena modificação para podermos utilizar este critério. Dado o vetor de observações $\mathbf{X} = (x_1, \dots, x_n)$ dividiremos este vetor em duas partes uma será a amostra de treinamento $\mathbf{X}(k) = (x_1, \dots, x_k)$, $k < n$ e a outra parte $\mathbf{X}(r) = (x_{k+r}, \dots, x_n)$, $r = 1, \dots, n - k$, considere também o vetor $\mathbf{X}(k, r) = (x_1, \dots, x_{k+r-1})$ será utilizada para testar a adequabilidade do modelo. Assim a CPO será dada por

$$p(x_{k+r}|\mathbf{X}(k), \mathbf{X}(k, r)) = \int p(x_{k+r}|\mathbf{X}(k, r), \Theta) p(\Theta|\mathbf{X}(k)) d\Theta \quad (\text{B.13})$$

e o critério de escolha é dado por

$$\alpha = \prod_{r=1}^{n-k} \alpha_r \quad (\text{B.14})$$

Apendice C

Elementos da Análise Estocástica Discreta

Nesta seção daremos alguns resultados que serão necessários ao desenvolvimento dos modelos a tempo discreto.

Desejamos caracterizar o modelo de mercado viável, isto é, um mercado sem oportunidades de arbitragem em termos do comportamento dos incrementos do processo preço descontado $\tilde{S}_t = (1 + r)^t S_t$. Para podermos fazer isto, precisamos primeiro recordar alguns resultados importantes principalmente alguns resultados sobre martingales. Somente os resultados mais básicos necessários às nossas propostas são descritas aqui. Existe uma vasta literatura sobre a teoria dos Martingales a tempo discreto. Duas boas referências neste assunto são [Neveu, 79] e [Doob, 53]

C.1 Base Estocástica Discreta

Utilizaremos uma base estocástica $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$, que é o espaço de probabilidade $(\Omega, \mathcal{F}, \mathbb{P})$ com a filtragem $\mathbb{F} = (\mathcal{F}_t)_{t \in [0, T]}$ com $T \in \mathbb{Z}^+$ isto é, uma coleção não decrescente de sub- σ -álgebras $\mathcal{F}_t \subset \mathcal{F}$.

Uma filtragem determina uma estrutura de informação e evolução do processo, que nos possibilita uma maneira expressar um processo estocástico como uma função de duas variáveis definidas sobre $\Omega \times [0, T]$. Todos os subconjuntos de $\Omega \times [0, T]$ são chamados conjuntos aleatórios.

Uma variável aleatória $\tau : \Omega \rightarrow [0, T] \cup \{\infty\}$ é usualmente chamada um tempo de parada, um tempo de Markov ou um tempo que não depende do futuro, se $\{\tau \leq t\} \in \mathcal{F}_t$ para todo $t \in [0, T]$ com $T \in \mathbb{Z}^+$. A condição $\{\tau \leq t\} \in \mathcal{F}_t$ pode ser substituída por $\{\tau = t\} \in \mathcal{F}_t$ note entretanto que esta troca depende da enumerabilidade de $[0, T]$. Para o caso dos modelos a tempo contínuo $[0, T]$ com $T \in \mathbb{Z}^+$ é substituído por $[0, T]$ com $T \in \mathbb{R}^+$ e aí temos que usar

a condição $\{\tau \leq t\} \in \mathcal{F}_t$, para todo $t \in \mathbb{R}^+$ na definição de tempo de parada. Nos modelos discretos a condição $\{\tau = t\}$ é geralmente mais simples de se verificar. Não obstante, muito dos resultados básicos sobre tempos de parada e suas provas são idênticas em ambas estruturas discreta e contínua.

Definimos a σ -álgebra dos eventos A que precedem os tempos de parada τ como sendo:

$$\mathcal{F}_\tau = \{A \in \mathcal{F} : \{\tau \leq t\} \cap A \in \mathcal{F}_t, t \in [0, T], T \in \mathbb{Z}^+\} \quad (\text{C.1})$$

Dentre todas as sequências $(X_t)_{t \in [0, T]}$ de variáveis aleatórias distinguimos aquelas para as quais X_t são $\mathcal{F}_t$ -mesurável. Tais sequências adaptadas a filtragem $\mathbb{F}$ são chamadas sequências estocásticas.

Uma sequência estocástica $(X_t)_{t \in [0, T]}$ é chamada previsível ou não antecipativa se $X_t \in \mathcal{F}_{t-1}$ para cada $t \in [0, T]$, onde $\mathcal{F}_{-1} = \mathcal{F}_0$.

C.2 Martingales

Definição C.1 Um processo $X = (X_t)_{t \in [0, T]}$ é chamado um martingale relativo a $(\mathbb{F}, \mathbb{P})$ se

1. X é adaptado a $\mathbb{F}$, isto é, X_t é $\mathcal{F}_t$ -mensurável, para todo $t \in [0, T]$;
2. $E(|X_t|) < \infty$ q.s, para todo $t \in [0, T]$;
3. $E(X_t | \mathcal{F}_{t-1}) = X_{t-1}$ q.s para $t \geq 1$.

Notação: $E^{\mathcal{F}_{t-1}}(X_t) = E(X_t | \mathcal{F}_{t-1})$

Assim, X será um supermartingale se no lugar de (3) tivermos

$$E^{\mathcal{F}_{t-1}}(X_t) \leq X_{t-1} \text{ q.s para } t \geq 1 \quad (\text{C.2})$$

e X será um submartingale se no lugar de (3) tivermos

$$E^{\mathcal{F}_{t-1}}(X_t) \geq X_{t-1} \text{ q.s para } t \geq 1 \quad (\text{C.3})$$

Os submartingales e supermartingales são também conhecidos como semimartingales.

De (3) vemos que a melhor previsão de valores futuros não observados de (X_k) baseado na informação $\mathcal{F}_t$ é X_t , em termos mais matemáticos, a variável aleatória Y que minimiza $E^{\mathcal{F}_t}[(X_{t+1} - Y)^2]$ é X_t .

² Isto significa que $E(|X_n|) < \infty$ só não é satisfeita em conjuntos de medida nula.

Note que se X é um martingale se e somente se $\mathbf{E}^{\mathcal{F}_t} [\Delta X_{t+1}] = 0$ para todo $t \in [0, T]$, conseqüentemente $\mathbf{E}^{\mathcal{F}_t} [X_{t+1}] = \mathbf{E}^{\mathcal{F}_t} [X_t]$, portanto um martingale é em média constante. Similarmente, em média um submartingale cresce, e um supermartingale decresce.

Se X é um martingale, então para $p > 1$, usando a esperança condicional iterada e a propriedade martingale repetidamente temos,

$$\begin{aligned}\mathbf{E}^{\mathcal{F}_t} [X_{t+p}] &= \mathbf{E}^{\mathcal{F}_t} [\mathbf{E}^{\mathcal{F}_{t+p-1}} [X_{t+p}]] = \mathbf{E}^{\mathcal{F}_t} [X_{t+p-1}] \\ \mathbf{E}^{\mathcal{F}_t} [X_{t+p-1}] &= \mathbf{E}^{\mathcal{F}_t} [\mathbf{E}^{\mathcal{F}_{t+p-2}} [X_{t+p-1}]] = \mathbf{E}^{\mathcal{F}_t} [X_{t+p-2}]\end{aligned}$$

$$\mathbf{E}^{\mathcal{F}_t} [X_{t+1}] = X_t$$

portanto

$$\mathbf{E}^{\mathcal{F}_t} [X_{t+p}] = X_t \quad (\text{C.4})$$

e similarmente para os semimartingales.

C.2.1 Convergência de Martingales

Daremos agora alguns resultados sobre de convergência, resultados estes que fazem dos martingales uma ferramenta tão poderosa.

Um supermartingale é decrescente na média. Recorde que uma sequência de números reais que é limitado inferiormente converge, isto é, decresce para seu maior limitante inferior, conhecido como ínfimo. Isto sugere que um supermartingale que é limitado inferiormente convirja q.s.. Dizemos que $X \in L^1$ se

$$\sup_t \mathbf{E} |X_t| < \infty \quad (\text{C.5})$$

Daremos a seguir três resultados sobre convergência de martingales. As provas destes teoremas podem ser encontrados em [Doob, 53] e [Neveu, 79].

Teorema C.1 *Um supermartingale $X \in L^1$ é convergente. Portanto existe X_∞ finito tal que*

$$X_t \rightarrow X_\infty \text{ quando } t \rightarrow \infty \text{ q.s.}$$

Teorema C.2 *Um martingale $X \in L^1$ converge q.s., isto é*

$$X_t \rightarrow X_\infty \text{ em } L^1$$

se

$$\lim_{t \rightarrow \infty} E[X_t - X_\infty] = 0$$

Teorema C.3 *As seguintes afirmações são equivalentes para martingales $X = (X_t)_{t \in [0, T]}$:*

1. X é convergente em L^1 ;
2. X é L^1 -limitado, e seu limite X_∞ satisfaz

$$X_t = E[X_\infty | \mathcal{F}_t];$$
3. Existe um variável aleatória X integrável com

$$X_t = E[X | \mathcal{F}_t].$$

A classe de martingales que tem estas propriedades são chamados de martingales uniformemente integráveis ou regulares.

C.2.1 Decomposição de Doob

O próximo resultado mostrar que toda sequência de variáveis aleatórias tendo média finita pode ser transformada em um martingale.

Considere $X = (X_t)_{t \in [0, T]}$ uma sequência de variáveis aleatórias com média finita, adaptada a filtragem $(\mathcal{F}_t)_{t \in [0, T]}$. Para $t > 0$, defina

$$U_t = E(X_t - X_{t-1} | \mathcal{F}_{t-1}) \quad (\text{C.6})$$

e

$$Y_t = X_t - \sum_{k=1}^t U_k \quad (\text{C.7})$$

com $Y_0 = X_0$. Então a sequência Y_t é um martingale com respeito a filtragem $(\mathcal{F}_t)_{t \in [0, T]}$.

Teorema C.4 (Decomposição de Doob) *Seja X um submartingale com respeito a filtragem $(\mathcal{F}_t)_{t \in [0, T]}$. Então existe uma única sequência aleatória Y e V tal que*

1. para todo $t \geq 0$, $X_t = Y_t + V_t$;
2. Y é um martingale com respeito a filtragem $(\mathcal{F}_t)_{t \in [0, T]}$;
3. V é uma sequência crescente a partir do zero, isto é, $0 = V_0 \leq V_1 \leq \dots$;
4. V é uma sequência previsível ou não antecipativa, isto é, $V_t \in \mathcal{F}_{t-1}$.

Deste modo, o teorema acima nos diz que todo submartingale pode ser escrito como a soma de um martingale com uma sequência aleatória não antecipativa crescente.

Embora o teorema da decomposição de Doob seja um resultado simples na estrutura a tempo discreto, o análogo a tempo contínuo, a decomposição de Doob-Meyer, é profunda. Isto ilustra os contraste que podem surgir entre as teorias dos processos estocásticos nos cenários do tempo discreto e do tempo contínuo.

C.2.1 Transformações de Martingales

Daremos agora um resultado que mostra como martingales ou semimartingales podem ser transformados em outros martingales ou semimartingales.

Consideremos inicialmente um exemplo de um jogo ou uma série de investimentos especulativos, que se desenvolve ao longo do tempo numa estrutura temporal discretizada. Vamos assumir que em $t = 0$ não ocorre nenhuma jogada, o resultado líquido a cada rodada por unidade apostado no período t é dado por

$$\Delta X_t = X_t - X_{t-1} \quad (\text{C.8})$$

Deste modo se X_t é um martingale o jogo é justo na média. Denotemos por φ_t o processo $\mathcal{F}_{t-1}$ -mensurável, como nossa aposta na rodada t , portanto φ_t é não antecipativo. Ser não antecipativo significa que devemos decidir quanto vamos apostar na rodada t baseado na informação que temos até a rodada $t - 1$. O resultado na rodada t é

$$\varphi_t \Delta X_t \quad (\text{C.9})$$

O resultado, que pode ser positivo ou negativo, acumulado até t é dado por

$$Y_t = \sum_{k=1}^t \varphi_k \Delta X_k \quad (\text{C.10})$$

Notação:

$$Y_t = (\varphi \bullet X)_{t \in [0, T]} \quad (\text{C.11})$$

$$Y = \varphi \bullet X \quad (\text{C.12})$$

Definição C.2 (Transformação Martingale) *Seja $(X_t)_{t \in [0, T]}$ um martingale e $(\varphi_t)_{t \in [0, T]}$ um processo limitado e não antecipativo, portanto $\mathcal{F}_{t-1}$ -mensurável, definidos sobre $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$.*

O processo $Y_t = (\varphi \bullet X)_{t \in [0, T]}$, dado para $t \geq 1$ é dado por

$$Y_t = \sum_{k=1}^t \varphi_k \Delta X_k$$

com $Y_0 = 0$ é um martingale e é chamado a transformação martingale de X por φ .

Se φ for um processo não negativo e limitado e X for um supermartingale então $\varphi \bullet X$ é supermartingale nulo no zero e se X for um submartingale então $\varphi \bullet X$ será um submartingale nulo no zero.

Lema C.1 (Transformação Martingale) *Uma sequência adaptada de variáveis aleatórias $(X_t)_{t \in [0, T]}$ é um martingale se e somente se para toda sequência limitada e não antecipativa $(\varphi_t)_{t \in [0, T]}$, temos*

$$E \left[\sum_{k=1}^t \varphi_k \Delta X_k \right] = 0$$

A não antecipatividade de φ significa que não podemos prever o futuro, o que é realista e justo, portanto não esperamos ganhar nem perder nada ao longo do tempo.

As transformações de martingales são o análogo discreto da integral estocástica em que o martingale X é usado como integrador. O cálculo de Itô baseado nesta teoria de integração, forma a estrutura da precificação martingale a tempo contínuo. Portanto um entendimento desta, tecnicamente mais simples, transformação martingale fornece uma valiosa compreensão dentro dos princípios básicos do cálculo estocástico.

C.2.1 Tempos de Parada

Um dos pontos centrais da teoria dos martingales diz respeito a obtenção de subsequências de variáveis aleatórias de martingales pela amostragem deles numa sequência crescente de tempos de parada. Daremos dois resultados que serão úteis para as propostas neste trabalho.

Definição C.3 (Tempo de Parada) *Uma variável aleatória $\tau \in \mathbb{Z}^+$ é chamada um tempo de parada, um tempo opcional ou uma amostra opcional, se*

$$\{\tau \leq t\} = \{\omega : \tau(\omega) \leq t\} \in \mathcal{F}_t, \forall t \in \mathbb{Z}^+$$

e $\Upsilon = \{\tau_1, \tau_2, \dots, \tau_t\}$ é o conjunto de tempos de parada até t .

Pense em τ como o tempo em que decidimos sair do jogo. Se saímos ou não no tempo τ esta decisão é tomada utilizando somente a informação até o tempo t . Deste modo, tempos de parada modelam situações em que não há conhecimento antecipado do futuro. Em particular,

no contexto financeiro onde não existem grupos por dentro. Os grupos por dentro são aquelas pessoas que possam ter alguma informação privilegiada sobre o que ocorrerá no futuro. Infelizmente sabemos que existem tais grupos de especuladores que usam tais informação para auferir lucros sem correr riscos, no entanto os modelos serão construídos assumindo que não existem tais grupos.

Teorema C.5 (Tempo Opcional de Doob) *Seja τ um tempo de parada e $(X_t)_{t \in \mathbb{Z}^+}$ um supermartingale. Assumindo que um dos itens abaixo é satisfeito:*

1. τ é limitado, isto é, $\tau(\omega) \leq k$ para alguma constante k para todo $\omega \in \Omega$;
2. $(X_t)_{t \in \mathbb{Z}^+}$ é limitado, isto é $|X_t(\omega)| \leq k$ para alguma constante k para todo $t \in \mathbb{Z}^+$, $\omega \in \Omega$;
3. $E(\tau) < \infty$ e $(X_t - X_{t-1})$ é limitado.

Então X_τ é integrável, e

$$E(X_\tau) \leq E(X_0)$$

se $(X_t)_{t \in \mathbb{Z}^+}$ for um martingale, então

$$E(X_\tau) = E(X_0)$$

Notação: denotemos $(X_{t \wedge \tau})_{t \in \mathbb{Z}^+}$ para uma sequência de paradas.

Proposição C.1 (Tempo de Parada) *Considere a seguinte sequência $(X_t)_{t \in \mathbb{Z}^+}$, então:*

1. Se $(X_t)_{t \in \mathbb{Z}^+}$ for adaptada à $(\mathcal{F}_t)_{t \in \mathbb{Z}^+}$ e τ é um tempo de parada então a sequência de parada $(X_{t \wedge \tau})_{t \in \mathbb{Z}^+}$ é adaptada à $(\mathcal{F}_t)_{t \in \mathbb{Z}^+}$;
2. Se $(X_t)_{t \in \mathbb{Z}^+}$ é um martingale (supermartingale) e τ é um tempo de parada então $(X_{t \wedge \tau})_{t \in \mathbb{Z}^+}$ é um martingale (supermartingale).

A segunda afirmação da proposição nos diz que se um jogo é justo ele permanece justo se a partir de um dado momento τ decidimos parar.

O fato de que a sequência de parada $(X_{t \wedge \tau})_{t \in \mathbb{Z}^+}$ continuar sendo um martingale se $(X_t)_{t \in \mathbb{Z}^+}$ for um martingale é fundamental quando se estuda o problema da precificação da opção de venda americana.

Apendice D

Elementos da Análise Estocástica Contínua

Vamos agora dar alguns resultados que serão importantes no desenvolvimento na análise do problema a tempo contínuo. As referências consultadas que foram a base para seção são as seguintes: [Bingham & Kiesel, 98], [Elliott & Kopp, 99], [Gihman & Skorohod, 72], [Karatzas & Shreve, 91] e [Oksendal, 92]. Outras referências utilizadas serão citadas quando necessário.

D.1 Considerações Iniciais

Faremos agora um pequeno resumo, alguns pontos que foram tratado no caso discreto não serão tratados aqui pois seria apenas repetição. A diferença agora é que teremos $t \in \mathbb{R}^+$. Assim onde tínhamos t e $t+1$, $t \in \mathbb{Z}^+$ temos agora s e t com $s < t$, $s, t \in \mathbb{R}^+$.

Assim teremos, uma base estocástica $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$, que é o espaço de probabilidade $(\Omega, \mathcal{F}, \mathbb{P})$ com a filtragem $\mathbb{F} = (\mathcal{F}_t)_{t \in \mathbb{R}^+}$. A filtragem $\mathbb{F}$ para o caso contínuo deve agora satisfazer a certas “condições usuais”. Isto significa que a filtragem $\mathbb{F}$ é:

- completa; isto é, todo conjunto nulo em $\mathcal{F}$ pertence a $\mathcal{F}_0$, e portanto a cada $\mathcal{F}_t$;
- contínuo à direita; isto é, $\mathcal{F}_t = \bigcap_{s>t} \mathcal{F}_s$.

Exatamente como no caso discreto $\mathcal{F}_t$ representa a história do processo até o tempo t só que agora a filtragem consiste de uma coleção não enumerável de σ -álgebras. A continuidade à direita é específica para a estrutura a tempo contínuo.

Definição D.1 (Tempo de Parada) *Seja $t \in \mathbb{R}^+$. Uma variável aleatória τ tomando valores em $\mathbb{R}^+$ é um tempo de parada se $\{\tau \leq t\} \in \mathcal{F}_t$ para todo $t \geq 0$. Consequentemente, o evento*

$\{\tau \leq t\}$ depende somente da história do processo até o tempo t . A primeira vez que preço da ação (considerando o nosso caso) chega a um certo nível é um tempo de parada, neste estado em que se encontra, dizemos, a primeira vez que o preço atinge um certo nível superior após passar por um determinado valor. Portanto, a última vez, antes de alguma determinada data, na qual o preço da ação, atinge um certo nível não é um tempo de parada, por que dizer “a última vez” requer informação sobre o futuro.

Note agora que diferentemente do caso discreto, no caso contínuo não faz sentido trocar a condição $\{\tau \leq t\} \in \mathcal{F}_t$ por $\{\tau = t\} \in \mathcal{F}_t$.

D.2 Martingales

O conceito de martingales é igual ao do caso discreto a diferença agora é que $t \in \mathbb{R}^+$.

Definição D.2 Considere a base estocástica $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$. Um processo $X = (X_t)_{t \in \mathbb{R}^+}$ é chamado um martingale relativo a $(\mathbb{F}, \mathbb{P})$ se

1. X é adaptado a $\mathbb{F}$, isto é, X_t é $\mathcal{F}_t$ -mensurável, para todo $t \in \mathbb{R}^+$;
2. $E(|X_t|) < \infty$ q.s.³, para todo $t \in \mathbb{R}^+$;
3. $E(X_t | \mathcal{F}_s) = X_s$ q.s para $s \leq t$.

O mais importante exemplo de um martingale a tempo contínuo é o movimento Browniano ou processo de Wiener.

Definição D.3 (Processo de Wiener) Dado a base estocástica $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$. O processo de Wiener padrão $(B_t)_{t \in \mathbb{R}^+}$ é um processo estocástico que tem as seguintes propriedades:

1. As trajetórias são contínuas para quase todo $\omega \in \Omega$, isto é, o conjunto de realizações ω que não possuem esta propriedade é um conjunto de medida nula, ou seja, $t \rightarrow B_t(\omega)$ é contínuo q.s.;
2. Para $t > s$ então os acréscimos $(B_t - B_s)$ são variáveis aleatórias independentes. Isto é o que chamamos de incrementos independentes;
3. Para $t > s$ então os acréscimos $(B_t - B_s)$ são variáveis aleatórias com distribuição normal de média zero e variância $t - s$;
4. O ponto inicial do processo é $B_0 = 0$;
5. $(B_t)_{t \in \mathbb{R}^+}$ é um martingale.

³ Isto significa que $E(|X_n|) < \infty$ só não é satisfeita em conjuntos de medida nula.

Uma construção matemática rigorosa da medida de Wiener não é trivial e está fora do escopo deste trabalho, no entanto pode-se mostrar através da conexão entre uma equação diferencial estocástica

$$dX_t = a(X_t) dt + b(X_t) dB_t \quad (D.1)$$

e uma equação diferencial parabólica

$$\frac{\partial p}{\partial t} + Lp = 0 \quad (D.2)$$

e L é o operador diferencial elíptico

$$L = A(x) \frac{\partial}{\partial x} + B(x) \frac{\partial^2}{\partial x^2} \quad (D.3)$$

onde

$$A(x) = a(x) \quad \text{e} \quad B(x) = \frac{[b(x)]^2}{2} \quad (D.4)$$

que a medida de Wiener é dada por

$$p(x, t) = \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{x^2}{2t}\right) \quad (D.5)$$

mais adiante mostraremos esta conexão entre as EDE's e EDP's. A equação (D.2) também é conhecida como Equação reversa de Kolmogorov (Backward equation).

Partindo do conhecimento que a medida de Wiener é dada pela equação (D.5) então é fácil provar as propriedades (2) a (5) do processo de Wiener descritas acima. Para mostrar a propriedade (2) por exemplo, basta lembrar que

$$E[(B_t)^2] = t \quad (D.6)$$

e

$$E(B_s B_t) = \min(s, t) \quad (D.7)$$

e usar o fato que variáveis aleatórias gaussianas são independentes se e só se elas são não correlacionadas. Portanto, basta mostrar que

$$E[(B_{t_{i+1}} - B_{t_i})(B_{t_{j+1}} - B_{t_j})] = 0 \quad (D.8)$$

assim temos que

$$\begin{aligned}
E[(B_{t_{i+1}} - B_{t_i})(B_{t_i} - B_{t_{i-1}})] &= E[B_{t_{i+1}}B_{t_i} - B_{t_{i+1}}B_{t_{i-1}} - B_{t_i}B_{t_i} + B_{t_i}B_{t_{i-1}}] \\
&= E(B_{t_{i+1}}B_{t_i}) - E(B_{t_{i+1}}B_{t_{i-1}}) - \\
&\quad E[(B_{t_i})^2] + E(B_{t_i}B_{t_{i-1}}) \\
&= t_i - t_{i-1} - t_i + t_{i-1} \\
&= 0 \blacksquare
\end{aligned} \tag{D.9}$$

a propriedade (3) e (5) é trivial e a propriedade (4) é somente uma convenção. Restou somente a propriedade (1). Para que possamos verificá-la precisaremos do teorema de Kolmogorov para isso, veja [Oksendal, 92] e [Wong, 71].

Teorema D.1 (Critério de Continuidade de Kolmogorov) *Suponha que o processo $(X_t)_{t \in \mathbb{R}^+}$ satisfaça a seguinte condição. Para todo $t \in \mathbb{R}^+$ existe constantes positivas α , β e D tal que*

$$E[|X_t - X_s|^\alpha] \leq D|t - s|^{1+\beta}; \quad 0 \leq s < t$$

então

$$\lim_{h \rightarrow 0} \sup_{|t-s| < h} |X_t - X_s| = 0$$

portanto existe uma versão contínua de $(X_t)_{t \in \mathbb{R}^+}$.

Portanto para o processo de Wiener a condição acima é satisfeita para $D = 3$, $\alpha = 4$ e $\beta = 1$. Para isto basta ver que

$$E[(B_t - B_s)^4] = 3(t - s)^2 \tag{D.10}$$

Teorema D.2 (Caracterização do Processo de Wiener) *Seja $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ uma base estocástica e $(X_t)_{t \in \mathbb{R}^+}$ um processo definido sobre esta base. Se X_t é um martingale (local) contínuo sobre esta base tal que*

$$X_t^2 - t$$

é um martingale (local) contínuo. Então $(X_t)_{t \in \mathbb{R}^+}$ é um processo de Wiener.

Para prova deste teorema, veja [Karatzas & Shreve, 91] ou [Elliott & Kopp, 99]. Esta caracterização do processo de Wiener é devida a Lévy. estas propriedades implicam outras propriedades bem conhecidas do processo de Wiener, como por exemplo, que $(X_t)_{t \in \mathbb{R}^+}$ é um processo gaussiano com incrementos independentes.

D.3 Equações Diferenciais Estocásticas

Vamos inicialmente considerar o problema hipotético de um investidor que deseja investir suas economias na poupança onde o juro mais a correção monetária r , neste caso hipotético, sejam constantes e a capitalização seja contínua no tempo, assim

$$\frac{dX_t}{dt} = rX_t \quad (\text{D.11})$$

desta modo, o montante que é o capital inicial mais os juros e correção monetária a qualquer tempo t será

$$X_t = X_0 e^{rt} \quad (\text{D.12})$$

onde X_0 é o capital inicial.

No entanto, sabemos que este modelo não é realista. Um modelo mais adequado pode ser construído se considerarmos que r varia no tempo e que além disso essa variação seja aleatória, isto é

$$r_t = \mu + \sigma Z_t \quad (\text{D.13})$$

onde Z_t é o ruído no tempo t . Este ruído é conhecido como ruído branco.

Assim (D.11) toma a seguinte forma

$$\frac{dX_t}{dt} = rX_t + \sigma X_t Z_t \quad (\text{D.14})$$

A equação acima é conhecida como equação de Langevin. desta maneira agora temos um modelo mais realista. A questão é como resolver esta equação, pois posto desta maneira, as regras do cálculo clássico não podem ser aplicados aqui. De uma maneira mais genérica podemos ter

$$\frac{dX_t}{dt} = a(t, X_t) + b(t, X_t)Z_t \quad (\text{D.15})$$

Colocando de outra forma, temos

$$dX_t = a(t, X_t)dt + b(t, X_t)dB_t \quad (\text{D.16})$$

onde $dB_t = Z_t dt$ é o processo de Wiener. Na forma integral, temos

$$X_t = X_0 + \int_0^t a(s, X_s)ds + \int_0^t b(s, X_s)dB_s \quad (\text{D.17})$$

O processo estocástico X_t descrito desta pelas equações (D.16) e (D.17) é conhecido como processo de difusão de Itô. Ressaltamos no entanto que, a forma diferencial da equação (D.16) é apenas simbólica pois no cálculo de Itô tem-se apenas o conceito de integral estocástica. A questão agora é como definir a segunda integral da equação (D.17). Deste modo, vamos a partir de agora provar a existência de integrais do tipo

$$\int_0^t f(s, X_s) dB_s \quad (D.18)$$

para uma classe de funções $f : [0, \infty] \times \Omega \rightarrow \mathbb{R}$, onde B_t é o movimento browniano ou processo de Wiener.

Suponha $0 \leq S < T$ e $f(t, X_t)$ é uma função conhecida. Queremos definir

$$\int_S^T f(s, X_s) dB_s \quad (D.19)$$

Vamos mostrar algumas dificuldades que são encontradas ao se tentar definir a integral estocástica, será dado também um resultado importante que é utilizado quando definimos a integral estocástica, este resultado é conhecido como a Isometria de Itô, será mostrado apenas para a classe de funções simples. A extensão para uma classe mais geral e uma construção rigorosa da integral estocástica pode ser encontrada por exemplo em [Friedman, 75], [Ikeda & Watanabe, 81], [Karatzas & Shreve, 91] ou [Oksendal, 92].

D.3.1 Dificuldades na Definição da Integral Estocástica

Uma função f é simples se ela pode ser escrita na seguinte forma

$$f(t, X_t) = \sum_{j \geq 0} e_j(\omega) \chi_{(t_j, t_{j+1}]}(t) \quad (D.20)$$

onde $\chi_{(t_j, t_{j+1}]}$ denota a função característica e $n \in \mathbb{Z}^+$. Para tais funções é razoável definir

$$\int_S^T f(t, X_t) dB_t = \sum_{j \geq 0} e_j(\omega) [B_{t_{j+1}} - B_{t_j}] \quad (D.21)$$

onde

$$t_k = t_k^{(n)} = S + k \frac{T-S}{n} \quad (D.22)$$

porém, sem suposições adicionais sobre as funções e_j teremos dificuldades em determinar

$$\sum_{j \geq 0} e_j(\omega) [B_{t_{j+1}} - B_{t_j}] \quad (D.23)$$

O motivo é que, como mostraremos a seguir, ao contrário de uma somatória de Riemann em que a função e_j pode ser tomada em qualquer ponto do intervalo $(t_j, t_{j+1}]$ que o resultado do somatório será sempre o mesmo, aqui o resultado depende da escolha do ponto aonde será tomada a função e_j . Vamos então, calcular

$$E \left[\int_0^T f(t, X_t) dB_t \right] \quad (\text{D.24})$$

para mostrar as dificuldades em se determinar um valor para uma integral em que o integrando e o integrador são um processo estocástico.

Consideremos o caso em que o funcional f é $f(t, X_t) = B_t$. Assim, vamos considerar duas possíveis aproximações para esta função:

$$f_1(t, x_t) = \sum_{j=0}^n B_{t_j} \chi_{(t_j, t_{j+1}]}(t) \quad (\text{D.25})$$

$$f_2(t, x_t) = \sum_{j=0}^n B_{t_{j+1}} \chi_{(t_j, t_{j+1}]}(t) \quad (\text{D.26})$$

Deste modo, para f_1 temos:

$$\begin{aligned} E \left[\int_0^T f_1 dB_t \right] &= E \left[\sum_{j=0}^{n-1} B_{t_j} (B_{t_{j+1}} - B_{t_j}) \right] \\ &= \sum_{j=0}^{n-1} E \left[B_{t_j} B_{t_{j+1}} - B_{t_j}^2 \right] \\ &= \sum_{j=0}^{n-1} \left[E(B_{t_j} B_{t_{j+1}}) - E(B_{t_j}^2) \right] \\ &= \sum_{j=0}^{n-1} (t_{j+1} - t_j) = 0 \end{aligned} \quad (\text{D.27})$$

e para f_2 temos

$$\begin{aligned}
E \left[\int_0^T f_2 dB_t \right] &= E \left[\sum_{j=0}^{n-1} B_{t_{j+1}} (B_{t_{j+1}} - B_{t_j}) \right] \\
&= \sum_{j=0}^{n-1} E \left[B_{t_{j+1}}^2 - B_{t_{j+1}} B_{t_j} \right] \\
&= \sum_{j=0}^{n-1} \left[E(B_{t_{j+1}}^2) - E(B_{t_{j+1}} B_{t_j}) \right] \\
&= \sum_{j=0}^{n-1} (t_{j+1} - t_j) = \sum_{j=0}^{n-1} \left[(j+1) \frac{T}{n} - j \frac{T}{n} \right] \\
&= \sum_{j=0}^{n-1} \frac{T}{n} = T
\end{aligned} \tag{D.28}$$

Portanto, apesar de ambos f_1 e f_2 serem boas aproximações para o funcional f , a integral de acordo com (D.21) não é fechada, isto é, depende do ponto aonde será tomado o funcional f , não importando quão grande seja o n escolhido.

Isto reflete unicamente o fato de que as variações das trajetórias de B_t são tão grandes que nos impossibilita construir a integral (D.19) no sentido de Riemann-Stieltjes. Pode ser mostrado que as trajetórias $t \rightarrow B_t$ do movimento browniano são não diferenciáveis, e que, através da lei do logaritmo iterado, a variação total das trajetórias é infinita, veja [Karatzas & Shreve, 91].

Em geral é natural aproximar uma dada função $f(t, X_t)$ por:

$$\sum_j f(t_j^*, X_{t_j^*}) \chi_{(t_j, t_{j+1}]}(t) \tag{D.29}$$

aonde os pontos t_j^* pertencem aos intervalos $[t_j, t_{j+1}]$, e então definimos $\int_S^T f(s, X_s) dB_s$ como o limite (num sentido que será explicado mais adiante) de $\sum_j f(t_j^*, X_{t_j^*}) [B_{t_{j+1}} - B_{t_j}]$ quando $n \rightarrow \infty$. Entretanto como foi mostrado, diferentemente da integral de Riemann-Stieltjes, aqui faz diferença quais pontos t_j^* dos intervalos $[t_j, t_{j+1}]$ são escolhidos.

Duas interpretações para esta integral que correspondem a duas escolhas de t_j^* tem sido muito utilizadas.

1. $t_j^* = t_j$ (o ponto extremo esquerdo), que nos leva a integral de Itô, e será denotado de agora em diante por

$$\int_S^T f(t, X_t) dB_t \tag{D.30}$$

2. $t_j^* = \frac{t_j + t_{j+1}}{2}$ (o ponto médio), que nos leva a integral de Stratonovich, e será denotado de agora em diante por

$$\int_S^T f(t, X_t) \circ dB_t \quad (D.31)$$

A grande vantagem da integral no sentido de Itô é que ela é uma extensão natural da transformação martingale do caso discreto.

Lema D.1 (A Isometria de Itô) Se $f(t, X_t)$ é limitada e simples então

$$E \left[\left(\int_S^T f(t, X_t) dB_t \right)^2 \right] = E \left[\int_S^T f(t, X_t)^2 dt \right]$$

Prova: Temos que

$$\int_S^T f(t, X_t) dB_t = \sum_{j \geq 0} e_j(X_{t_j}) \Delta B_{t_j}$$

onde $\Delta B_{t_j} = B_{t_{j+1}} - B_{t_j}$, portanto

$$\begin{aligned} E \left[\left(\int_S^T f(t, X_t) dB_t \right)^2 \right] &= E \left[\sum_{i,j} e_i(X_{t_i}) e_j(X_{t_j}) \Delta B_{t_i} \Delta B_{t_j} \right] \\ &= \sum_{i,j} E [e_i(X_{t_i}) e_j(X_{t_j}) \Delta B_{t_i} \Delta B_{t_j}] \end{aligned}$$

como sabemos os incrementos do movimento browniano são independentes, portanto:

Se $i \neq j$ com $i < j$

$$\begin{aligned} E [e_i(X_{t_i}) e_j(X_{t_j}) \Delta B_{t_i} \Delta B_{t_j}] &= e_i(X_{t_i}) E [e_j(X_{t_j}) \Delta B_{t_i}] E [\Delta B_{t_j}] \\ &= 0 \end{aligned}$$

Se $i = j$, então

$$\begin{aligned} E [e_i(X_{t_i}) e_j(X_{t_j}) \Delta B_{t_i} \Delta B_{t_j}] &= E [(e_j(X_{t_j}))^2] E [(\Delta B_{t_i})^2] \\ &= E [(e_j(X_{t_j}))^2] (t_{j+1} - t_j) \end{aligned}$$

Assim

$$E [e_i(X_{t_i}) e_j(X_{t_j}) \Delta B_{t_i} \Delta B_{t_j}] = \begin{cases} 0 & \text{se } i \neq j \\ E [e_j(X_{t_j})^2] (t_{j+1} - t_j) & \text{se } i = j \end{cases}$$

Portanto

$$\begin{aligned} E \left[\left(\int_S^T f(t, X_t) dB_t \right)^2 \right] &= \sum_{j \geq 0} E [e_j(X_{t_j})^2] (t_{j+1} - t_j) \\ &= E \left[\int_S^T f(t, X_t)^2 dt \right] \blacksquare \end{aligned}$$

Vamos agora mostrar como definir a integral estocástica para duas funções bem simples

$$f(t, B_t) = B_t \quad e \quad f(t, B_t) = t B_t \quad (D.32)$$

para estas funções determinaremos suas respectivas Integrais Estocásticas e então deduziremos o Lema de Itô, que será de extrema utilidade no cálculo de integrais mais complicadas.

Portanto temos

1. $\int_0^t B_s dB_s$
Desde que $f(t, X_t) = B_t$ é contínua q.s., definiremos ϕ_n da seguinte maneira

$$\phi_n^{(t)} = \sum_{j=0}^{n-1} B_{t_j} \chi_{(t_j, t_{j+1}]}(t) \quad \beta_t \quad (D.33)$$

onde $\chi_{(t_j, t_{j+1}]}(t)$ é a função característica para uma partição qualquer do intervalo $[0, t]$, $0 = t_0 < t_1 < \dots < t_n$, com $t_j = a + j \frac{b-a}{n}$, neste caso $a = 0$ e $b = t$, $j = 0, 1, \dots, n$, onde ϕ_n converge uniformemente para f em $[0, t]$. Desde modo, temos:

$$I_n = \int_0^t B_s dB_s = \sum_{j=0}^{n-1} B_{t_j}^+ (B_{t_{j+1}} - B_{t_j})$$

observe que

$$\begin{aligned} \frac{(b^2 - a^2) - (b - a)^2}{2} &= \frac{(b - a)(b + a) - (b - a)(b - a)}{2} \\ &= \frac{(b - a)(b + a - b + a)}{2} = \frac{(b - a)2a}{2} \\ &= a(b - a) \end{aligned} \quad (D.34)$$

portanto de D.34 temos que

$$\begin{aligned} I_n &= \frac{1}{2} \sum_{j=0}^{n-1} \left[(B_{t_{j+1}}^2 - B_{t_j}^2) - (B_{t_{j+1}} - B_{t_j})^2 \right] \\ &= \frac{1}{2} (B_{t_n}^2 - B_{t_0}^2) - \frac{1}{2} \sum_{j=0}^{n-1} Z_j^2 \end{aligned} \quad (D.35)$$

onde $Z_j = B_{t_{j+1}} - B_{t_j}$ e portanto das propriedades do processo de Wiener temos que $Z_j \sim N(0, (t_{j+1} - t_j))$.

Mostremos agora que $\sum_{j=0}^{n-1} Z_j^2 \rightarrow t$. Sabemos que

$$E \left(\sum_{j=0}^{n-1} Z_j^2 \right) = \sum_{j=0}^{n-1} (t_{j+1} - t_j) = t \quad (D.36)$$

e

$$\begin{aligned} Var \left(\sum_{j=0}^{n-1} Z_j^2 \right) &= E \left(\sum_{j=0}^{n-1} Z_j^4 \right) - \left[E \left(\sum_{j=0}^{n-1} Z_j^2 \right) \right]^2 \\ &\leq E \left(\sum_{j=0}^{n-1} Z_j^4 \right) \end{aligned} \quad (D.37)$$

e

$$E \left(\sum_{j=0}^{n-1} Z_j^4 \right) = 3 \sum_{j=0}^{n-1} (t_{j+1} - t_j)^2 = \frac{3t^2}{n} \quad (\text{D.38})$$

Da desigualdade de Chebyshev vem que

$$\mathbf{P} \left[\left| \left(\sum_{j=0}^{n-1} Z_j^2 \right) - E \left(\sum_{j=0}^{n-1} Z_j^2 \right) \right| > \varepsilon \right] \leq \frac{\text{Var} \left(\sum_{j=0}^{n-1} Z_j^2 \right)}{\varepsilon^2} \quad (\text{D.39})$$

De D.36, D.37 e D.38, temos que

$$\begin{aligned} \mathbf{P} \left[\left| \left(\sum_{j=0}^{n-1} Z_j^2 \right) - t \right| > \varepsilon \right] &\leq E \left(\sum_{j=0}^{n-1} Z_j^4 \right) \\ &\leq \frac{3t^2}{n} \rightarrow 0 \text{ quando } n \rightarrow \infty \end{aligned} \quad (\text{D.40})$$

Portanto

$$\int_0^t B_s dB_s = \frac{1}{2} B_t^2 - \frac{1}{2} t \quad (\text{D.41})$$

observe que o resultado é diferente da integral clássica. Passemos agora a segunda função.

$$2. \int_0^t s B_s dB_s$$

Vamos começar integrando por partes, assim fazemos

$$u = t \Rightarrow du = dt \quad (\text{D.42})$$

$$dv = B_s dB_s \Rightarrow v = \frac{1}{2} B_t^2 - \frac{1}{2} t \quad (\text{D.43})$$

Portanto

$$\begin{aligned} \int_0^t s B_s dB_s &= \int_0^t d \left(\frac{s B_s^2}{2} - \frac{s^2}{2} \right) - \int_0^t \left(\frac{B_s^2}{2} - \frac{s}{2} \right) dt \\ 2 \int_0^t s B_s dB_s &= \int_0^t d(s B_s^2) - \int_0^t d(s^2) - \int_0^t B_s^2 ds + \int_0^t s ds \\ \int_0^t d(s B_s^2) &= \int_0^t B_s^2 ds + 2 \int_0^t s B_s dB_s + \int_0^t d(s^2) - \int_0^t s ds \\ \int_0^t d(s B_s^2) &= \int_0^t B_s^2 ds + 2 \int_0^t s B_s dB_s + \int_0^t s ds \end{aligned}$$

colocando na forma diferencial temos:

$$d(t B_t^2) = B_t^2 dt + 2t B_t dB_t + t dt \quad (\text{D.44})$$

Denotando $F = t B_t^2$ temos pelas regras do cálculo clássico:

$$\frac{\partial F}{\partial t} dt = B_t^2 dt \quad \frac{\partial F}{\partial B_t} dB_t = 2t B_t dB_t \quad \frac{1}{2} \frac{\partial^2 F}{\partial B_t^2} (dB_t)^2 = t dt$$

A partir destes resultados podemos agora deduzir a fórmula da mudança da variável de Itô

$$\partial F = \frac{\partial F}{\partial t} dt + \frac{\partial F}{\partial B} dB + \frac{1}{2} \frac{\partial^2 F}{\partial B^2} (dB)^2 \quad (\text{D.45})$$

Note que $(dB)^2 = dt$. Numa forma mais genérica temos o seguinte

Lema D.2 (Fórmula da Mudança de Variável de Itô) *Seja o processo estocástico $(X_t)_{t \in \mathbb{R}^+}$ que é solução da EDE*

$$dX_t = a(X_t) dt + b(X_t) dB_t$$

e $f(t, x) \in C_0^2(\mathbb{R}^+ \times \mathbb{R})$. Então

$$Y_t = f(t, x)$$

e portanto

$$dY_t = \left[\frac{\partial f}{\partial t} + a(x) \frac{\partial f}{\partial x} + \frac{[b(x)]^2}{2} \frac{\partial^2 f}{\partial x^2} \right] dt + b(x) \frac{\partial f}{\partial x} dB_t$$

A prova completa deste teorema pode ser encontrado por exemplo em [Friedman, 75], [Ikeda & Watanabe, 81], [Karatzas & Shreve, 91] ou [Oksendal, 92].

D.4 Conecção Entre o Cálculo Estocástico e as Equações Diferenciais Parciais

Nesta seção queremos mostrar que a densidade de transição de um processo $(X_t)_{t \in \mathbb{R}^+}$ definido sob a base estocástica $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ satisfazendo uma equação diferencial estocástica

$$dX_t = a(X_t) dt + b(X_t) dB_t \quad (\text{D.46})$$

pode ser obtida resolvendo a equação diferencial parabólica

$$\frac{\partial u(t, x)}{\partial t} + a(x) \frac{\partial u(t, x)}{\partial x} + \frac{(b(x))^2}{2} \frac{\partial^2 u(t, x)}{\partial x^2} = 0 \quad (\text{D.47})$$

conhecida como Equação reversa de Kolmogorov (backward equation), ou alternativamente

$$\frac{\partial u(t, x)}{\partial t} + \frac{\partial [u(t, x) a(x)]}{\partial x} + \frac{1}{2} \frac{\partial^2 [u(t, x) (b(x))^2]}{\partial x^2} = 0 \quad (\text{D.48})$$

conhecida como Equação direta de Kolmogorov (forward equation), às vezes também chamada de equação de Fokker-Planck.

Os resultados que daremos agora podem ser encontrados com mais detalhes em

[Karatzas & Shreve, 91], [Oksendal, 92], [Gihman & Skorohod, 72], [Wong, 71] e [Gillespie, 92].

D.4.1 Relação entre EDP e EDE via Teoria dos Processos de Markov

A última referência é um estudo sobre os processos Markovianos. Nesta obra mostra-se como a partir da equação de Chapman-Kolmogorov

$$P(x, t | x_0, t_0) = \int_{-\infty}^{\infty} P(x, t | z, s) P(z, s | x_0, t_0) dz \quad (D.49)$$

com $\{t_0 \leq s \leq t\}$, chegamos as equações de Kolmogorov. Note que a equação acima pode ser reescrita como

$$P(x, t | x_0, t_0) = E[P(x, t | z, s) | \mathcal{F}_0] \quad (D.50)$$

Iremos utilizar a teoria dos processo markovianos para mostrar que podemos encontrar a solução da equação diferencial parabólica

$$\frac{\partial u}{\partial t} + Lu = 0 \quad (D.51)$$

onde L é o operador diferencial elíptico veja (D.3), resolvendo a equação diferencial estocástica análoga

$$dX_t = a(X_t) dt + b(X_t) dB_t \quad (D.52)$$

onde X_t é um processo de Itô e $a(X_t)$ e $b(X_t)$ satisfazem as condições

$$\begin{aligned} |a(t, x) - a(t, y)| + |b(t, x) - b(t, y)| &\leq K|x - y| \\ |a(t, x)|^2 + |b(t, x)|^2 &\leq K(1 + |x|^2) \end{aligned} \quad (D.53)$$

para algum K . Como veremos mais adiante, as equações de evolução temporal para $P(x, t | x_0, t_0)$ podem ser derivadas a partir de duas versões da equação de Chapman-Kolmogorov, obtidas pela renomeação das variáveis conforme a figura (D.54)

(D.54)

Estas duas equações de Chapman-Kolmogorov são as seguintes, do gráfico (D.54a), temos

$$P(x, t + \Delta t | x_0, t_0) = \int_{-\infty}^{\infty} P(x, t + \Delta t | x - \xi, t) P(x - \xi, t | x_0, t_0) d\xi \quad (D.55)$$

do gráfico (D.54b), temos

$$P(x, t | x_0, t_0) = \int_{-\infty}^{\infty} P(x, t | x_0 + \xi, t_0 + \Delta t_0) P(x_0 + \xi, t_0 + \Delta t_0 | x_0, t_0) d\xi \quad (D.56)$$

A equação (D.55) nos dá a equação que governa o comportamento de $P(x, t | x_0, t_0)$ para fixos x_0, t_0 , é a chamada equação de evolução temporal direta (forward equation). Já a equação (D.56) nos dá a equação que governa o comportamento de $P(x, t | x_0, t_0)$ para fixos x, t , é a chamada equação de evolução temporal reversa (backward equation).

Vamos começar mostrando em linhas gerais como, a partir da equação acima podemos mostrar que a densidade de transição de um processo $(X_t)_{t \in \mathbb{R}^+}$ satisfazendo uma equação diferencial estocástica

$$dX_t = a(X_t) dt + b(X_t) dB_t \quad (D.57)$$

satisfaz as equações (D.47) e (D.48). Aqui estamos considerando apenas o caso homogêneo no tempo, portanto $a(X_t)$ e $b(X_t)$ dependem somente de X_t .

A primeira etapa é verificar que o processo $(X_t)_{t \in \mathbb{R}^+}$ satisfaz a propriedade de Markov: O comportamento futuro do processo depende apenas do estado do processo no instante presente, isto é, o comportamento futuro do processo é independente do comportamento passado do mesmo. O teorema (7.2) em [Oksendal, 92], confirma o fato que X_t é um processo markoviano, portanto

$$P(x, t | \mathcal{F}_0) = P(x, t | x_0, t_0) \quad (D.58)$$

A partir deste fato podemos então utilizar a equação de Chapman-Kolmogorov (D.49) para descrever a evolução temporal do processo $(X_t)_{t \in \mathbb{R}^+}$. Vamos começar falando um pouco sobre o propagador de um processo de Markov.

Suponha que nosso processo está no estado x no tempo t , isto é, $X_t = x$. Então após um infinitésimo de tempo mais tarde $t + \Delta t$ o processo terá evoluído para algum estado que está deslocado de x por um valor

$$\Xi(dt; x, t) = X_{t+dt} - X_t, \text{ dado } X_t = x \quad (D.59)$$

Note que esta condição " $X_t = x$ " nesta definição afeta tanto o termo à direita como o termo à esquerda. Este deslocamento $\Xi(dt; x, t)$ do estado x durante o tempo $[t, t + dt]$ é claramente um variável aleatória. Aqui vamos utilizar a mesma denominação adotada por [Gillespie, 92], assim $\Xi(dt; x, t)$ será denominado o propagador do processo $(X_t)_{t \in \mathbb{R}^+}$. Semelhantemente a qualquer variável aleatória o propagador é completamente especificado pela sua função densidade. Vamos denotar a função densidade do propagador $\Xi(dt; x, t)$ por $P(\xi|dt, x, t)$ e referir-se a ele como a função densidade do propagador. Deste modo temos por definição que

$$P(\xi|dt, x, t) d\xi = P(x + \xi, t + dt|x, t) d\xi = \text{Prob}\{\Xi(dt; x, t) \in [\xi, \xi + d\xi]\} \quad (\text{D.60})$$

Evidentemente, o propagador $\Xi(dt; x, t)$ nos informa onde o processo no estado x no tempo t estará num infinitésimo de tempo mais tarde $t + \Delta t$. Especificamente, o processo estará no estado $x + \Xi(dt; x, t)$. Observe que da definição do propagador temos que

$$\Xi(dt; x, t) = X_{t+dt} - X_t = dX_t \quad (\text{D.61})$$

Temos ainda que os momentos de $\Xi(dt; x, t)$ tem a seguinte estrutura analítica

$$\langle \Xi^n(dt; x, t) \rangle = \int_{-\infty}^{\infty} \xi^n P(\xi|dt, x, t) d\xi = \mathbf{B}_n(x, t) dt + o(dt) \quad (\text{D.62})$$

considerando o deslocamento $\Xi(dt; x, t)$ do estado x em t igual a ξ .

onde $\mathbf{B}_1(x, t), \mathbf{B}_2(x, t), \dots$ são todas funções bem comportadas de x e t e

$$\frac{o(dt)}{dt} \rightarrow 0 \text{ quando } \Delta t \rightarrow 0 \quad (\text{D.63})$$

Assim, de (D.62) temos que

$$\mathbf{B}_n(x, t) = \frac{1}{dt} \int_{-\infty}^{\infty} \xi^n P(\xi|dt, x, t) d\xi, \quad n \in \mathbb{Z}^+ \quad (\text{D.64})$$

A partir deste conceitos básicos desenvolvidos aqui e mais a equação de Chapman-Kolmogorov, vamos mostrar a relação entre as equações diferenciais parabólicas e as equações diferenciais estocásticas.

Vamos primeiro considerar a seguinte função

$$f(x) = P(x + \xi, t + \Delta t|x, t) P(x, t|x_0, t_0) \quad (\text{D.65})$$

desta maneira, então podemos escrever o primeiro fator do integrando de (D.55) como $f(x - \xi)$. Assumindo que $f \in C_0^\infty$ e fazendo a expansão em séries de Taylor de $f(x - \xi)$ temos

$$f(x - \xi) = f(x) + \sum_{n=1}^{\infty} \frac{(-\xi)^n}{n!} \frac{\partial^n f(x)}{\partial x^n} \quad (D.66)$$

substituindo isto pelo primeiro fator do integrando na equação (D.55), temos

$$\begin{aligned} P(x, t + \Delta t | x_0, t_0) &= \int_{-\infty}^{\infty} P(x, t + \Delta t | x - \xi, t) P(x - \xi, t | x_0, t_0) d\xi \\ &= \int_{-\infty}^{\infty} P(x + \xi, t + \Delta t | x, t) P(x, t | x_0, t_0) d\xi + \\ &\quad + \sum_{n=1}^{\infty} \frac{(-1)^n}{n!} \frac{\partial^n}{\partial x^n} \int_{-\infty}^{\infty} \xi^n P(x + \xi, t + \Delta t | x, t) P(x, t | x_0, t_0) d\xi \\ &= P(x, t | x_0, t_0) \underbrace{\int_{-\infty}^{\infty} \xi^n P(x + \xi, t + \Delta t | x, t) d\xi}_{=1} + \\ &\quad + \sum_{n=1}^{\infty} \frac{(-1)^n}{n!} \frac{\partial^n}{\partial x^n} \left\{ \left[\int_{-\infty}^{\infty} \xi^n P(x + \xi, t + \Delta t | x, t) d\xi \right] P(x, t | x_0, t_0) \right\} \\ &= P(x, t | x_0, t_0) + \\ &\quad + \sum_{n=1}^{\infty} \frac{(-1)^n}{n!} \frac{\partial^n}{\partial x^n} \left\{ \left[\int_{-\infty}^{\infty} \xi^n P(x + \xi, t + \Delta t | x, t) d\xi \right] P(x, t | x_0, t_0) \right\} \end{aligned}$$

subtraindo $P(x, t | x_0, t_0)$ de ambos os lados e dividindo por Δt , temos que

$$\begin{aligned} \frac{P(x, t + \Delta t | x_0, t_0) - P(x, t | x_0, t_0)}{\Delta t} &= \sum_{n=1}^{\infty} \frac{(-1)^n}{n!} \frac{\partial^n}{\partial x^n} \times \\ &\quad \left\{ \left[\frac{1}{\Delta t} \int_{-\infty}^{\infty} \xi^n P(x + \xi, t + \Delta t | x, t) d\xi \right] \times \right. \\ &\quad \left. P(x, t | x_0, t_0) \right\} \end{aligned} \quad (D.67)$$

tomando o limite $\Delta t \rightarrow 0$

$$\frac{\partial P(x, t | x_0, t_0)}{\partial t} = \sum_{n=1}^{\infty} \frac{(-1)^n}{n!} \frac{\partial^n}{\partial x^n} [B_n(x, t) P(x, t | x_0, t_0)] \quad (D.68)$$

esta equação diferencial parcial de ordem infinita é conhecida como Equação direta (forward equation) de Kramers-Moyal. Se todas as funções $B_n(x, t)$ são conhecidas, isto é, se todos os momentos do Propagador do processo $\Xi(dt; x, t)$ são conhecidos, então a equação direta de Kramers-Moyal (D.68) constitui uma completa equação de evolução temporal do processo $(X_t)_{t \in \mathbb{R}^+}$ para fixos x_0 e t_0 .

A dedução da equação reversa de Kramers-Moyal inicia-se observando que se definirmos uma função h como

$$h(x_0) = P(x, t | x_0, t_0 + \Delta t_0) \quad (D.69)$$

então podemos escrever o primeiro fator do integrando de (D.56) como $h(x_0 + \xi)$. Assumindo que $h \in C_0^\infty$ e fazendo a expansão em séries de Taylor de $h(x_0 + \xi)$ temos que

$$h(x_0 + \xi) = h(x_0) + \sum_{n=1}^{\infty} \frac{\xi^n}{n!} \frac{\partial^n h(x_0)}{\partial x_0^n} \quad (D.70)$$

substituindo isto pelo primeiro fator do integrando em (D.56) temos que

$$\begin{aligned} P(x, t | x_0, t_0) &= \int_{-\infty}^{\infty} P(x, t | x_0 + \xi, t_0 + \Delta t_0) P(x_0 + \xi, t_0 + \Delta t_0 | x_0, t_0) d\xi \\ &= \int_{-\infty}^{\infty} P(x, t | x_0, t_0 + \Delta t_0) P(x_0 + \xi, t_0 + \Delta t_0 | x_0, t_0) d\xi + \\ &\quad \sum_{n=1}^{\infty} \frac{1}{n!} \left\{ \int_{-\infty}^{\infty} \xi^n \left[\frac{\partial^n}{\partial x_0^n} P(x, t | x_0, t_0 + \Delta t_0) \right] P(x_0 + \xi, t_0 + \Delta t_0 | x_0, t_0) d\xi \right\} \\ &= P(x, t | x_0, t_0 + \Delta t_0) \underbrace{\int_{-\infty}^{\infty} P(x_0 + \xi, t_0 + \Delta t_0 | x_0, t_0) d\xi}_{=1} + \\ &\quad \sum_{n=1}^{\infty} \frac{1}{n!} \left\{ \left[\frac{\partial^n}{\partial x_0^n} P(x, t | x_0, t_0 + \Delta t_0) \right] \int_{-\infty}^{\infty} \xi^n P(x_0 + \xi, t_0 + \Delta t_0 | x_0, t_0) d\xi \right\} \end{aligned}$$

subtraindo $P(x, t | x_0, t_0 + \Delta t_0)$ de ambos os lados e dividindo por Δt , temos que

$$\frac{P(x, t | x_0, t_0) - P(x, t | x_0, t_0 + \Delta t_0)}{\Delta t} = \sum_{n=1}^{\infty} \frac{1}{n!} \left[\frac{\partial^n}{\partial x_0^n} P(x, t | x_0, t_0 + \Delta t_0) \right] \times \frac{1}{\Delta t} \int_{-\infty}^{\infty} \xi^n P(x_0 + \xi, t_0 + \Delta t_0 | x_0, t_0) d\xi$$

tomando o limite $\Delta t \rightarrow 0$, temos que

$$-\frac{\partial P(x, t | x_0, t_0)}{\partial t} = \sum_{n=1}^{\infty} \frac{1}{n!} B_n(x_0, t_0) \frac{\partial^n}{\partial x_0^n} P(x, t | x_0, t_0) \quad (D.71)$$

Esta é a equação reversa (backward equation) de Kramers-Moyal.

Portanto as equações direta e reversa de Kramers-Moyal (D.68) e (D.71), nos dão todas as informações probabilísticas a respeito do processo $(X_t)_{t \in \mathbb{R}^+}$. Estas equações de evolução descrevem o fluxo da informação nas duas direções do tempo, a equação direta de Kramers-Moyal propagando para o futuro as condições presentes e a equação reversa de Kramers-Moyal propagando para o presente as condições de um valor esperado futuro.

A próxima etapa agora é determinar a densidade do propagador $\Xi(dt; x, t)$ que designamos por $P(\xi|dt, x, t)$.

Da equação (D.61) temos que

$$\Xi(dt; x, t) = X_{t+dt} - X_t = dX_t \quad (\text{D.72})$$

Ora, sabemos que o nosso processo é regido pela equação diferencial estocástica

$$dX_t = a(X_t) dt + b(X_t) dB_t \quad (\text{D.73})$$

onde B_t é o processo de Wiener, que tem como densidade

$$P(b, t) = \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{b^2}{2t}\right) \quad (\text{D.74})$$

portanto a densidade do propagador $\Xi(dt; x, t)$ é dado por

$$P(\xi|dt, x, t) = \frac{1}{\sqrt{2\pi [b(X_t)]^2 dt}} \exp\left\{-\frac{[\xi - a(X_t) dt]^2}{2 [b(X_t)]^2 dt}\right\} \quad (\text{D.75})$$

Deste modo, a função geradora dos momentos do propagador é dada por

$$\mathbf{M}_\xi(s) = \exp\left[a(X_t) dt s + \frac{[b(X_t)]^2 dt}{2} s^2\right] \quad (\text{D.76})$$

assim temos

$$\mathbf{B}_n(x, t) = \frac{1}{dt} \int_{-\infty}^{\infty} \xi^n P(\xi|dt, x, t) d\xi = \frac{1}{dt} \left[\frac{\partial^n}{\partial s^n} \mathbf{M}_\xi(s) \right]_{s=0} \quad (\text{D.77})$$

portanto temos para $n = 1$

$$\begin{aligned} \mathbf{B}_1(x, t) &= \frac{1}{dt} \left[\left(a(X_t) dt + \underbrace{[b(X_t)]^2 dt}_{=0} \right) \underbrace{\mathbf{M}_\xi(s)}_{=1} \right]_{s=0} \\ &= \frac{1}{dt} a(X_t) dt \end{aligned} \quad (\text{D.78})$$

$$\therefore \mathbf{B}_1(x, t) = a(X_t) \quad (\text{D.79})$$

para $n = 2$

$$\begin{aligned}
B_2(x, t) &= \frac{1}{dt} \left[\underbrace{([b(X_t)]^2 dt)}_{=1} \underbrace{M_\xi(s)}_{=1} + \left(a(X_t) dt + \underbrace{[b(X_t)]^2 dts}_{=0} \right)^2 \underbrace{M_\xi(s)}_{=1} \right]_{s=0} \\
&= \frac{1}{dt} [[b(X_t)]^2 dt + [a(X_t)]^2 (dt)^2] \\
&= [b(X_t)]^2
\end{aligned} \tag{D.80}$$

para $n = 3$

$$\begin{aligned}
B_3(x, t) dt &= ([b(X_t)]^2 dt) \left(a(X_t) dt + \underbrace{[b(X_t)]^2 dts}_{=0} \right) \underbrace{M_\xi(s)}_{=1} + \\
&\quad 2 \left(a(X_t) dt + \underbrace{[b(X_t)]^2 dts}_{=0} \right) ([b(X_t)]^2 dt) \underbrace{M_\xi(s)}_{=1} + \\
&\quad \left(a(X_t) dt + \underbrace{[b(X_t)]^2 dts}_{=0} \right)^3 \underbrace{M_\xi(s)}_{=1} \\
&= a(X_t) [b(X_t)]^2 (dt)^2 + 2a(X_t) [b(X_t)]^2 (dt)^2 + a(X_t) (dt)^3 \tag{D.81}
\end{aligned}$$

$$B_3(x, t) = \frac{o(dt)}{dt} = 0 \tag{D.82}$$

Se continuarmos a calcular os elementos seguintes $B_4(x, t)$, $B_5(x, t)$, $\dots$ verificaremos que

$$B_n(x, t) = \begin{cases} a(X_t), & \text{para } n = 1 \\ [b(X_t)]^2, & \text{para } n = 2 \\ 0, & \text{para } n \geq 3 \end{cases} \tag{D.83}$$

Deste modo as equações direta e reversa de Kramers-Moyal (D.68) e (D.71) tomam a seguinte forma

$$\frac{\partial}{\partial t} P(x, t | x_0, t_0) = \frac{\partial}{\partial x} [a(x) P(x, t | x_0, t_0)] + \frac{1}{2} \frac{\partial}{\partial x} [[b(x)]^2 P(x, t | x_0, t_0)] \tag{D.84}$$

e

$$-\frac{\partial}{\partial t_0} P(x, t | x_0, t_0) = a(x_0) \frac{\partial}{\partial x_0} P(x, t | x_0, t_0) + \frac{[b(x_0)]^2}{2} \frac{\partial^2}{\partial x_0^2} P(x, t | x_0, t_0) \tag{D.85}$$

A equação (D.84) é conhecida como equação direta (forward equation) de Kolmogorov ou equação direta de Fokker-Planck-Kolmogorov (F-P-K) e a equação (D.85) como equação

reversa (backward equation) de kolmogorov ou equação reversa de Fokker-Planck-Kolmogorov (F-P-K).

Uma questão que pode ser levantada agora é como garantir que $P(x, t | x_0, t_0)$ é diferenciável, isto é,

$$P(x, t | x_0, t_0) \in C_0^2(\mathbb{R}^+ \times \mathbb{R}) \quad (D.86)$$

Esta condição é satisfeita se $a(X_t) = a(t, x)$ e $b(X_t) = b(t, x)$ satisfizerem as seguintes condições, veja [Wong, 71] ou [Gihman & Skorohod, 72] por exemplo:

Existe uma constante K tal que para $t \in \mathbb{R}^+$ e $x, y \in \mathbb{R}$, temos

$$\begin{aligned} |a(t, x) - a(t, y)| + |b(t, x) - b(t, y)| &\leq K|x - y| \\ |a(t, x)|^2 + |b(t, x)|^2 &\leq K(1 + |x|^2) \end{aligned} \quad (D.87)$$

portanto satisfeitas estas condições existe $P(x, t | x_0, t_0)$ que é solução de (D.84) e (D.85).

Portanto mostramos a afirmação que queríamos: Dado uma equação diferencial parabólica podemos encontrar sua solução resolvendo a equação diferencial estocástica análoga e vice-versa.

D.4.2 Relação entre EDP e EDE via Cálculo Estocástico

É interessante em muitas aplicações que possamos associar um operador diferencial $\mathcal{A}$ a um processo estocástico $(X_t)_{t \in \mathbb{R}^+}$. A conexão entre $\mathcal{A}$ e $(X_t)_{t \in \mathbb{R}^+}$ é que $\mathcal{A}$ é o gerador infinitesimal de $(X_t)_{t \in \mathbb{R}^+}$.

Definição D.4 (Gerador Infinitesimal) *Seja $(X_t)_{t \in \mathbb{R}^+}$ um processo de difusão de Itô homogêneo no tempo. O gerador infinitesimal $\mathcal{A}$ de $(X_t)_{t \in \mathbb{R}^+}$ é definido por*

$$\begin{aligned} \mathcal{A}f(X_t) &= \lim_{\Delta t \rightarrow 0} \frac{E[f(X_{t+\Delta t}) | \mathcal{F}_t] - E[f(X_t) | \mathcal{F}_t]}{\Delta t} \\ &= \lim_{\Delta t \rightarrow 0} \frac{E[f(X_{t+\Delta t}) | \mathcal{F}_t] - f(X_t)}{\Delta t} \end{aligned}$$

$X_t = x; x \in \mathbb{R}$.

O conjunto de funções $f : \mathbb{R} \rightarrow \mathbb{R}$ tal que o limite existe em x é denotado por $\mathcal{D}_{\mathcal{A}}(x)$. Enquanto $\mathcal{D}_{\mathcal{A}}$ denota o conjunto de funções para o qual o limite existe para todo $x \in \mathbb{R}$.

Mostremos agora o resultado que associa o operador $\mathcal{A}$ com o operador diferencial elíptico L , onde L é dado pela equação (D.3).

Teorema D.3 (Gerador Infinitesimal de um Processo de Difusão) *Seja L o operador diferencial elíptico em $\mathbb{R}$*

$$L = A(x) \frac{\partial}{\partial x} + B(x) \frac{\partial^2}{\partial x^2}$$

E o processo de difusão de Itô $(X_t)_{t \in \mathbb{R}^+}$ que é solução da EDE

$$dX_t = a(X_t) dt + b(X_t) dB_t$$

com

$$\begin{aligned} |a(t, x) - a(t, y)| + |b(t, x) - b(t, y)| &\leq K|x - y| \\ |a(t, x)|^2 + |b(t, x)|^2 &\leq K(1 + |x|^2) \end{aligned}$$

para alguma constante K . Então

$$\mathcal{A}f(X_t) = Lf(X_t)$$

Prova: *Considere uma função f que dependa somente de X_t , assim, pelo lema de Itô (D.2), temos que*

$$\begin{aligned} df(X_t) &= \left[\underbrace{\frac{\partial f}{\partial t}}_{=0} + a(x) \frac{\partial f}{\partial x} + \frac{[b(x)]^2}{2} \frac{\partial^2 f}{\partial x^2} \right] dt + b(x) \frac{\partial f}{\partial x} dB_t \\ &= \left[a(x) \frac{\partial f}{\partial x} + \frac{[b(x)]^2}{2} \frac{\partial^2 f}{\partial x^2} \right] dt + b(x) \frac{\partial f}{\partial x} dB_t \\ &= (Lf) dt + b(x) \frac{\partial f}{\partial x} dB_t \end{aligned}$$

onde

$$A(x) = a(x) \text{ e } B(x) = \frac{[b(x)]^2}{2}$$

integrando de t a $t + \Delta t$ e tomando o valor esperado em ambos os lados, temos que

$$\begin{aligned} E[f(X_{t+\Delta t}) | \mathcal{F}_t] - E[f(X_t) | \mathcal{F}_t] &= E \left[\int_t^{t+\Delta t} (Lf) ds \middle| \mathcal{F}_t \right] + \\ &\quad \underbrace{E \left[\int_t^{t+\Delta t} b(x) \frac{\partial f}{\partial x} dB_s \middle| \mathcal{F}_t \right]}_{=0} \\ E[f(X_{t+\Delta t}) | \mathcal{F}_t] - f(X_t) &= E \left[\int_t^{t+\Delta t} (Lf) ds \middle| \mathcal{F}_t \right] \end{aligned}$$

dividindo ambos os lados por Δt , temos

$$\frac{E[f(X_{t+\Delta t}) | \mathcal{F}_t] - f(X_t)}{\Delta t} = \frac{1}{\Delta t} E \left[\int_t^{t+\Delta t} (Lf) ds \middle| \mathcal{F}_t \right]$$

e tomando o limite

$$\lim_{\Delta t \rightarrow 0} \frac{E[f(X_{t+\Delta t})|\mathcal{F}_t] - f(X_t)}{\Delta t} = Lf(t, x) \blacksquare$$

O próximo teorema completa a conexão entre EDE's e EDP's. Este é a conhecida fórmula de Feynman-Kac.

Teorema D.4 (Fórmula de Feynman-Kac) *Seja $F(t, x) = F(X_t)$ um martingale com X_t satisfazendo a EDE*

$$dX_t = a(X_t)dt + b(X_t)dB_t$$

com

$$\begin{aligned} |a(t, x) - a(t, y)| + |b(t, x) - b(t, y)| &\leq K|x - y| \\ |a(t, x)|^2 + |b(t, x)|^2 &\leq K(1 + |x|^2) \end{aligned}$$

para alguma constante K . Então $F(t, x)$ é a solução da equação diferencial parcial

$$\frac{\partial F(t, x)}{\partial t} + a(x) \frac{\partial F(t, x)}{\partial x} + \frac{1}{2} [b(x)]^2 \frac{\partial^2 F(t, x)}{\partial x^2} = 0$$

com condição final $F(T, y) = h(X_T)$ e tem a seguinte representação estocástica

$$F(t, x) = E[h(X_T) | X_t = x]$$

Portanto a fórmula de Feynman-Kac nos fornece uma representação estocástica para a solução de uma EDP.

Prova: Pelo lema de Itô (D.2), temos que

$$dF(t, x) = \left[\frac{\partial F(t, x)}{\partial t} + a(x) \frac{\partial F(t, x)}{\partial x} + \frac{1}{2} [b(x)]^2 \frac{\partial^2 F(t, x)}{\partial x^2} \right] dt + b(x) \frac{\partial F(t, x)}{\partial x} dB_t$$

como $F(t, x)$ é um martingale então

$$\frac{\partial F(t, x)}{\partial t} + a(x) \frac{\partial F(t, x)}{\partial x} + \frac{1}{2} [b(x)]^2 \frac{\partial^2 F(t, x)}{\partial x^2} = 0$$

assim

$$dF(t, x) = b(x) \frac{\partial F(t, x)}{\partial x} dB_t$$

integrando

$$\begin{aligned} \int_t^T dF(X_s) &= \int_t^T b(X_s) \frac{\partial F(X_s)}{\partial X_s} dB_s \\ F(T, y) &= F(t, x) + \int_t^T b(X_s) \frac{\partial F(X_s)}{\partial X_s} dB_s \end{aligned}$$

e calculando o valor esperado com a condição final $F(T, y) = h(X_T)$ temos que

$$E[h(X_T) | X_t = x] = E[F(t, x) | X_t = x] + \underbrace{E\left[\int_t^T b(X_s) \frac{\partial F(X_s)}{\partial X_s} dB_s \mid X_t = x\right]}_{=0}$$

assim

$$F(t, x) = E[h(X_T) | X_t = x] \blacksquare$$

Este teorema é chave na solução do problema da precificação de opções.

Uma das etapas na construção da solução do problema da precificação de opções é a determinação da medida martingale equivalente (MME), pois é através dela que determinamos o preço justo ou preço do risco neutro. A ferramenta necessária para se determinar esta MME é o teorema da transformação de medida.

D.5 Transformações de Medida: A Formula de Cameron-Martin-Girsanov

Os primeiros trabalhos sobre transformações de medida se devem a Cameron e Martin veja, e [Cameron & Martin, 49]. No primeiro artigo [Cameron & Martin, 44], foi mostrado como a integral de Wiener é transformada sobre translações; no segundo artigo [Cameron & Martin, 45], foi mostrado como a integral de Wiener é transformada para uma classe de transformações lineares homogêneas, no mesmo artigo foi combinado o resultado das transformações lineares com a translação do artigo anterior dando assim uma fórmula para transformações lineares não homogêneas; no terceiro artigo é mostrado o resultado para transformações não lineares. Todos estes trabalhos foram feitos a partir do conceito da integral de Wiener, veja [Yeh, 73], esta referência é interessante pois contém vários resultados dados por Wiener.

Alguns anos mais tarde foi publicado um trabalho dando resultados semelhantes aos de Cameron e Martin, veja [Girsanov, 60]. Os resultados conseguidos por Girsanov tiveram como base a teoria desenvolvida por Itô, veja [Itô, 44] e [Itô, 46]. Os resultados conseguidos por Girsanov são mais gerais que os de Cameron e Martin, talvez por isso alguns autores quando citam o teorema da transformação de medida citem apenas Girsanov.

Antes de darmos este importante resultado, comecemos com outro que é base pra este, o teorema de Radon-Nikodym, para a prova deste resultado veja por exemplo [Billingsley, 79], [Doob, 94] ou [Halmos, 88]. As duas últimas referências tratam de medidas quaisquer a diferença é que na medida de probabilidade temos $\mathbb{P}(\Omega) = 1$. No entanto qualquer medida pode ser normalizada desde que $\mathbb{P}(\Omega) < \infty$.

Teorema (Radon-Nikodym) D.5 *Seja $\mathbb{P}$ e $\mathbb{Q}$ duas medidas de probabilidade sobre o espaço $(\Omega, \mathcal{F})$. Se para qualquer $A \in \mathcal{F}$ satisfazendo $\mathbb{P}(A) = 0$ tivermos também $\mathbb{Q}(A) = 0$, dizemos que $\mathbb{Q}$ é absolutamente contínua com relação a $\mathbb{P}$. Sobre estas condições, existe uma variável aleatória não negativa tal que*

$$\mathbb{Q}(A) = \int_A Z \mathbb{P}, \text{ qualquer } A \in \mathcal{F}$$

e Z é chamada a derivada de Radon-Nikodym de $\mathbb{Q}$ com respeito a $\mathbb{P}$, ou a densidade com respeito a $\mathbb{P}$ da medida $\mathbb{Q}$. A equação acima também é representada da seguinte maneira

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = Z$$

Corolário D.1 *Seja $\mathbb{Q}$ uma medida absolutamente contínua com relação a $\mathbb{P}$. Se Z é a derivada de Radon-Nikodym de $\mathbb{Q}$ com respeito a $\mathbb{P}$ então*

$$E_{\mathbb{Q}}(X) = E_{\mathbb{P}}(XZ)$$

para qualquer variável aleatória X para o qual $E_{\mathbb{P}}(XZ) < \infty$.

Corolário D.2 *Se $\mathbb{Q}$ é absolutamente contínua com relação a $\mathbb{P}$ e $\mathbb{P}$ é absolutamente contínua com relação a $\mathbb{Q}$, dizemos que $\mathbb{P}$ e $\mathbb{Q}$ são equivalentes, então Z é a derivada de Radon-Nikodym de $\mathbb{Q}$ com respeito a $\mathbb{P}$ e $\frac{1}{Z}$ é a derivada de Radon-Nikodym de $\mathbb{P}$ com respeito a $\mathbb{Q}$. O próximo resultado será importante na demonstração da fórmula de Cameron-Martin-Girsanov.*

Lema D.3 *$X_t Z_t$ é um martingale (local) sobre $\mathbb{P}$ se e somente se X_t é um martingale (local) sobre $\mathbb{Q}$, onde $Z_t = \frac{d\mathbb{Q}}{d\mathbb{P}}|_{\mathcal{F}_t}$.*

Prova: *Daremos apenas a prova para o caso de martingales. A extensão para martingales locais pode ser encontrada em [Jacob & Shiryaev, 87]. Vamos começar pela primeira implicação: X_t martingale sobre $\mathbb{Q} \Rightarrow X_t Z_t$ martingale sobre $\mathbb{P}$.*

suponha $s \leq t$ e $A \in \mathcal{F}_s$, então

$$\int_A X_t Z_t d\mathbb{P} = \int_A X_t d\mathbb{Q} = \int_A X_s d\mathbb{Q} = \int_A X_s Z_s d\mathbb{P}$$

agora mostremos que: $X_t Z_t$ martingale sobre $\mathbb{P} \Rightarrow X_t$ martingale sobre $\mathbb{Q}$.

$$\int_A X_t d\mathbb{Q} = \int_A X_t Z_t d\mathbb{P} = \int_A X_s Z_s d\mathbb{P} = \int_A X_s d\mathbb{Q} \blacksquare$$

Daremos a seguir o principal resultado desta seção, a fórmula de Cameron-Martin-Girsanov.

Teorema (Fórmula de Cameron-Martin-Girsanov) D.6 Seja $(B_t)_{t \in [0, T]}$, um processo de Wiener definido sobre a base estocástica $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ e seja $(\theta_t)_{t \in [0, T]}$ um processo $\mathcal{F}_t$ -adaptado, tal que

$$\int_0^T (\theta_s)^2 ds < \infty$$

e também tal que o processo

$$Z_t = \exp \left\{ - \int_0^t \theta_s dB_s - \frac{1}{2} \int_0^t (\theta_s)^2 ds \right\}$$

é um martingale sobre $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$. Defina uma nova medida de probabilidade $\mathbb{Q}$ sobre $(\Omega, \mathcal{F}, \mathbb{F})$ da seguinte maneira.

$$\left. \frac{d\mathbb{Q}}{d\mathbb{P}} \right|_{\mathcal{F}_T} = Z_T$$

então o processo

$$B_t^* = \int_0^t \theta_s ds + B_t$$

é um processo de Wiener sobre esta nova medida.

Prova:

1. Z_t é um martingale, de fato, usando o lema de Itô (D.2) e notando

$$dY_t = -\theta_t dB_t - \frac{1}{2} (\theta_t)^2 dt, Y_0 = 0$$

então

$$Z_t = e^{Y_t}, Z_0 = 1$$

portanto pelo lema de Itô, temos que

$$\begin{aligned} dZ_t &= \frac{\partial Z}{\partial Y} dY + \frac{1}{2} \frac{\partial^2 Z}{\partial Y^2} (dY)^2 \\ &= Z_t \left[-\theta_t dB_t - \frac{1}{2} (\theta_t)^2 dt \right] + \frac{1}{2} Z_t \left[-\theta_t dB_t - \frac{1}{2} (\theta_t)^2 dt \right]^2 \\ &= -Z_t \theta_t dB_t - \frac{1}{2} Z_t (\theta_t)^2 dt + \frac{1}{2} Z_t (\theta_t)^2 dt \\ &= -Z_t \theta_t dB_t \end{aligned}$$

portanto Z_t é um martingale.

2. $\mathbb{Q}$ é uma medida de probabilidade. De fato, desde que

$$dZ_t = Z_t \theta_t dB_t, \text{ com } Z_0 = 1$$

então

$$Z_t = Z_0 - \int_0^t Z_s \theta_s dB_s = 1 - \int_0^t Z_s \theta_s dB_s$$

tomando o valor esperado temos

$$E_{\mathbb{P}}(Z_t) = 1 - E \left[\int_0^t Z_s \theta_s dB_s \right] = 0$$

para qualquer $t \in [0, T]$, em particular para $t = T$, temos que

$$E_{\mathbb{P}}(Z_T) = \int_{\Omega} Z_T d\mathbb{P} = \mathbb{Q}(\Omega) = 1$$

portanto $\mathbb{Q}$ é uma medida de probabilidade.

3. Para mostra que B_t^* é um processo de Wiener, será utilizado a caracterização do processo de Wiener dada por Lévy, veja teorema (D.2).

- (a) Primeiro por definição B_t^* é um processo é contínuo q.s., pois B_t é contínuo q.s. e

$\int_0^t \theta_s ds$ é um processo contínuo q.s..

- (b) Para mostrar que B_t^* é um martingale local sobre $\mathbb{Q}$ basta pelo lema (D.3), mostrar

que $\{Z_t B_t^*\}$ é um martingale local. Primeiro note que, utilizando o lema de Itô

podemos obter uma fórmula geral de integração por partes

$$\int_0^t X_s dY_s = \int_0^t d(X_s Y_s) - \int_0^t Y_s dX_s - \int_0^t dX_s dY_s$$

portanto

$$X_t Y_t = X_0 Y_0 + \int_0^t X_s dY_s + \int_0^t Y_s dX_s + \int_0^t dX_s dY_s$$

aplicando isto em $Z_t B_t^*$ vem que

$$\begin{aligned} Z_t \tilde{B}_t &= Z_0 B_0^* + \int_0^t Z_s dB_s^* + \int_0^t B_s^* dZ_s + \int_0^t dZ_s dB_s^* \\ &= B_0^* + \int_0^t Z_s (dB_s + \theta_s ds) + \int_0^t B_s^* (-Z_s \theta_s dB_s) + \\ &\quad \int_0^t (-Z_s \theta_s dB_s) (dB_s + \theta_s ds) \\ &= B_0^* + \int_0^t Z_s dB_s + \int_0^t Z_s \theta_s ds - \int_0^t B_s^* Z_s \theta_s dB_s \\ &\quad - \int_0^t Z_s \theta_s ds \\ &= B_0^* + \int_0^t Z_s dB_s - \int_0^t B_s^* Z_s \theta_s dB_s \\ &= B_0^* + \int_0^t Z_s (1 - B_s^* \theta_s) dB_s \end{aligned}$$

e como a integral estocástica com relação a B_t é um martingale local, segue que

$\{Z_t B_t^*\}$ é um martingale local e pelo lema (D.3) segue o resultado.

- (c) Para finalizar precisamos mostrar que $(B_t^*)^2 - t$ é um martingale local. Para ver isto basta utilizar o lema de Itô para $f(B_t^*) = (B_t^*)^2$

$$\begin{aligned} d((B_t^*)^2) &= 2B_t^*dB_t^* + \frac{1}{2}2(dB_t^*)^2 \\ &= 2B_t^*dB_t^* + dt \end{aligned}$$

$$(B_t^*)^2 - t = 2 \int_0^t B_s^* dB_s^*$$

foi mostrado que B_t^* é um mrtingale local com relação a medida $\mathbb{Q}$,

conseqüentemente, a integral estocástica é um martingale local sobre $\mathbb{Q}$ e portanto

o resultado segue. ■